

## APPENDIX E: FCHECK RESULTS LOOSE DEFINITION OF ANY

### SUCCESSFUL PARSES

['Arabic', 'Total', 'speakers', 'Approx', '.', '340', 'million', 'native', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Approx', '.', 'NUM', 'million', 'native', 'speakers']], ['NUM', '340']], 9, True)

['Arabic', 'Dialects', 'Western', '(', 'Maghrebi', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Arabic', 'Dialects', 'Central', '(', 'incl', '.', 'Egyptian', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'incl', '.', 'Type: ANY', '\\)']], 8, True)

['Arabic', 'Dialects', 'Northern', '(', 'incl', '.', 'Levantine', ',', 'Iraqi', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'incl', '.', 'Type: ANY', ',', 'Type: ANY', '\\)']], 10, True)

['Arabic', 'Dialects', 'Southern', '(', 'incl', '.', 'Gulf', ',', 'Hejazi', ',', 'Yemeni', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'incl', '.', 'Type: ANY', ',', 'Type: ANY', ',', 'Type: ANY', '\\)']], 12, True)

['Arabic', 'Spoken', 'in', 'liturgical', 'language', 'of', 'Islam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'liturgical', 'language', 'of', 'Islam']], 7, True)

['Arabic', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Arabic', 'Writing', 'system', 'Syriac', 'alphabet', '(', 'Garshuni', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Arabic', 'Writing', 'system', 'Latin', 'Alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Arabic', 'Pronunciation', '\xc3', '\xa6', 'l', '\xca', '\x95', '\xc9', '\x91', 'r', '\xc9', '\x91', '\xcb', '\x88', 'bijj', '\xc9', '\x90']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc3', '\xa6', 'l', '\xca', '\x95', '\xc9', '\x91', 'r', '\xc9', '\x91', '\xcb', '\x88', 'bijj', '\xc9', '\x90']], 17, True)

['Arabic', 'Official', 'language', 'in', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Bahrain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Comoros']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Djibouti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Arabic', 'Official', 'language', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Kuwait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Libya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Mauritania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Palestine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Qatar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Arabic', 'Official', 'language', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Southern', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Arabic', 'Official', 'language', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'United', 'Arab', 'Emirates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'United', 'Arab', 'Emirates']]], 7, True)

['Arabic', 'Official', 'language', 'in', 'Western', 'Sahara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Arabic', 'Official', 'language', 'in', 'Yemen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'African', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Arabic', 'Official', 'language', 'in', 'Arab', 'League']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Arabic', 'Official', 'language', 'in', 'OIC']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Arabic', 'Official', 'language', 'in', 'United', 'Nations']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Arabic', 'Regulated', 'by', 'Jordan', ':', 'Jordan', 'Academy', 'of', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',

'Jordan', ':', 'Jordan', 'Academy', 'of', 'Type: ANY']]]], 9, True)

['Arabic', 'Regulated', 'by', 'Syria', ':', 'Arab', 'Academy', 'of', 'Damascus', '(', 'the', 'oldest', ')']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Syria', ':', 'Arab', 'Academy', 'of', 'Damascus', '\\(', 'the', 'oldest', '\\)']]]], 13, True)

['Arabic', 'Regulated', 'by', 'Tunisia', ':', 'Beit', 'Al-Hikma', 'Foundation']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Tunisia', ':', 'Beit', 'Al-Hikma', 'Foundation']]]], 8, True)

['Arabic', 'Language', 'codes', 'ISO', '639-2', 'ara']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Arabic', 'Language', 'codes', 'ISO', '639-1', 'ar']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Austrian', 'Spoken', 'in', 'Austria']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Austrian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['American', 'Region', 'Anglophone', 'North', 'America']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 5, True)

['American', 'Signed', 'in', 'United', 'States', ',', 'Canada']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'United', 'States', ',', 'Canada']]]], 7, True)

['American', 'Total', 'signers', 'unknown', ',', 'perhaps', '500', ',', '000', 'or', 'so', 'primary', 'users']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'unknown', ',', 'perhaps', 'NUM', ',', 'NUM', 'or', 'so', 'primary', 'users']]], ['NUM', '500'], ['NUM', '000']], 13, True)

['American', 'Dialects', 'Black', 'ASL']]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['American', 'Dialects', 'Bolivian', 'Sign']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afrikaans', 'Total', 'speakers', 'over', '7', 'million', '(', 'native', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'over', 'NUM', 'million', '\\\\(', 'native', '\\\\)']], ['NUM', '7']], 9, True)

['Afrikaans', 'Total', 'speakers', 'between', '15', '\\xe2', '\\x80', '\\x93', '23', 'million', '(', 'total', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'between', 'NUM', '\\xe2', '\\x80', '\\x93', 'NUM', 'million', '\\\\(', 'total', '\\\\)']], ['NUM', '15'], ['NUM', '23']], 13, True)

['Afrikaans', 'Pronunciation', '\\xc9', '\\x90', 'fri', '\\xcb', '\\x88', 'k', '\\xc9', '\\x91', '\\xcb', '\\x90', 'ns']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xc9', '\\x90', 'fri', '\\xcb', '\\x88', 'k', '\\xc9', '\\x91', '\\xcb', '\\x90', 'ns']], 13, True)

['Afrikaans', 'Region', 'Southern', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afrikaans', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Afrikaans', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afrikaans', 'Spoken', 'in', 'Elsewhere', 'in', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Elsewhere', 'in', 'Africa']], 6, True)

['Afrikaans', 'Spoken', 'in', 'notably', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Afrikaans', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afrikaans', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Afrikaans', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Afrikaans', 'Spoken', 'in', 'Lesotho', 'and', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Afrikaans', 'Spoken', 'in', 'Emigrant', 'and', 'expatriate',  
'communities', 'worldwide']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Emigrant', 'and', 'expatriate', 'communities', 'worldwide']]], 8, True)

['Afrikaans', 'Spoken', 'in', 'notably', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Afrikaans', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Afrikaans', 'Spoken', 'in', 'the', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the',  
'United', 'Kingdom']]], 6, True)

['Afrikaans', 'Spoken', 'in', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the',  
'United', 'States']]], 6, True)

['Afrikaans', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Afrikaans', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Afrikaans', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Afrikaans', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Afrikaans', 'Spoken', 'in', 'the', 'Netherlands', 'and', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the', 'Netherlands', 'and', 'Belgium']]], 7, True)

['Afrikaans', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Afrikaans', 'Regulated', 'by', 'Die', 'Taalkommissie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Afrikaans', 'Language', 'codes', 'ISO', '639-2', 'afr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Afrikaans', 'Language', 'codes', 'ISO', '639-3', 'afr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Afrikaans', 'Language', 'codes', 'Linguasphere', '52-ACB-ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Afrikaans', 'Language', 'codes', 'ISO', '639-1', 'af']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Armenian', 'Spoken', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Nagorno-Karabakh', '(', 'not', 'recognized', 'internationally', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nagorno-Karabakh', '\\\\(', 'not', 'recognized', 'internationally', '\\\\)']]], 9, True)

['Armenian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'United', 'States']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Armenian', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Armenian', 'Writing', 'system', 'Armenian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Armenian', 'Pronunciation', 'h', '\xc9', '\x91', 'j', '\xc9', '\x9b', '\xcb', '\x88', '\xc9', '\xbe', '\xc9', '\x9b', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'h', '\xc9', '\x91', 'j', '\xc9', '\x9b', '\xcb', '\x88', '\xc9', '\xbe', '\xc9', '\x9b', 'n']], 15, True)

['Armenian', 'Official', 'language', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Official', 'language', 'in', 'Nagorno-Karabakh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Official', 'language', 'in', '(', 'not', 'recognized', 'internationally', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', '\\(', 'not', 'recognized', 'internationally', '\\)']], 9, True)

['Armenian', 'Official', 'language', 'in', 'Minority', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Armenian', 'Official', 'language', 'in', ':', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Armenian', 'Official', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences', 'of', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences', 'of', 'Armenia']]], 9, True)

['Armenian', 'Language', 'codes', 'ISO', '639-2', 'arm', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'arm', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Armenian', 'Language', 'codes', 'ISO', '639-2', 'hye', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hye', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Armenian', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Armenian', 'Language', 'codes', 'ISO', '639-1', 'hy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Aramaic', 'Spoken', 'in', 'Assyrian', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Aramaic', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Palestine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Aramaic', 'Writing', 'system', 'Aramaic', 'abjad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aramaic', 'Writing', 'system', 'Syriac', 'abjad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aramaic', 'Writing', 'system', 'Hebrew', 'abjad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aramaic', 'Writing', 'system', 'Mandaic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aramaic', 'Pronunciation', '\xc9', '\x91', 'r', '\xc9', '\x91', 'm',  
'\xc9', '\x91', 'j', '\xc9', '\x91', ',', '\xc9', '\x94', 'r', '\xc9',  
'\x94', 'm', '\xc9', '\x94', 'j', '\xc9', '\x94']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',  
'\x91', 'r', '\xc9', '\x91', 'm', '\xc9', '\x91', 'j', '\xc9', '\x91', ',',  
'\xc9', '\x94', 'r', '\xc9', '\x94', 'm', '\xc9', '\x94', 'j', '\xc9',  
'\x94']]], 25, True)
```

```
['Aramaic', 'Pronunciation', 'ar', '\xc9', '\x91', 'mi', '\xce', '\xb8']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ar',  
'\xc9', '\x91', 'mi', '\xce', '\xb8']]], 8, True)
```

```
['Aramaic', 'Pronunciation', 'ar', '\xc9', '\x91', 'mit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ar',  
'\xc9', '\x91', 'mit']]], 6, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'hrt', '\xe2', '\x80',  
'\x93', 'H', '\xc3', '\xa9', 'rtevin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hrt', '\xe2', '\x80', '\x93', 'H', '\xc3', '\xa9',  
'rtevin']], ['NUM', '639-3']], 13, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'huy', '\xe2', '\x80',  
'\x93', 'Hulaul', '\xc3', '\xa1']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

['ISO', 'NUM', 'huy', '\xe2', '\x80', '\x93', 'Hulaul', '\xc3', '\xa1']],  
['NUM', '639-3']], 12, True)

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'lhs', '\xe2', '\x80',  
\x93', 'Mlahs', '\xc3', '\xb4']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lhs', '\xe2', '\x80', '\x93', 'Mlahs', '\xc3', '\xb4']],  
['NUM', '639-3']], 12, True)

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'syn', '\xe2', '\x80',  
\x93', 'Senaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'syn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'trg', '\xe2', '\x80',  
\x93', 'Lish', '\xc3', '\xa1', 'n', 'Did', '\xc3', '\xa1', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'trg', '\xe2', '\x80', '\x93', 'Lish', '\xc3', '\xa1', 'n',  
'Did', '\xc3', '\xa1', 'n']], ['NUM', '639-3']], 17, True)

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'tru', '\xe2', '\x80',  
\x93', 'Turoyo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tru', '\xe2', '\x80', '\x93', 'Turoyo']], ['NUM', '639-3']],  
10, True)

['Aramaic', 'Language', 'codes', 'Linguasphere', '12-AAA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Aragonese', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aragonese', 'Region', 'Aragon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Aragonese', 'Language', 'codes', 'ISO', '639-2', 'arg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Aragonese', 'Language', 'codes', 'ISO', '639-3', 'arg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Aragonese', 'Language', 'codes', 'Linguasphere', '51-AAA-d']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Aragonese', 'Language', 'codes', 'ISO', '639-1', 'an']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Basque', 'Total', 'speakers', '665', ',', '800', 'in', '2006']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '665'], ['NUM', '800']], 8,  
True)
```

```
['Basque', 'Spoken', 'in', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Basque', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Basque', 'Region', 'Basque', 'Country']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Basque', 'Region', 'Navarre']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Basque', 'Region', 'French', 'Basque', 'Country']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Basque', 'Official', 'language', 'in', 'Basque', 'Country']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Basque', 'Official', 'language', 'in', 'Navarre']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Basque', 'Regulated', 'by', 'Euskaltzaindia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Basque', 'Language', 'codes', 'ISO', '639-2', 'baq', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'baq', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Basque', 'Language', 'codes', 'ISO', '639-2', 'eus', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'eus', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Basque', 'Language', 'codes', 'ISO', '639-3', 'eus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Basque', 'Language', 'codes', 'Linguasphere', '40-AAA-a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Basque', 'Language', 'codes', 'ISO', '639-1', 'eu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Bulgarian', 'Total', 'speakers', '12', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '12']], 5, True)
```

```
['Bulgarian', 'Region', 'The', 'Balkans']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bulgarian', 'Spoken', 'in', 'Bulgaria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bulgarian', 'Spoken', 'in', 'Greece']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bulgarian', 'Spoken', 'in', 'Serbia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bulgarian', 'Spoken', 'in', 'Turkey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bulgarian', 'Spoken', 'in', 'Ukraine']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Moldova']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Romania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Albania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Kosovo', 'and', 'among', 'emigrant',  
'communities', 'worldwide']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kosovo',  
'and', 'among', 'emigrant', 'communities', 'worldwide']]], 9, True)

['Bulgarian', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Bulgarian', 'Official', 'language', 'in', 'Bulgaria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Bulgarian', 'Official', 'language', 'in', 'European', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bulgarian', 'Official', 'language', 'in', 'Mount', 'Athos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bulgarian', 'Recognised', 'minority', 'language', 'in', 'Serbia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bulgarian', 'Recognised', 'minority', 'language', 'in', 'Romania', '(',  
'Banatian', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Romania', '\\\\(', 'Banatian', 'variant', '\\\\)']]], 10,  
True)

['Bulgarian', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Bulgarian', 'Language', 'codes', 'ISO', '639-2', 'bul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Bulgarian', 'Language', 'codes', 'ISO', '639-3', 'bul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bulgarian', 'Language', 'codes', 'ISO', '639-1', 'bg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Bambara', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bambara', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bambara', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bambara', 'Spoken', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bambara', 'Spoken', 'in', 'The', 'Gambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bambara', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']]], 7, True)

['Bambara', 'Language', 'codes', 'ISO', '639-2', 'bam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Bambara', 'Language', 'codes', 'ISO', '639-3', 'bam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bambara', 'Language', 'codes', 'ISO', '639-1', 'bm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Breton', 'Region', 'Brittany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Breton', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Breton', 'Pronunciation', 'b', '\xc9', '\xbe', 'e', '.', '\xcb', '\x88',  
'z', '\xc3', '\xb5', '\xcb', '\x90', '.', 'n', '\xc9', '\x9b', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'b',  
'\xc9', '\xbe', 'e', '.', '\xcb', '\x88', 'z', '\xc3', '\xb5', '\xcb',  
'\x90', '.', 'n', '\xc9', '\x9b', 'k']], 19, True)

['Breton', 'Language', 'codes', 'ISO', '639-2', 'bre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Breton', 'Language', 'codes', 'ISO', '639-3', 'bre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Breton', 'Language', 'codes', 'ISO', '639-1', 'br']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Blissymbols', 'Regulated', 'by', 'Blissymbolics', 'Communication',  
'International']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Blissymbolics', 'Communication', 'Type: ANY']], 6, True)

['Blissymbols', 'Total', 'speakers', 'some', 'thousands', 'of', 'users']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'some', 'thousands', 'of', 'users']], 7, True)

['Blissymbols', 'Date', 'founded', '1949']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Blissymbols', 'Category', '(', 'purpose', ')', 'constructed',  
'language']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Blissymbols', 'Category', '(', 'purpose', ')', 'Blissymbols']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)
```

```
['Blissymbols', 'Created', 'by', 'Charles', 'K', '.', 'Bliss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
'Charles', 'K', '.', 'Bliss']]], 7, True)
```

```
['Blissymbols', 'Category', '(', 'sources', ')', 'Ideographic', 'written',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Ideographic', 'written', 'language']]], 8, True)
```

```
['Blissymbols', 'Setting', 'and', 'usage', 'Augmentative', 'and',  
'Alternative', 'Communication']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Augmentative', 'and', 'Alternative', 'Communication']]], 8, True)
```

```
['Blissymbols', 'Language', 'codes', 'ISO', '639-2', 'zbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Blissymbols', 'Language', 'codes', 'ISO', '639-3', 'zbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Brahui', 'Spoken', 'in', 'Balochistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Brahui', 'Writing', 'system', 'Perso-Arabic', ',', 'Roman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Brahui', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Brahui', 'Language', 'codes', 'ISO', '639-3', 'brh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Khmer', 'Spoken', 'in', 'Cambodia']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'France']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khmer', 'Writing', 'system', 'Khmer', 'script', '(', 'abugida', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Khmer', 'Official', 'language', 'in', 'Cambodia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Khmer', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Khmer', 'Language', 'codes', 'ISO', '639-2', 'khm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Khmer', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khmer', 'Language', 'codes', 'ISO', '639-1', 'km']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Catalan,', 'Spoken', 'in', 'Andorra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Catalan,', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Catalan,', 'Spoken', 'in', 'Catalonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Catalan,', 'Spoken', 'in', 'Valencian', 'Community']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Catalan,', 'Spoken', 'in', 'Balearic', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Catalan,', 'Spoken', 'in', 'La', 'Franja', '(', 'Aragon', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'La',  
'Franja', '\\\\(', 'Aragon', '\\\\)']]], 8, True)

['Catalan,', 'Spoken', 'in', 'Carche', '(', 'Murcia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)

['Catalan,', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Catalan,', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Catalan,', 'Spoken', 'in', 'Alghero', '(', 'Sardinia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)

['Catalan,', 'Spoken', 'in', 'See', 'geographic', 'distribution', 'of',  
'Catalan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'See',  
'geographic', 'distribution', 'of', 'Type: ANY']]], 8, True)

```
['Catalan,', 'Writing', 'system', 'Latin', '(', 'Catalan', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]]], 8, True)
```

```
['Catalan,', 'Pronunciation', 'k', '\xc9', '\x99', 't', '\xc9', '\x99',  
'\xcb', '\x88', '\xc9', '\xab', 'a', 'or', 'kata', '\xcb', '\x88', 'la']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k',  
'\xc9', '\x99', 't', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xab', 'a',  
'or', 'kata', '\xcb', '\x88', 'la']]]], 18, True)
```

```
['Catalan,', 'Pronunciation', 'valensi', '\xcb', '\x88', 'a']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation',  
'valensi', '\xcb', '\x88', 'a']]]], 6, True)
```

```
['Catalan,', 'Standard', 'forms', 'Catalan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Catalan,', 'Standard', 'forms', 'Valencian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Catalan,', 'Official', 'language', 'in', 'Andorra']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['Catalan,', 'Official', 'language', 'in', 'Italy', ':', 'Alghero']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Italy', ':', 'Alghero']]]], 7, True)
```

```
['Catalan,', 'Official', 'language', 'in', 'Latin', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Catalan,', 'Official', 'language', 'in', 'Spain', ':', 'Catalonia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Spain', ':', 'Catalonia']]]], 7, True)
```

```
['Catalan,', 'Official', 'language', 'in', 'Valencian', 'Community']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Catalan,', 'Official', 'language', 'in', 'Balearic', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Catalan,', 'Regulated', 'by', 'Acad', '\xc3', '\xa8', 'mia',  
'Valenciana', 'de', 'la', 'Llengua']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Acad', '\xc3', '\xa8', 'mia', 'Valenciana', 'de', 'la', 'Llengua']]], 11,  
True)
```

```
['Catalan,', 'Regulated', 'by', 'Acad', '\xc3', '\xa8', 'mia', 'Aragonesa',  
'del', 'Catal', '\xc3', '\xa0']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Acad', '\xc3', '\xa8', 'mia', 'Aragonesa', 'del', 'Catal', '\xc3',  
'\xa0']]], 12, True)
```

```
['Catalan,', 'Language', 'codes', 'ISO', '639-2', 'cat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Catalan,', 'Language', 'codes', 'ISO', '639-3', 'cat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Catalan,', 'Language', 'codes', 'Linguasphere', '51-AAA-e']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Catalan,', 'Language', 'codes', 'ISO', '639-1', 'ca']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Chinese', 'Dialects', 'Mandarin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Chinese', 'Dialects', 'Jin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Chinese', 'Dialects', 'Huizhou']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Chinese', 'Dialects', 'Wu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Chinese', 'Dialects', 'Hunanese']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Chinese', 'Dialects', 'Jiangxinese']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chinese', 'Dialects', 'Hakka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chinese', 'Dialects', 'Yue', '(', 'including', 'Cantonese-Taishanese', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'including', 'Type: ANY', '\\)']]], 7, True)

['Chinese', 'Dialects', 'Pinghua']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chinese', 'Dialects', 'Shaojiang']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chinese', 'Dialects', 'Northern', 'Min']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Dialects', 'Central', 'Min']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Dialects', 'Pu', 'Xian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Dialects', 'Teochew', '(', 'including', 'Swatow', ',', 'Chaozhou', ',', 'Jieyang', ',', 'parts', 'of', 'Shanwei', '/', 'Meizhou', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'including', 'Swatow', ',', 'Chaozhou', ',', 'Jieyang', ',', 'parts', 'of', 'Shanwei', '/', 'Meizhou', '\\)']]], 17, True)

['Chinese', 'Region', '(', 'majorities', ')', ':', 'Mainland', 'China', ',', 'Hong', 'Kong', ',', 'Macau', ',', 'Taiwan', ',', 'Singapore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', '\\(', 'majorities', '\\)', ':', 'Mainland', 'China', ',', 'Hong', 'Kong', ',', 'Hong', 'Kong', ',', 'Macau', ',', 'Taiwan', ',', 'Singapore']]], 17, True)



'Macau', ',', 'Taiwan', ',', 'Type: ANY']]], 17, True)

['Chinese', 'Writing', 'system', 'Chinese', 'characters']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Chinese', 'Writing', 'system', 'zhuyin', 'fuhao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Chinese', 'Writing', 'system', 'pinyin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chinese', 'Writing', 'system', "Xiao'erjing"]:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chinese', 'Standard', 'forms', 'Mandarin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chinese', 'Official', 'language', 'in', 'United', 'Nations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Chinese', 'Official', 'language', 'in', 'Hong', 'Kong', 'Macau',  
'Republic', 'of', 'China', '(', 'Taiwan', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Hong', 'Kong', 'Macau', 'Republic', 'of', 'China', '\\\\(', 'Taiwan',  
'\\\)']]], 13, True)

['Chinese', 'Regulated', 'by', 'In', 'the', 'ROC', ':', 'National',  
'Languages', 'Committee']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'In',  
'the', 'ROC', ':', 'National', 'Languages', 'Committee']]], 10, True)

['Chinese', 'Recognised', 'minority', 'language', 'in', 'Mauritius']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Chinese', 'Recognised', 'minority', 'language', 'in', 'United', 'States',  
'(', 'minority', 'and', 'auxiliary', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'United', 'States', '\\\\(', 'minority', 'and',  
'auxiliary', '\\\)']]], 12, True)

```
['Chinese', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'zho', '\\xe2', '\\x80',  
'\\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'cjl', '\\xe2', '\\x80',  
'\\x93', 'Jinyu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cjl', '\\xe2', '\\x80', '\\x93', 'Jinyu']], ['NUM', '639-3']],  
10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'cmn', '\\xe2', '\\x80',  
'\\x93', 'Mandarin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cmn', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'cpx', '\\xe2', '\\x80',  
'\\x93', 'Pu', 'Xian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cpx', '\\xe2', '\\x80', '\\x93', 'Pu', 'Xian']], ['NUM',  
'639-3']], 11, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'czh', '\\xe2', '\\x80',  
'\\x93', 'Huizhou']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'czh', '\\xe2', '\\x80', '\\x93', 'Huizhou']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'gan', '\\xe2', '\\x80',  
'\\x93', 'Gan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

```
'ISO', 'NUM', 'gan', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'hak', '\xe2', '\x80',  
'\x93', 'Hakka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hak', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'hsn', '\xe2', '\x80',  
'\x93', 'Xiang']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hsn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'wuu', '\xe2', '\x80',  
'\x93', 'Wu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'wuu', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'yue', '\xe2', '\x80',  
'\x93', 'Yue']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'yue', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'lzh', '\xe2', '\x80',  
'\x93', 'Literary', 'Chinese']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lzh', '\xe2', '\x80', '\x93', 'Literary', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'Map', 'of', 'the',  
'Sinophone', 'world']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Map', 'of', 'the', 'Sinophone', 'world']], ['NUM',  
'639-3']], 10, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'Information']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chinese', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Cornish', 'Total', 'speakers', '2', ',', '000', 'fluent']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'fluent']], ['NUM', '2'], ['NUM', '000']], 7, True)
```

```
['Cornish', 'Region', 'Cornwall']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Cornish', 'Spoken', 'in', 'United', 'Kingdom']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Cornish', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Cornish', 'Pronunciation', 'k', '\xc9', '\x99', 'r', '\xcb', '\x88',
'nu', '\xcb', '\x90', '\xc9', '\x99', 'k']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k',
'\xc9', '\x99', 'r', '\xcb', '\x88', 'nu', '\xcb', '\x90', '\xc9', '\x99',
'k']]], 14, True)
```

```
['Cornish', 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Cornish', 'Recognised', 'minority', 'language', 'in', 'United',
'Kingdom']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'United', 'Kingdom']]], 7, True)
```

```
['Cornish', 'Language', 'codes', 'ISO', '639-2', 'cor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Cornish', 'Language', 'codes', 'ISO', '639-3', 'cor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cornish', 'Language', 'codes', 'Linguasphere', '50-ABB-a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Cornish', 'Language', 'codes', 'ISO', '639-1', 'kw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Czech', 'Total', 'speakers', '12', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '12']], 5, True)

['Czech', 'Region', 'Central', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Czech', 'Spoken', 'in', 'Czech', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Czech', 'Spoken', 'in', 'Austria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Czech', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Czech', 'Spoken', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Czech', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Czech', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Czech', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Czech', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']], 6, True)

['Czech', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```

['Czech', 'Writing', 'system', 'Czech', 'variant', 'of', 'Latin',
'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Czech', 'Official', 'language', 'in', 'Czech', 'Republic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Czech', 'Official', 'language', 'in', 'European', 'Union']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Czech', 'Language', 'codes', 'ISO', '639-2', 'cze', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'cze', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Czech', 'Language', 'codes', 'ISO', '639-2', 'ces', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'ces', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Czech', 'Language', 'codes', 'ISO', '639-3', 'ces']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Czech', 'Language', 'codes', 'ISO', '639-1', 'cs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Corsican', 'Total', 'speakers', 'Current', 'number', 'not', 'known', '.',
'In', '1990', ',', '225', ',', '000', 'in', 'Corsica', 'alone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'Current', 'number', 'not', 'known', '.', 'In', 'NUM', ',', 'NUM', ',',
'NUM', 'in', 'Corsica', 'alone']]], ['NUM', '1990'], ['NUM', '225'], ['NUM',
'000']], 17, True)

['Corsican', 'Spoken', 'in', 'Corsica', ',', 'northern', 'Sardinia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Corsican', 'Pronunciation', '\\xcb', '\\x88', 'k', '\\xc9', '\\x94', 'rsu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb',
'\\x88', 'k', '\\xc9', '\\x94', 'rsu']]], 8, True)

['Corsican', 'Language', 'codes', 'ISO', '639-2', 'cos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Corsican', 'Language', 'codes', 'ISO', '639-3', 'cos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Corsican', 'Language', 'codes', 'Linguasphere', '51-AAA-p']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Corsican', 'Language', 'codes', 'Linguasphere', 'Corsican', 'dialects']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY', 'dialects']]], 6, True)

['Corsican', 'Language', 'codes', 'ISO', '639-1', 'co']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Danish', 'Total', 'speakers', 'c', '.', '6', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c',  
'.', 'NUM', 'million']], ['NUM', '6']], 7, True)

['Danish', 'Spoken', 'in', 'Denmark', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',']], 5, True)

['Danish', 'Spoken', 'in', 'Faroe', 'Islands', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Faroe',  
'Islands', ',']], 6, True)

['Danish', 'Spoken', 'in', 'Greenland', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',']], 5, True)

['Danish', 'Spoken', 'in', 'Schleswig-Holstein', '(', 'Germany', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Danish', 'Writing', 'system', 'Latin', '(', 'Danish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Danish', 'Pronunciation', 'dan', '\\xcb', '\\x80', 's', '\\xc9', '\\xa1',  
'\\xcc', '\\x8a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'dan',  
'\\xcb', '\\x80', 's', '\\xc9', '\\xa1', '\\xcc', '\\x8a']]], 10, True)

['Danish', 'Official', 'language', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Danish', 'Official', 'language', 'in', 'Faroe', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Danish', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Danish', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Danish', 'Official', 'language', 'in', 'Minority', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Danish', 'Official', 'language', 'in', ':', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Danish', 'Language', 'codes', 'ISO', '639-2', 'dan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Danish', 'Language', 'codes', 'ISO', '639-3', 'dan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Danish', 'Language', 'codes', 'ISO', '639-1', 'da']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Dioula', 'Total', 'speakers', '2', ',', '700', ',', '000', '(',  
'several', 'millions', 'more', 'including', 'second', 'language',  
'speakers', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', '\\\\(', 'several', 'millions', 'more',  
'including', 'second', 'language', 'speakers', '\\\\)']]], ['NUM', '2'],  
['NUM', '700'], ['NUM', '000']], 17, True)

['Dioula', 'Spoken', 'in', 'Burkina', 'Faso']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dioula', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dioula', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']]], 7, True)

['Dioula', 'Language', 'codes', 'ISO', '639-2', 'dyu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Dioula', 'Language', 'codes', 'ISO', '639-3', 'dyu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Esperanto', 'Regulated', 'by', 'Akademio', 'de', 'Esperanto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Akademio', 'de', 'Type: ANY']]], 6, True)

['Esperanto', 'Date', 'founded', '1887']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Esperanto', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Esperanto', 'Category', '(', 'purpose', ')', 'Esperanto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Esperanto', 'Created', 'by', 'L', '.', 'L', '.', 'Zamenhof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'L', '.', 'L', '.', 'Zamenhof']]], 8, True)

['Esperanto', 'Category', '(', 'sources', ')', 'Vocabulary', 'from', 'Romance', 'and', 'Germanic', 'languages', ';', 'phonology', 'from', 'Slavic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'sources', '\\\\)', 'Vocabulary', 'from', 'Romance', 'and', 'Germanic', 'languages', ';', 'phonology', 'from', 'Slavic', 'languages']]], 16, True)

['Esperanto', 'Language', 'codes', 'ISO', '639-2', 'epo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Esperanto', 'Language', 'codes', 'ISO', '639-3', 'epo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Esperanto', 'Language', 'codes', 'ISO', '639-1', 'eo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Etruscan', 'Spoken', 'in', 'Ancient', 'Etruria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Etruscan', 'Region', 'Italian', 'Peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Etruscan', 'Language', 'codes', 'ISO', '639-3', 'ett']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Estonian', 'Region', 'Northern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Estonian', 'Spoken', 'in', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Estonian', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Estonian', 'Writing', 'system', 'Historically', 'Runic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Estonian', 'Official', 'language', 'in', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Estonian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Estonian', 'Language', 'codes', 'ISO', '639-2', 'est']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Estonian', 'Language', 'codes', 'ISO', '639-3', 'est', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'est', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Estonian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Estonian', 'Language', 'codes', 'ISO', '639-3', 'vro', '\xe2', '\x80',  
'\x93', 'V', '\xc3', '\xb5', 'ro']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'vro', '\xe2', '\x80', '\x93', 'V', '\xc3', '\xb5', 'ro']],  
['NUM', '639-3']], 13, True)

['Estonian', 'Language', 'codes', 'ISO', '639-1', 'et']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Frisian', 'Region', 'Friesland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Frisian', 'Region', 'Groningen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Frisian', 'Region', 'Lower', 'Saxony']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Frisian', 'Region', 'Schleswig-Holstein']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Frisian', 'Spoken', 'in', 'Netherlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Frisian', 'Spoken', 'in', 'Germany']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Frisian', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Frisian', 'Official', 'language', 'in', 'Netherlands']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Frisian', 'Official', 'language', 'in', 'Germany']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Frisian', 'Regulated', 'by', 'NL', ':', 'Fryske', 'Akademy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'NL',
':', 'Fryske', 'Akademy']]], 7, True)
```

```
['Frisian', 'Regulated', 'by', 'D', ':', 'no', 'official', 'regulation']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'D',
':', 'no', 'official', 'regulation']]], 8, True)
```

```
['Frisian', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Frisian', 'Language', 'codes', 'Linguasphere', '52-ACA']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['French', 'Pronunciation', 'f', '\xca', '\x81', '\xc9', '\x91', '\xcc',
'\x83', 's', '\xc9', '\x9b']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'f',
'\xca', '\x81', '\xc9', '\x91', '\xcc', '\x83', 's', '\xc9', '\x9b']]], 12,
True)
```

```
['French', 'Total', 'speakers', 'native', 'speakers', ':', '110',
'million', ',', 'total', '(', 'both', 'native', 'and', 'second',
'language', ')', 'speakers', ':', '265', 'million-', '270', 'million']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'native', 'speakers', ':', 'NUM', 'million', ',', 'total', '\\(', 'both',
'native', 'and', 'second', 'language', '\\)'], 'speakers', ':', 'NUM',
'million-', 'NUM', 'million']]], ['NUM', '110'], ['NUM', '265'], ['NUM',
'270']], 23, True)
```

['French', 'Official', 'language', 'in', '30', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['French', 'Official', 'language', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Benin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['French', 'Official', 'language', 'in', 'Burundi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Comoros']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of',  
'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 9, True)

['French', 'Official', 'language', 'in', 'Djibouti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Equatorial', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['French', 'Official', 'language', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Haiti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Luxembourg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Madagascar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['French', 'Official', 'language', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Seychelles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Togo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Administrative', '/',  
'cultural']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Administrative', '/', 'cultural']]], 7, True)

['French', 'Official', 'language', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Mauritius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Mauritania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['French', 'Official', 'language', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

```
['French', 'Official', 'language', 'in', '14', 'dependent', 'entities']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'NUM', 'dependent', 'entities']], ['NUM', '14']], 7, True)
```

```
['French', 'Official', 'language', 'in', 'Aosta', 'Valley']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['French', 'Official', 'language', 'in', 'Clipperton']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['French', 'Official', 'language', 'in', 'French', 'Polynesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['French', 'Official', 'language', 'in', 'Guernsey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['French', 'Official', 'language', 'in', 'Jersey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['French', 'Official', 'language', 'in', 'Mayotte']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['French', 'Official', 'language', 'in', 'New', 'Caledonia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['French', 'Official', 'language', 'in', 'Puducherry']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['French', 'Official', 'language', 'in', 'Saint-Barth', '\xc3', '\xa9',  
'lemy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Saint-Barth', '\xc3', '\xa9', 'lemy']], 8, True)
```

```
['French', 'Official', 'language', 'in', 'Saint-Martin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```



```
['French', 'Official', 'language', 'in', 'Saint-Pierre', 'and',  
'Miquelon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Saint-Pierre', 'and', 'Miquelon']]]], 7, True)
```

```
['French', 'Official', 'language', 'in', 'Wallis', 'and', 'Futuna']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Wallis', 'and', 'Futuna']]]], 7, True)
```

```
['French', 'Official', 'language', 'in', 'Numerous', 'international',  
'organisations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Numerous', 'international', 'organisations']]]], 7, True)
```

```
['French', 'Language', 'codes', 'ISO', '639-2', 'fre', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'fre', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)
```

```
['French', 'Language', 'codes', 'ISO', '639-2', 'fra', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'fra', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)
```

```
['French', 'Language', 'codes', 'ISO', '639-3', 'fra']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['French', 'Language', 'codes', 'Linguasphere', '51-AAA-i']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]]], 5, True)
```

```
['French', 'Language', 'codes', 'ISO', '639-1', 'fr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['West', 'Spoken', 'in', 'Belgium']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['West', 'Spoken', 'in', 'Netherlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['West', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['West', 'Region', 'West', 'Flanders']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['West', 'Language', 'codes', 'ISO', '639-3', 'vls']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['West', 'Language', 'codes', 'Linguasphere', '52-ACB-ag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Faroese', 'Spoken', 'in', 'Faroe', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Faroese', 'Spoken', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Faroese', 'Spoken', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Faroese', 'Writing', 'system', 'Latin', '(', 'Faroese', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Faroese', 'Pronunciation', '\xcb', '\x88', 'f', '\xc3', '\xb8', '\xcb',  
'\x90', '\xc9', '\xb9', '\xc9', '\xaa', 'st', ',', '\xcb', '\x88', 'f',  
'\xc3', '\xb8', '\xcb', '\x90', '\xc9', '\xb9', '\xc9', '\xaa', '\xca',  
'\x82', 't']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'f', '\xc3', '\xb8', '\xcb', '\x90', '\xc9', '\xb9', '\xc9',  
'\xaa', 'st', ',', '\xcb', '\x88', 'f', '\xc3', '\xb8', '\xcb', '\x90',  
'\xc9', '\xb9', '\xc9', '\xaa', '\xca', '\x82', 't']]], 29, True)
```

```
['Faroese', 'Official', 'language', 'in', 'Faroe', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Faroese', 'Recognised', 'minority', 'language', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)
```

```
['Faroese', 'Language', 'codes', 'ISO', '639-2', 'fao']:
```

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Faroese', 'Language', 'codes', 'ISO', '639-3', 'fao']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Faroese', 'Language', 'codes', 'Linguasphere', '52-AAA-ab']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Faroese', 'Language', 'codes', 'ISO', '639-1', 'fo']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Persian', 'Dialects', 'Tehrani']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Dari']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Tajik']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Khorasani']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Khuzestani']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Aimaq']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Hazaragi']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Persian', 'Dialects', 'Jud', '\xc3', '\xa6', 'o-Persian']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Jud', '\xc3', '\xa6', 'o-Persian']], 6, True)

['Persian', 'Dialects', 'Bukhori']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Persian', 'Dialects', 'Dezfuli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Persian', 'Dialects', 'Juhuri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Persian', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Bahrain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Persian', 'Spoken', 'in', 'Iranian', 'diaspora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Persian', 'Writing', 'system', 'Perso-Arabic', 'script', ',', 'Cyrillic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'script', ',', 'Type: ANY']]], 7, True)

['Persian', 'Pronunciation', 'p', '\xc9', '\x92', '\xcb', '\x90', '\xc9', '\xbe', '\xcb', '\x88', 'si']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\xc9', '\x92', '\xcb', '\x90', '\xc9', '\xbe', '\xcb', '\x88', 'si']]], 12, True)

['Persian', 'Official', 'language', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Persian', 'Official', 'language', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Persian', 'Official', 'language', 'in', 'Tajikistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Persian', 'Language', 'codes', 'ISO', '639-2', 'per', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'per', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Persian', 'Language', 'codes', 'ISO', '639-2', 'fas', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fas', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Persian', 'Language', 'codes', 'ISO', '639-3', 'fas', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fas', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)

['Persian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Persian', 'Language', 'codes', 'ISO', '639-3', ':', 'pes', '\xe2', '\x80', '\x93', 'Iranian', 'Persian']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'pes', '\xe2', '\x80', '\x93', 'Iranian', 'Type: ANY']],  
['NUM', '639-3']], 12, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'prs', '\xe2', '\x80',  
\x93', 'Afghan', 'Persian']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'prs', '\xe2', '\x80', '\x93', 'Afghan', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'tgk', '\xe2', '\x80',  
\x93', 'Tajik']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tgk', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'aiq', '\xe2', '\x80',  
\x93', 'Aimaq']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'aiq', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'bhh', '\xe2', '\x80',  
\x93', 'Bukharic']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bhh', '\xe2', '\x80', '\x93', 'Bukharic']], ['NUM',  
'639-3']], 10, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'drw', '\xe2', '\x80',  
\x93', 'Darwazi']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'drw', '\xe2', '\x80', '\x93', 'Darwazi']], ['NUM',  
'639-3']], 10, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'haz', '\xe2', '\x80',  
\x93', 'Hazaragi']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'haz', '\xe2', '\x80', '\x93', 'Hazaragi']], ['NUM',  
'639-3']], 10, True)
```

```
['Persian', 'Language', 'codes', 'ISO', '639-3', 'jpr', '\xe2', '\x80',  
\x93', 'Dzhidi']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'jpr', '\xe2', '\x80', '\x93', 'Dzhidi']], ['NUM', '639-3']],  
10, True)
```

['Persian', 'Language', 'codes', 'ISO', '639-3', 'phv', '\xe2', '\x80', '\x93', 'Pahlavani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'phv', '\xe2', '\x80', '\x93', 'Pahlavani']], ['NUM', '639-3']], 10, True)

['Persian', 'Language', 'codes', 'Linguasphere', 'Countries', 'with', 'significant', 'Persian-speaking', 'populations']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Countries', 'with', 'significant', 'Persian-speaking', 'populations']], 9, True)

['Persian', 'Language', 'codes', 'Linguasphere', '(', 'Click', 'on', 'image', 'for', 'the', 'legend', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', '\\(', 'Click', 'on', 'image', 'for', 'the', 'legend', '\\)']], 12, True)

['Persian', 'Language', 'codes', 'ISO', '639-1', 'fa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['German', 'Total', 'speakers', 'Native', 'speakers', ':', 'ca', '.', '120', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Native', 'speakers', ':', 'ca', '.', 'NUM', 'million']], ['NUM', '120']], 10, True)

['German', 'Total', 'speakers', 'Non-native', 'speakers', ':', 'ca', '.', '80', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Non-native', 'speakers', ':', 'ca', '.', 'NUM', 'million']], ['NUM', '80']], 10, True)

['German', 'Pronunciation', '\xcb', '\x88', 'd', '\xc9', '\x94', '\xca', '\x8f', 't', '\xca', '\x83']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'd', '\xc9', '\x94', '\xca', '\x8f', 't', '\xca', '\x83']], 12, True)

['German', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['German', 'Official', 'language', 'in', '(', 'official', 'and', 'working', 'language', ')']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', '\\\\(', 'official', 'and', 'working', 'language', '\\\\)']]], 10, True)

['German', 'Official', 'language', 'in', 'Austria']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['German', 'Official', 'language', 'in', 'Belgium']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['German', 'Official', 'language', 'in', 'Germany']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['German', 'Official', 'language', 'in', 'Liechtenstein']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['German', 'Official', 'language', 'in', 'Luxembourg']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['German', 'Official', 'language', 'in', 'Switzerland']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['German', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['German', 'Recognised', 'minority', 'language', 'in', 'Denmark']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['German', 'Recognised', 'minority', 'language', 'in', 'Hungary']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['German', 'Recognised', 'minority', 'language', 'in', 'Kazakhstan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['German', 'Recognised', 'minority', 'language', 'in', 'Romania']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)



```
['German', 'Recognised', 'minority', 'language', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-2', 'ger', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ger', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-2', 'deu', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'deu', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-3', 'gct', '\\xe2', '\\x80',  
'\\x93', 'Alem', '\\xc3', '\\xa1', 'n', 'Coloniero']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'gct', '\\xe2', '\\x80', '\\x93', 'Alem', '\\xc3', '\\xa1', 'n',  
'Coloniero']]], ['NUM', '639-3']], 14, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-3', 'bar', '\\xe2', '\\x80',  
'\\x93', 'Austro-Bavarian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bar', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-3', 'cim', '\\xe2', '\\x80',  
'\\x93', 'Cimbrian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cim', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-3', 'ksh', '\\xe2', '\\x80',  
'\\x93', 'K', '\\xc3', '\\xb6', 'lsch']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ksh', '\\xe2', '\\x80', '\\x93', 'K', '\\xc3', '\\xb6', 'lsch']]],  
['NUM', '639-3']], 13, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-3', 'ltz', '\\xe2', '\\x80',  
'\\x93', 'Luxembourgish']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ltz', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)
```

['German', 'Language', 'codes', 'ISO', '639-3', 'vmf', '\xe2', '\x80', '\x93', 'Main-Franconian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'vmf', '\xe2', '\x80', '\x93', 'Main-Franconian']], ['NUM', '639-3']], 10, True)

['German', 'Language', 'codes', 'ISO', '639-3', 'mhn', '\xe2', '\x80', '\x93', 'M', '\xc3', '\xb3', 'cheno']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mhn', '\xe2', '\x80', '\x93', 'M', '\xc3', '\xb3', 'cheno']], ['NUM', '639-3']], 13, True)

['German', 'Language', 'codes', 'ISO', '639-3', 'pfl', '\xe2', '\x80', '\x93', 'Palatinate', 'German']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pfl', '\xe2', '\x80', '\x93', 'Palatinate', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['German', 'Language', 'codes', 'ISO', '639-3', 'pdt', '\xe2', '\x80', '\x93', 'Plautdietsch']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pdt', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['German', 'Language', 'codes', 'ISO', '639-3', 'uln', '\xe2', '\x80', '\x93', 'Unserdeutsch']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'uln', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['German', 'Language', 'codes', 'ISO', '639-3', 'wep', '\xe2', '\x80', '\x93', 'Westphalian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wep', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['German', 'Language', 'codes', 'ISO', '639-1', 'de']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Greek', 'Total', 'speakers', 'c', '.', '13', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c', '.', 'NUM', 'million']], ['NUM', '13']], 7, True)

['Greek', 'Spoken', 'in', 'Greece']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Cyprus']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Albania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Egypt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Greek', 'diaspora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Greek', 'Spoken', 'in', ':', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':', 'Type: ANY', 'Type: ANY']]], 6, True)

['Greek', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'United', 'Kingdom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Greek', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Greek', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Greek', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Greek', 'Pronunciation', 'elini', '\xcb', '\x88', 'ka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'elini', '\xcb', '\x88', 'ka']], 6, True)

['Greek', 'Official', 'language', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Greek', 'Official', 'language', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Greek', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Greek', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Language', 'codes', 'ISO', '639-2', 'gre', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gre', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Greek', 'Language', 'codes', 'ISO', '639-2', 'ell', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ell', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Greek', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Greek', 'Language', 'codes', 'ISO', '639-1', 'el']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Gallurese', 'Spoken', 'in', 'Gallura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gallurese', 'Spoken', 'in', 'northeastern', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Gallurese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gallurese', 'Language', 'codes', 'ISO', '639-3', 'sdn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gallurese', 'Language', 'codes', 'Linguasphere', '51-AAA-pd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Hebrew', 'Total', 'speakers', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hebrew', 'Total', 'speakers', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'United', 'States']], 5, True)

['Hebrew', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hebrew', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Hebrew', 'Pronunciation', 'Yemenite', ':', '\xca', '\x95', 'iv', '\xcb', '\x88', 'ri', '\xcb', '\x90', '\xce', '\xb8', ',']:   
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'Yemenite', ':', '\xca', '\x95', 'iv', '\xcb', '\x88', 'ri', '\xcb', '\x90', '\xce', '\xb8', ',']], 15, True)

['Hebrew', 'Pronunciation', 'Ashkenazi', ':', '\xcb', '\x88', 'iv', '\xca', '\x80', 'is']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation',

'Ashkenazi', ':', '\xcb', '\x88', 'iv', '\xca', '\x80', 'is']]], 10, True)

['Hebrew', 'Official', 'language', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Hebrew', 'Official', 'language', 'in', 'Judaism']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Hebrew', 'Language', 'codes', 'ISO', '639-2', 'heb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hebrew', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hebrew', 'Language', 'codes', 'Linguasphere', '12-AAB-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Hebrew', 'Language', 'codes', 'ISO', '639-1', 'he']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Hindi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Mauritius', 'and', 'significant', 'communities', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mauritius', 'and', 'significant', 'communities', 'in', 'USA']]], 9, True)

['Hindi', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hindi', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'United', 'Arab', 'Emirates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'Arab', 'Emirates']], 6, True)

['Hindi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Hindi', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hindi', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Hindi', 'Writing', 'system', 'Devanagari']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hindi', 'Official', 'language', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Hindi', 'Language', 'codes', 'ISO', '639-2', 'hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Hindi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Hindi', 'Language', 'codes', 'ISO', '639-3', 'hin-hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hindi', 'Language', 'codes', 'ISO', '639-1', 'hi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Hungarian', 'Total', 'speakers', '16', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '16']], 5, True)

['Hungarian', 'Spoken', 'in', 'Hungary', 'and', 'areas', 'of', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Hungary', 'and', 'areas', 'of', 'Romania']]], 8, True)

['Hungarian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hungarian', 'Spoken', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hungarian', 'Spoken', 'in', 'Slovenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hungarian', 'Spoken', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hungarian', 'Spoken', 'in', 'Croatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hungarian', 'Spoken', 'in', 'Austria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hungarian', 'Spoken', 'in', 'and', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hungarian', 'Pronunciation', '\xcb', '\x88', 'm', '\xc9', '\x92', '\xc9', '\x9f', '\xc9', '\x92', 'r']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'm', '\xc9', '\x92', '\xc9', '\x9f', '\xc9', '\x92', 'r']]], 12, True)

['Hungarian', 'Official', 'language', 'in', 'Hungary']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Hungarian', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Hungarian', 'Official', 'language', 'in', 'Slovakia', '(', 'regional', 'language', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Slovakia', '\\(', 'regional', 'language', '\\)']]], 9, True)

['Hungarian', 'Official', 'language', 'in', 'Slovenia', '(', 'regional', 'language', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Slovenia', '\\(', 'regional', 'language', '\\)']]], 9, True)

['Hungarian', 'Official', 'language', 'in', 'Serbia', '(', 'regional', 'language', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Serbia', '\\(', 'regional', 'language', '\\)']]], 9, True)

['Hungarian', 'Official', 'language', 'in', 'Austria', '(', 'regional', 'language', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Austria', '\\\\(', 'regional', 'language', '\\\\)']]], 9, True)

['Hungarian', 'Official', 'language', 'in', 'Some', 'official', 'rights', 'in', 'Romania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Some', 'official', 'rights', 'in', 'Romania']]], 9, True)

['Hungarian', 'Official', 'language', 'in', 'Ukraine', 'and', 'Croatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Ukraine', 'and', 'Croatia']]], 7, True)

['Hungarian', 'Language', 'codes', 'ISO', '639-2', 'hun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hungarian', 'Language', 'codes', 'ISO', '639-3', 'hun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hungarian', 'Language', 'codes', 'ISO', '639-1', 'hu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Hausa', 'Spoken', 'in', 'Benin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Spoken', 'in', 'Burkina', 'Faso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hausa', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Spoken', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Spoken', 'in', 'Niger']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Hausa', 'Spoken', 'in', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Spoken', 'in', 'Togo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Total', 'speakers', '25', 'million', 'as', 'a', 'first', 'language', ',', '18', 'million', 'as', 'a', 'second', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'as', 'a', 'first', 'language', ',', 'NUM', 'million', 'as', 'a', 'second', 'language']]], ['NUM', '25'], ['NUM', '18']], 16, True)

['Hausa', 'Writing', 'system', 'Latin', ',', 'Arabic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Hausa', 'Language', 'codes', 'ISO', '639-2', 'hau']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Hausa', 'Language', 'codes', 'ISO', '639-3', 'hau']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hausa', 'Language', 'codes', 'ISO', '639-1', 'ha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Hawaiian', 'Official', 'language', 'in', 'recognized', 'as', 'minority', 'language', 'in', 'parts', 'of']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognized', 'as', 'minority', 'language', 'in', 'parts', 'of']]], 11, True)

['Hawaiian', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Hawaiian', 'Language', 'codes', 'ISO', '639-2', 'haw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

```

['Hawaiian', 'Language', 'codes', 'ISO', '639-3', 'haw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Italian', 'Total', 'speakers', '62', 'million', 'or', '65', 'million',
'in', 'the', 'European', 'Union', ';', 'more', 'than', '20', 'million',
'as', 'second', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'or', 'NUM', 'million', 'in', 'the', 'European', 'Union',
';', 'more', 'than', 'NUM', 'million', 'as', 'second', 'language']],
['NUM', '62'], ['NUM', '65'], ['NUM', '20']], 20, True)

['Italian', 'Region', 'al', 'in', 'Slovenia', 'and', 'Croatia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'al', 'in',
'Slovenia', 'and', 'Croatia']]], 7, True)

['Italian', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'Switzerland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'San', 'Marino']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Italian', 'Spoken', 'in', 'Vatican', 'City']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Italian', 'Spoken', 'in', 'Malta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'Eritrea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'Somalia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Italian', 'Pronunciation', 'ita', '\xcb', '\x88', 'ljano']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ita',

```

'\xcb', '\x88', 'ljano']]], 6, True)

['Italian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Italian', 'Official', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Italian', 'Official', 'language', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Italian', 'Official', 'language', 'in', 'San', 'Marino']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Italian', 'Official', 'language', 'in', 'Vatican', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Italian', 'Official', 'language', 'in', 'Sovereign', 'Military', 'Order',  
'of', 'Malta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Sovereign', 'Military', 'Order', 'of', 'Malta']]], 9, True)

['Italian', 'Official', 'language', 'in', 'Croatia', '(', 'Istria',  
'County', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Croatia', '\\(', 'Istria', 'County', '\\)']]], 9, True)

['Italian', 'Official', 'language', 'in', 'Slovenia', '(', 'Slovenian',  
'Istria', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Slovenia', '\\(', 'Slovenian', 'Istria', '\\)']]], 9, True)

['Italian', 'Regulated', 'by', 'not', 'officially', 'by', 'Accademia',  
'della', 'Crusca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'not',  
'officially', 'by', 'Accademia', 'della', 'Crusca']]], 9, True)

['Italian', 'Language', 'codes', 'ISO', '639-2', 'ita']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Italian', 'Language', 'codes', 'ISO', '639-3', 'ita']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Italian', 'Language', 'codes', 'Linguasphere', '51-AAA-q']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]]], 5, True)
```

```
['Italian', 'Language', 'codes', 'ISO', '639-1', 'it']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Ido', 'Regulated', 'by', 'Uniono', 'por', 'la', 'Linguo', 'Internaciona',
'Ido']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Uniono', 'por', 'la', 'Linguo', 'Internaciona', 'Type: ANY']]]], 9, True)
```

```
['Ido', 'Total', 'speakers', '(', 'but', 'see', 'community', 'section',
')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\
(', 'but', 'see', 'community', 'section', '\\)']]]], 9, True)
```

```
['Ido', 'Date', 'founded', '1907']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['Ido', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\
(', 'purpose', '\\)', 'Type: ANY', 'language']]]], 7, True)
```

```
['Ido', 'Category', '(', 'purpose', ')', 'Ido']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\
(', 'purpose', '\\)', 'Type: ANY']]]], 6, True)
```

```
['Ido', 'Category', '(', 'sources', ')', 'based', 'on', 'Esperanto']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\
(', 'sources', '\\)', 'based', 'on', 'Type: ANY']]]], 8, True)
```

```
['Ido', 'Language', 'codes', 'ISO', '639-2', 'ido']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Ido', 'Language', 'codes', 'ISO', '639-3', 'ido']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```

['Ido', 'Language', 'codes', 'Linguasphere', '51-AAB-db']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Ido', 'Language', 'codes', 'ISO', '639-1', 'io']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'io']], ['NUM', '639-1']], 6, True)

['Interlingua', 'Regulated', 'by', 'no', 'regulating', 'body']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'no',
'regulating', 'body']]], 6, True)

['Interlingua', 'Total', 'speakers', 'a', 'few', 'hundred']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a',
'few', 'hundred']]], 6, True)

['Interlingua', 'Date', 'founded', '1951']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Interlingua', 'Category', '(', 'purpose', ')', 'international',
'auxiliary', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'international', 'auxiliary', 'language']]], 8, True)

['Interlingua', 'Category', '(', 'purpose', ')', 'Interlingua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Interlingua', 'Writing', 'system', 'Latin', 'alphabet', 'French']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Interlingua', 'Writing', 'system', 'Italian']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'Portuguese']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'Spanish']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'German']:

```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'English']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'Russian', 'and', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Interlingua', 'Pronunciation', '/', '\xc9', '\xaa', 'nt', '\xc9', '\x99', 'r', '\xcb', '\x88', 'l', '\xc9', '\xaa', '\xc5', '\x8b', '\xc9', '\xa1', 'w', '\xc9', '\x99', '/', ';', 'IA', ':', 'inte', '\xc9', '\xbe', '\xcb', '\x88', 'li', '\xc5', '\x8b', '\xc9', '\xa1', 'wa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', '\xc9', '\xaa', 'nt', '\xc9', '\x99', 'r', '\xcb', '\x88', 'l', '\xc9', '\xaa', '\xc5', '\x8b', '\xc9', '\xa1', 'w', '\xc9', '\x99', '/', ';', 'IA', ':', 'inte', '\xc9', '\xbe', '\xcb', '\x88', 'li', '\xc5', '\x8b', '\xc9', '\xa1', 'wa']]], 36, True)

['Interlingua', 'Setting', 'and', 'usage', 'Scientific', 'registration', 'of', 'international', 'vocabulary', ';', 'international', 'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Scientific', 'registration', 'of', 'international', 'vocabulary', ';', 'international', 'auxiliary', 'language']]], 13, True)

['Interlingua', 'Language', 'codes', 'ISO', '639-2', 'ina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Interlingua', 'Language', 'codes', 'ISO', '639-3', 'ina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Interlingua', 'Language', 'codes', 'ISO', '639-1', 'ia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Japanese', 'Total', 'speakers', '130', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '130']], 5, True)

['Japanese', 'Spoken', 'in', 'Majority', ':', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',

'Majority', ':', 'Japan']]], 6, True)

['Japanese', 'Writing', 'system', 'Chinese', 'characters']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Japanese', 'Writing', 'system', 'r', '\xc5', '\x8d', 'maji']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'r',  
'\xc5', '\x8d', 'maji']]], 7, True)

['Japanese', 'Writing', 'system', 'Siddham', 'script', '(', 'occasionally',  
'in', 'Buddhist', 'temples', '.', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Siddham', 'script', '\\(', 'occasionally', 'in', 'Buddhist', 'temples',  
'.', '\\)']]], 12, True)

['Japanese', 'Pronunciation', 'n', '\xca', '\xb2', 'iho', '\xc5', '\x8b',  
'\xc9', '\xa1', 'o']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n',  
'\xca', '\xb2', 'iho', '\xc5', '\x8b', '\xc9', '\xa1', 'o']]], 11, True)

['Japanese', 'Official', 'language', 'in', 'Japan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Japanese', 'Official', 'language', 'in', 'Palau', '(', 'Angaur', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Palau', '\\(', 'Angaur', '\\)']]], 8, True)

['Japanese', 'Regulated', 'by', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Japanese', 'Language', 'codes', 'ISO', '639-2', 'jpn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Japanese', 'Language', 'codes', 'ISO', '639-3', 'jpn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Japanese', 'Language', 'codes', 'ISO', '639-1', 'ja']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Korean', 'Spoken', 'in', 'South', 'Korea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Korean', 'Spoken', 'in', 'North', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Korean', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Korean', 'Writing', 'system', 'Exclusive', 'use', 'of', 'Hangul', '(', 'Both', 'Koreas', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Exclusive', 'use', 'of', 'Hangul', '\\(', 'Both', 'Koreas', '\\)']]], 11, True)

['Korean', 'Writing', 'system', 'mix', 'of', 'Hangul', 'and', 'hanja', '(', 'some', 'professional', 'scripts', 'in', 'S', '.', 'Korea', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'mix', 'of', 'Hangul', 'and', 'hanja', '\\(', 'some', 'professional', 'scripts', 'in', 'S', '.', 'Korea', '\\)']]], 17, True)

['Korean', 'Writing', 'system', 'or', 'Cyrillic', 'alphabet', '(', 'lesser', 'used', 'in', 'Goryeomal', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'or', 'Cyrillic', 'alphabet', '\\(', 'lesser', 'used', 'in', 'Goryeomal', '\\)']]], 12, True)

['Korean', 'Official', 'language', 'in', 'North', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Korean', 'Official', 'language', 'in', 'South', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Korean', 'Regulated', 'by', 'South', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Korean', 'Regulated', 'by', '\\xea', '\\xb5', '\\xad', '\\xeb', '\\xa6', '\\xbd', '\\xea', '\\xb5', '\\xad', '\\xec', '\\x96', '\\xb4', '\\xec', '\\x9b', '\\x90', '/', '\\xe5', '\\x9c', '\\x8b', '\\xe7', '\\xab', '\\x8b', '\\xe5', '\\x9c', '\\x8b', '\\xe8', '\\xaa', '\\x9e', '\\xe9', '\\x99', '\\xa2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',

```
'\xea', '\xb5', '\xad', '\xeb', '\xa6', '\xbd', '\xea', '\xb5', '\xad',
'\xec', '\x96', '\xb4', '\xec', '\x9b', '\x90', '/', '\xe5', '\x9c',
'\x8b', '\xe7', '\xab', '\x8b', '\xe5', '\x9c', '\x8b', '\xe8', '\xaa',
'\x9e', '\xe9', '\x99', '\xa2']]], 34, True)
```

```
['Korean', 'Regulated', 'by', 'North', 'Korea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Korean', 'Regulated', 'by', ':', 'Sahoe', 'Kwahagwon', '\xc5', '\x8e',
'hak', 'Y', '\xc5', '\x8f', 'nguso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', ':',
'Sahoe', 'Kwahagwon', '\xc5', '\x8e', 'hak', 'Y', '\xc5', '\x8f',
'nguso']]], 13, True)
```

```
['Korean', 'Regulated', 'by', '\xec', '\x82', '\xac', '\xed', '\x9a',
'\x8c', '\xea', '\xb3', '\xbc', '\xed', '\x95', '\x99', '\xec', '\x9b',
'\x90', '\xec', '\x96', '\xb4', '\xed', '\x95', '\x99', '\xec', '\x97',
'\xb0', '\xea', '\xb5', '\xac', '\xec', '\x86', '\x8c']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'\xec', '\x82', '\xac', '\xed', '\x9a', '\x8c', '\xea', '\xb3', '\xbc',
'\xed', '\x95', '\x99', '\xec', '\x9b', '\x90', '\xec', '\x96', '\xb4',
'\xed', '\x95', '\x99', '\xec', '\x97', '\xb0', '\xea', '\xb5', '\xac',
'\xec', '\x86', '\x8c']]], 33, True)
```

```
['Korean', 'Language', 'codes', 'ISO', '639-2', 'kor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Korean', 'Language', 'codes', 'ISO', '639-3', 'kor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Korean', 'Language', 'codes', 'Linguasphere', '45-AAA-a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Korean', 'Language', 'codes', 'ISO', '639-1', 'ko']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Klingon', 'Total', 'speakers', 'Unknown', '.', 'Around', '12', 'fluent',
'speakers', 'in', '1996', ',', 'according', 'to', 'Lawrence', 'Schoen',
',', 'director', 'of', 'the', 'KLI']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'Unknown', '.', 'Around', 'NUM', 'fluent', 'speakers', 'in', 'NUM', ',',

```

'according', 'to', 'Lawrence', 'Schoen', ',', 'director', 'of', 'the',  
'KLI']], ['NUM', '12'], ['NUM', '1996']], 21, True)

['Klingon', 'Date', 'founded', '1984']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Klingon', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']], 7, True)

['Klingon', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']], 7, True)

['Klingon', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']], 7, True)

['Klingon', 'Category', '(', 'purpose', ')', 'Klingon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']], 6, True)

['Klingon', 'Writing', 'system', 'a', 'priori', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Klingon', 'Pronunciation', '/', '\\xcb', '\\x88', 't', '\\xcd', '\\xa1',  
'\\xc9', '\\xac', '\\xc9', '\\xaa', '\\xc5', '\\x8b', '\\xc9', '\\x91', 'n', 'xol',  
'/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'\\xcb', '\\x88', 't', '\\xcd', '\\xa1', '\\xc9', '\\xac', '\\xc9', '\\xaa',  
'\\xc5', '\\x8b', '\\xc9', '\\x91', 'n', 'xol', '/']], 19, True)

['Klingon', 'Created', 'by', 'Marc', 'Okrand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Klingon', 'Language', 'codes', 'ISO', '639-2', 'tlh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Klingon', 'Language', 'codes', 'ISO', '639-3', 'tlh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kashubian', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kashubian', 'Region', 'Pomerania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kashubian', 'Official', 'language', 'in', 'In', 'official', 'use']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'In', 'official', 'use']]], 7, True)

['Kashubian', 'Official', 'language', 'in', 'as', 'a', 'regional', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'as', 'a', 'regional', 'language']]], 8, True)

['Kashubian', 'Official', 'language', 'in', 'in', 'some', 'communes', 'of', 'Pomeranian', 'Voivodeship']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'in', 'some', 'communes', 'of', 'Pomeranian', 'Voivodeship']]], 10, True)

['Kashubian', 'Official', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Kashubian', 'Language', 'codes', 'ISO', '639-2', 'csb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kashubian', 'Language', 'codes', 'ISO', '639-3', 'csb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kashubian', 'Language', 'codes', 'Linguasphere', '53-AAA-cb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Latin', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Latin', 'Pronunciation', 'la', '\xcb', '\x88', 'ti', '\xcb', '\x90', 'na']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'la',

'\xcb', '\x88', 'ti', '\xcb', '\x90', 'na']]], 9, True)

['Latin', 'Spoken', 'in', 'Vatican', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Latin', 'Official', 'language', 'in', 'Holy', 'See']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Latin', 'Regulated', 'by', 'In', 'antiquity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Latin', 'Regulated', 'by', 'Roman', 'schools', 'of', 'grammar', 'and', 'rhetoric', '.', 'Today']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Roman', 'schools', 'of', 'grammar', 'and', 'rhetoric', '.', 'Today']]], 11, True)

['Latin', 'Regulated', 'by', 'Opus', 'Fundatum', 'Latinitas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Opus', 'Fundatum', 'Latinitas']]], 6, True)

['Latin', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Latin', 'Language', 'codes', 'ISO', '639-3', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Latin', 'Language', 'codes', 'Linguasphere', '51-AAB-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Latin', 'Language', 'codes', 'Linguasphere', 'and', 'Koine', 'Greek']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'and', 'Koine', 'Type: ANY']]], 7, True)

['Latin', 'Language', 'codes', 'Linguasphere', 'Coptic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Latin', 'Language', 'codes', 'Linguasphere', 'Syriac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'Linguasphere', 'Type: ANY']]], 5, True)

['Latin', 'Language', 'codes', 'Linguasphere', 'and', 'other', 'native',  
'languages', 'dominated', 'the', 'eastern', 'half']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'and', 'other', 'native', 'languages', 'dominated', 'the',  
'eastern', 'half']]], 12, True)

['Latin', 'Language', 'codes', 'ISO', '639-1', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['West', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['West', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['West', 'Spoken', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['West', 'Total', 'speakers', '\xe2', '\x89', '\x88', '4', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x89', '\x88', 'NUM', 'million']], ['NUM', '4']], 8, True)

['West', 'Language', 'codes', 'ISO', '639-2', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Loglan', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Loglan', 'Date', 'founded', '1955']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Loglan', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',



'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Loglan', 'Category', '(', 'purpose', ')', 'engineered', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Loglan', 'Category', '(', 'purpose', ')', 'logical', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Loglan', 'Category', '(', 'purpose', ')', 'Loglan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Loglan', 'Created', 'by', 'James', 'Cooke', 'Brown']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'James',  
'Cooke', 'Brown']]], 6, True)

['Loglan', 'Setting', 'and', 'usage', 'engineered', 'language', 'for',  
'testing', 'the', 'Sapir', '\\e2', '\\x80', '\\x93', 'Whorf', 'Hypothesis',  
'and', 'other', 'linguistic', 'research']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'engineered', 'language', 'for', 'testing', 'the', 'Sapir',  
'\\e2', '\\x80', '\\x93', 'Whorf', 'Hypothesis', 'and', 'other',  
'linguistic', 'research']]], 19, True)

['Loglan', 'Language', 'codes', 'ISO', '639-3', '\\e2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\e2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Luxembourgish', 'Region', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Luxembourgish', 'Spoken', 'in', 'Luxembourg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Luxembourgish', 'Spoken', 'in', 'Belgium']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Luxembourgish', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Luxembourgish', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Luxembourgish', 'Pronunciation', '\xcb', '\x88', 'l', '\xc5', '\x93',  
't', '\xcd', '\xa1', 's', '\xc9', '\x99', 'bu', '\xc9', '\x90', '\xcc',  
'\xaf', 'j', '\xc9', '\x99', '\xca', '\x83']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'l', '\xc5', '\x93', 't', '\xcd', '\xa1', 's', '\xc9', '\x99',  
'bu', '\xc9', '\x90', '\xcc', '\xaf', 'j', '\xc9', '\x99', '\xca',  
'\x83']]], 23, True)
```

```
['Luxembourgish', 'Official', 'language', 'in', 'Luxembourg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Luxembourgish', 'Regulated', 'by', 'Conseil', 'Permanent', 'de', 'la',  
'Langue', 'Luxembourgeoise', '(', 'CPLL', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Conseil', 'Permanent', 'de', 'la', 'Langue', 'Luxembourgeoise', '\\(',  
'CPLL', '\\)']]], 12, True)
```

```
['Luxembourgish', 'Language', 'codes', 'ISO', '639-2', 'ltz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Luxembourgish', 'Language', 'codes', 'ISO', '639-3', 'ltz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Luxembourgish', 'Language', 'codes', 'ISO', '639-1', 'lb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Maltese', 'Spoken', 'in', 'Malta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Maltese', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Maltese', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Maltese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Maltese', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Maltese', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Maltese', 'Writing', 'system', 'Maltese-variant', 'of', 'the', 'Latin',  
'Alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Maltese', 'Official', 'language', 'in', 'Malta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Maltese', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Maltese', 'Regulated', 'by', 'Il-Kunsill', 'Nazżjonali', 'tal-Ilsien',  
'Malti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Il-  
Kunsill', 'Nazżjonali', 'tal-Ilsien', 'Malti']]], 7, True)
```

```
['Maltese', 'Language', 'codes', 'ISO', '639-2', 'mlt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Maltese', 'Language', 'codes', 'ISO', '639-3', 'mlt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Maltese', 'Language', 'codes', 'Linguasphere', '12-AAC-c']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Maltese', 'Language', 'codes', 'ISO', '639-1', 'mt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

['Macedonian', 'Region', 'Balkans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Macedonian', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']]], 6, True)

['Macedonian', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macedonian', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macedonian', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macedonian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macedonian', 'Spoken', 'in', 'Macedonian', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Macedonian', 'Writing', 'system', 'Cyrillic', '(', 'Macedonian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Macedonian', 'Pronunciation', 'ma', '\xcb', '\x88', 'k', '\xc9', '\x9b', 'd', '\xc9', '\x94', 'nski']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ma', '\xcb', '\x88', 'k', '\xc9', '\x9b', 'd', '\xc9', '\x94', 'nski']]], 12, True)

['Macedonian', 'Official', 'language', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'Macedonia']]], 7, True)

['Macedonian', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in', 'parts', 'of']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in', 'parts', 'of']]], 11, True)

['Macedonian', 'Official', 'language', 'in', ':', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Macedonian', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Macedonian', 'Official', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Macedonian', 'Language', 'codes', 'ISO', '639-2', 'mac', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mac', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Macedonian', 'Language', 'codes', 'ISO', '639-2', 'mkd', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mkd', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Macedonian', 'Language', 'codes', 'ISO', '639-3', 'mkd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Macedonian', 'Language', 'codes', 'Linguasphere', 'Countries', 'with', 'significant', 'Macedonian-speaking', 'populations']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Countries', 'with', 'significant', 'Macedonian-speaking', 'populations']]], 9, True)

['Macedonian', 'Language', 'codes', 'Linguasphere', '(', 'Click', 'on', 'image', 'for', 'the', 'legend', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', '\\\\(', 'Click', 'on', 'image', 'for', 'the', 'legend', '\\\\)']]], 12, True)

['Macedonian', 'Language', 'codes', 'ISO', '639-1', 'mk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Mandarin', 'Total', 'speakers', 'Native', ':', '845', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',

'Native', ':', 'NUM', 'million']], ['NUM', '845']], 7, True)

['Mandarin', 'Dialects', 'Northeastern']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Beijing']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Ji-Lu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Jiao-Liao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Lower', 'Yangtze']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Dialects', 'Central', 'Plains']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Dialects', 'Lan-Yin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Southwestern']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Spoken', 'in', 'Taiwan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Spoken', 'in', 'Singapore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Spoken', 'in', 'Malaysia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-3', 'cmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Manx', 'Total', 'speakers', 'Extinct', 'as', 'a', 'first', 'language',  
'in', '1974', ';', 'subsequently', 'revived', 'and', 'now', 'with',  
'about', 'a', 'hundred', 'competent', 'speakers', ',', 'including', 'a',  
'small', 'number', 'of', 'children', 'who', 'are', 'new', 'native',  
'speakers', ',', ']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Extinct', 'as', 'a', 'first', 'language', 'in', 'NUM', ';',  
'subsequently', 'revived', 'and', 'now', 'with', 'about', 'a', 'hundred',  
'competent', 'speakers', ',', 'including', 'a', 'small', 'number', 'of',  
'children', 'who', 'are', 'new', 'native', 'speakers', ',', ']]], ['NUM',  
'1974']], 34, True)

['Manx', 'Pronunciation', '\\xc9', '\\x99', '\\xcb', '\\x88', '\\xc9', '\\xa3',  
'ilk', ',', '\\xc9', '\\x99', '\\xcb', '\\x88', '\\xc9', '\\xa3', 'il', '\\xc9',  
'\\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xc9',  
'\\x99', '\\xcb', '\\x88', '\\xc9', '\\xa3', 'ilk', ',', '\\xc9', '\\x99', '\\xcb',  
'\\x88', '\\xc9', '\\xa3', 'il', '\\xc9', '\\xa1']]], 19, True)

['Manx', 'Spoken', 'in', 'Isle', 'of', 'Man']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle',  
'of', 'Man']]], 6, True)

['Manx', 'Official', 'language', 'in', 'Isle', 'of', 'Man']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Isle', 'of', 'Man']]], 7, True)

['Manx', 'Language', 'codes', 'ISO', '639-2', 'glv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Manx', 'Language', 'codes', 'ISO', '639-3', 'glv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Manx', 'Language', 'codes', 'Linguasphere', '50-AAA-aj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Manx', 'Language', 'codes', 'ISO', '639-1', 'gv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Marathi', 'Total', 'speakers', 'Total', '90', 'million', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Total', 'NUM', 'million', 'speakers']], ['NUM', '90']], 7, True)

['Marathi', 'Total', 'speakers', '70', 'million', 'native', ',', '20', 'million', 'second', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'native', ',', 'NUM', 'million', 'second', 'language']], ['NUM', '70'], ['NUM', '20']], 11, True)

['Marathi', 'Region', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Marathi', 'Region', 'parts', 'of', 'Goa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Marathi', 'Region', 'Gujarat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Marathi', 'Region', 'Madhya', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Region', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Marathi', 'Region', 'Andhra', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



['Marathi', 'Spoken', 'in', 'India', ',', 'Mauritius', 'and', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
,', 'Mauritius', 'and', 'Israel']]], 8, True)

['Marathi', 'Spoken', 'in', 'Marathi-speaking', 'populations', 'are',  
'found', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Marathi-  
speaking', 'populations', 'are', 'found', 'in', 'United', 'States']]], 10,  
True)

['Marathi', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Guyana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Trinidad', 'and', 'Tobago']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Marathi', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Marathi', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marathi', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marathi', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marathi', 'Writing', 'system', 'Devanagari', 'script', '(', 'standard', ')', ',', 'Modi', 'script', '(', 'shorthand', 'script', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devanagari', 'script', '\\\\(', 'standard', '\\\\)', ',', 'Modi', 'script', '\\\\(', 'shorthand', 'script', '\\\\)']], 15, True)

['Marathi', 'Pronunciation', 'm', '\\xc9', '\\x99', '\\xcb', '\\x88', '\\xc9', '\\xbe', 'a', '\\xca', '\\x88', '\\xca', '\\xb0', 'i']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm', '\\xc9', '\\x99', '\\xcb', '\\x88', '\\xc9', '\\xbe', 'a', '\\xca', '\\x88', '\\xca', '\\xb0', 'i']], 15, True)

['Marathi', 'Official', 'language', 'in', 'India', '(', 'State', 'of', 'Maharashtra', ',', 'Union', 'territories', 'of', 'Daman-Diu', ')', 'and', 'Dadra', 'Nagar', 'Haveli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\\\(', 'State', 'of', 'Maharashtra', ',', 'Union', 'territories', 'of', 'Daman-Diu', '\\\\)', 'and', 'Dadra', 'Nagar', 'Haveli']], 19, True)

['Marathi', 'Language', 'codes', 'ISO', '639-2', 'mar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Marathi', 'Language', 'codes', 'ISO', '639-3', 'mar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marathi', 'Language', 'codes', 'ISO', '639-1', 'mr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Mobilian', 'Language', 'codes', 'ISO', '639-3', 'mod']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Total', 'speakers', '5', 'million', 'Norwegians']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'Norwegians']], ['NUM', '5']], 6, True)

['Norwegian', 'Writing', 'system', 'Latin', '(', 'Norwegian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Norwegian', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']], 8, True)

['Norwegian', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Norwegian', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-2']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nor', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',

'639-3']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nob', '\xe2',  
\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5',  
'l']], ['NUM', '639-3']], 14, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nno', '\xe2', '\x80',  
\x93', 'Nynorsk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-3']],  
10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'no', '\xe2', '\x80',  
\x93', 'Norwegian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-1']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nb', '\xe2', '\x80',  
\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']],  
['NUM', '639-1']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nn', '\xe2', '\x80',  
\x93', 'Nynorsk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-1']],  
10, True)

['Northern', 'Dialects', 'Lobedu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Northern', 'Region', 'Gauteng']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Northern', 'Region', 'Limpopo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY'']], 3, True)

['Northern', 'Region', 'Mpumalanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Northern', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Northern', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-2', 'nso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'nso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['New', 'Spoken', 'in', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['New', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['New', 'Language', 'codes', 'ISO', '639-3', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['New', 'Language', 'codes', 'ISO', '639-3', 'European', 'countries', 'with', 'a', 'Neo-Latin', 'literary', 'tradition']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'European', 'countries', 'with', 'a', 'Neo-Latin', 'literary', 'tradition']], ['NUM', '639-3']], 12, True)

['New', 'Language', 'codes', 'ISO', '639-1', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Novial', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'\xe2', '\x80', '\x93']]], 6, True)

['Novial', 'Date', 'founded', '1928']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Novial', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY', 'language']], 7, True)

['Novial', 'Category', '(', 'purpose', ')', 'international', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'international', 'auxiliary', 'language']], 8, True)

['Novial', 'Category', '(', 'purpose', ')', 'Novial']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY']], 6, True)

['Novial', 'Created', 'by', 'Otto', 'Jespersen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Novial', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']], 7, True)

['Novial', 'Language', 'codes', 'ISO', '639-3', 'nov']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Novial', 'Language', 'codes', 'Linguasphere', '51-AAB-dc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Norn', 'Spoken', 'in', 'Shetland', ',', 'Orkney', 'and', 'Caithness']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Shetland', ',', 'Orkney', 'and', 'Caithness']], 8, True)

['Norn', 'Language', 'codes', 'ISO', '639-3', 'nrn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norn', 'Language', 'codes', 'Linguasphere', '52-AAA-ac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'Linguasphere', 'Type: ANY']]], 5, True)

['Occitan', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Occitan', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Occitan', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Occitan', 'Spoken', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Occitan', 'Dialects', 'Auvergnat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Occitan', 'Dialects', 'Gascon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Occitan', 'Dialects', 'Languedocian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Occitan', 'Dialects', 'Limousin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Occitan', 'Dialects', 'Proven', '\xc3', '\xa7', 'al']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Proven', '\xc3', '\xa7', 'al']]], 6, True)

['Occitan', 'Dialects', 'Vivaro-Alpine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Occitan', 'Regulated', 'by', 'Conselh', 'de', 'la', 'lengua', 'occitana', '(', 'in', 'progress', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Conselh', 'de', 'la', 'lengua', 'occitana', '\\(', 'in', 'progress', '\\

\)''''], 12, True)

['Occitan', 'Language', 'codes', 'ISO', '639-2', 'oci']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Occitan', 'Language', 'codes', 'ISO', '639-3', 'oci']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Occitan', 'Language', 'codes', 'ISO', '639-1', 'oc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Prussian', 'Spoken', 'in', 'Prussia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Prussian', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Prussian', 'Language', 'codes', 'ISO', '639-3', 'prg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Prussian', 'Language', 'codes', 'Linguasphere', '54-AAC-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Old', 'Spoken', 'in', 'Nordic', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Old', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Old', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Old', 'Spoken', 'in', 'England', 'and', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', 'Wales']], 6, True)



```
['Old', 'Spoken', 'in', 'Isle', 'of', 'Man']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle',  
'of', 'Man']]], 6, True)
```

```
['Old', 'Spoken', 'in', 'Normandy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Old', 'Spoken', 'in', 'Vinland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Old', 'Spoken', 'in', 'the', 'Volga', 'and', 'places', 'in-between']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the',  
'Volga', 'and', 'places', 'in-between']]], 8, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'non']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'non']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'England', '(', 'except', 'the', 'extreme',  
'southwest', 'and', 'northwest', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'England', '\\\\(', 'except', 'the', 'extreme', 'southwest', 'and',  
'northwest', '\\\\)']]], 12, True)
```

```
['Old', 'Spoken', 'in', 'parts', 'of', 'modern', 'Scotland', 'south-east',  
'of', 'the', 'Forth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',  
'of', 'modern', 'Scotland', 'south-east', 'of', 'the', 'Forth']]], 11,  
True)
```

```
['Old', 'Spoken', 'in', 'and', 'the', 'eastern', 'fringes', 'of', 'modern',  
'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',  
'the', 'eastern', 'fringes', 'of', 'modern', 'Wales']]], 10, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'ang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'ang']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Occidental', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Occidental', 'Category', '(', 'purpose', ')', 'constructed', 'language', 'based', 'on', 'Indo-European', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'constructed', 'language', 'based', 'on', 'Indo-European', 'languages']], 11, True)

['Occidental', 'Language', 'codes', 'ISO', '639-2', 'ile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Occidental', 'Language', 'codes', 'ISO', '639-3', 'ile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Occidental', 'Language', 'codes', 'ISO', '639-1', 'ie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Polish', 'Total', 'speakers', '40', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '40']], 5, True)

['Polish', 'Spoken', 'in', 'Poland', '.', 'Minorities', ':', 'Belarus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Poland', '.', 'Minorities', ':', 'Belarus']], 8, True)

['Polish', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Polish', 'Spoken', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Polish', 'Spoken', 'in', 'Latvia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Polish', 'Spoken', 'in', 'United', 'Kingdom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Polish', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Polish', 'Spoken', 'in', 'Czech', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Polish', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Writing', 'system', 'Latin', '(', 'Polish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Polish', 'Pronunciation', '\xcb', '\x88', 'p', '\xc9', '\x94', 'lski']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'p', '\xc9', '\x94', 'lski']]], 8, True)

['Polish', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Polish', 'Official', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Official', 'language', 'in', 'Minority', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Polish', 'Official', 'language', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Official', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Language', 'codes', 'ISO', '639-2', 'pol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Polish', 'Language', 'codes', 'ISO', '639-3', 'pol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Polish', 'Language', 'codes', 'ISO', '639-1', 'pl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Portuguese', 'Region', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Portuguese', 'Region', 'the', 'Americas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Portuguese', 'Region', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Portuguese', 'Region', 'Europe', 'and', 'Oceania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY']]], 5, True)

['Portuguese', 'Spoken', 'in', 'See', 'geographic', 'distribution', 'of', 'Portuguese']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'See', 'geographic', 'distribution', 'of', 'Type: ANY']]], 8, True)

['Portuguese', 'Pronunciation', 'pu', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\xca', '\x83', '(', 'European', ')', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pu', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\xca', '\x83', '\\(', 'European', '\\)', ',']]], 17, True)

['Portuguese', 'Pronunciation', 'po', '\xcf', '\x87', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '(', 'j', ')', '\xca', '\x83', '(', 'BP-carioca', ')', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xcf', '\x87', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\\(', 'j', '\\)', '\xca', '\x83', '\\(', 'BP-carioca', '\\)', ',']]], 20, True)

['Portuguese', 'Pronunciation', 'po', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '(', 'j', ')', 's', '(', 'BP-paulistano', ')', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\\(', 'j', '\\)', 's', '\\(', 'BP-paulistano', '\\)', ',']]], 19, True)

['Portuguese', 'Pronunciation', 'po', '\xc9', '\xb9', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'ejs', '(', 'BP-caipira', ')', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xc9', '\xb9', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'ejs', '\\(', 'BP-caipira', '\\)', ',']]], 15, True)

['Portuguese', 'Pronunciation', 'po', '\xcf', '\x87', '(', 'h', ')', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '(', 'j', ')', 's', '(', 'BP-mineiro', ')', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xcf', '\x87', '\\(', 'h', '\\)', 'tu', '\xcb', '\x88', '\xc9', '\xa1',

'e', '\\(', 'j', '\\)', 's', '\\(', 'BP-mineiro', '\\)', ',']]], 22, True)

['Portuguese', 'Pronunciation', 'p', '\\xc9', '\\x94', 'htu', '\\xcb', '\\x88', '\\xc9', '\\xa1', 'ejs', '(', 'BP-nordestino', ')', ',']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\\xc9', '\\x94', 'htu', '\\xcb', '\\x88', '\\xc9', '\\xa1', 'ejs', '\\(', 'BP-nordestino', '\\)', ',']]], 15, True)

['Portuguese', 'Official', 'language', 'in', '8', 'countries']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Portuguese', 'Official', 'language', 'in', 'Angola']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Brazil']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Cape', 'Verde']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Portuguese', 'Official', 'language', 'in', 'East', 'Timor']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Portuguese', 'Official', 'language', 'in', 'Guinea-Bissau']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Mozambique']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Portugal']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'S', '\\xc3', '\\xa3', 'o', 'Tom', '\\xc3', '\\xa9', 'and', 'Pr', '\\xc3', '\\xad', 'ncipe']]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'S', '\\xc3', '\\xa3', 'o', 'Tom', '\\xc3', '\\xa9', 'and', 'Pr', '\\xc3', '\\xad', 'ncipe']]], 16, True)

['Portuguese', 'Official', 'language', 'in', '1', 'dependent', 'entity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'NUM', 'dependent', 'entity']], ['NUM', '1']], 7, True)

['Portuguese', 'Official', 'language', 'in', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Numerous', 'international',  
'organisations']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Numerous', 'international', 'organisations']], 7, True)

['Portuguese', 'Language', 'codes', 'ISO', '639-2', 'por']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Portuguese', 'Language', 'codes', 'ISO', '639-3', 'por']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Portuguese', 'Language', 'codes', 'Linguasphere', '51-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Portuguese', 'Language', 'codes', 'ISO', '639-1', 'pt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Standard', 'Spoken', 'in', 'Republic', 'of', 'China', '(', 'Taiwan',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'China', '\\(', 'Taiwan', '\\)']], 9, True)

['Standard', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Standard', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']], 6, True)

['Standard', 'Official', 'language', 'in', 'Republic', 'of', 'China', '(',  
'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Republic', 'of', 'China', '\\(', 'Taiwan', '\\)']], 10, True)

['Standard', 'Regulated', 'by', 'In', 'the', 'ROC', ':', 'National',  
'Languages', 'Committee']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'In',  
'the', 'ROC', ':', 'National', 'Languages', 'Committee']]], 10, True)

['Standard', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Punjabi', 'Total', 'speakers', '57', ',', '129', ',', '000', '(',  
'Encarta', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'Encarta', '\\)']], ['NUM', '57'],  
['NUM', '129'], ['NUM', '000']], 11, True)

['Punjabi', 'Region', 'Punjab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Punjabi', 'Spoken', 'in', 'Pakistan', 'and', 'India', '.', 'Minor',  
'populations', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Pakistan', 'and', 'India', '.', 'Minor', 'populations', 'in', 'United',  
'Kingdom']]], 12, True)

['Punjabi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Punjabi', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Punjabi', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Punjabi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Punjabi', 'Spoken', 'in', 'United', 'Arab', 'Emirates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'Arab', 'Emirates']]], 6, True)



```

['Punjabi', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Punjabi', 'Writing', 'system', 'Gurmukhi', 'in', 'Punjab', '(', 'India',
')', 'and', 'Sikh', 'diaspora']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Gurmukhi', 'in', 'Punjab', '\\(', 'India', '\\)', 'and', 'Sikh',
'diaspora']]], 12, True)

['Punjabi', 'Writing', 'system', 'Shahmukhi', 'in', 'Punjab', '(',
'Pakistan', ')', 'and', 'Sarhad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Shahmukhi', 'in', 'Punjab', '\\(', 'Pakistan', '\\)', 'and', 'Sarhad']]],
11, True)

['Punjabi', 'Official', 'language', 'in', 'India', '(', 'Indian', 'Punjab',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'India', '\\(', 'Indian', 'Punjab', '\\)'']]], 9, True)

['Punjabi', 'Official', 'language', 'in', 'Pakistan', '(', 'Pakistani',
'Punjab', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Pakistan', '\\(', 'Pakistani', 'Punjab', '\\)'']]], 9, True)

['Punjabi', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Punjabi', 'Language', 'codes', 'ISO', '639-2', 'pan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Punjabi', 'Language', 'codes', 'ISO', '639-3', 'pan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Punjabi', 'Language', 'codes', 'ISO', '639-1', 'pa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Quechua', 'Total', 'speakers', '\\xe2', '\\x89', '\\x88', '10', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\\xe2', '\\x89', '\\x88', 'NUM', 'million']], ['NUM', '10']], 8, True)

```

['Quechua', 'Region', 'Central', 'Andes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quechua', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quechua', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quechua', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quechua', 'Spoken', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quechua', 'Spoken', 'in', 'Chile', 'and', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Quechua', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Quechua', 'Official', 'language', 'in', 'Peru', 'and', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Peru', 'and', 'Bolivia']]], 7, True)

['Quechua', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Quechua', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Quechua', 'Language', 'codes', 'ISO', '639-3', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quechua', 'Language', 'codes', 'ISO', '639-1', 'qu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Russian', 'Total', 'speakers', 'Primary', 'language', ':', 'about', '175', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Primary', 'language', ':', 'about', 'NUM', 'million']], ['NUM', '175']], 9, True)

['Russian', 'Total', 'speakers', 'Total', ':', '300', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Total', ':', 'NUM', 'million']], ['NUM', '300']], 7, True)

['Russian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Spoken', 'in', 'countries', 'of', 'the', 'former', 'Soviet', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'countries', 'of', 'the', 'former', 'Soviet', 'Union']], 9, True)

['Russian', 'Spoken', 'in', 'emigrant', 'communities', 'around', 'the', 'world']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'emigrant', 'communities', 'around', 'the', 'world']], 8, True)

['Russian', 'Spoken', 'in', 'notably', 'the', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'notably', 'the', 'UK']], 6, True)

['Russian', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Spoken', 'in', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the', 'United', 'States']], 6, True)

['Russian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Writing', 'system', 'Cyrillic', '(', 'Russian', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Russian', 'Pronunciation', '\xcb', '\x88', 'rus', '\xca', '\xb2', 'k', '\xca', '\xb2', '\xc9', '\xaa', 'j', 'j', '\xc9', '\xaa', '\xcb', '\x88', 'z', '\xc9', '\xa8', 'k']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'rus', '\xca', '\xb2', 'k', '\xca', '\xb2', '\xc9', '\xaa', 'j', 'j', '\xc9', '\xaa', '\xcb', '\x88', 'z', '\xc9', '\xa8', 'k']], 22, True)

['Russian', 'Official', 'language', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Russian', 'Official', 'language', 'in', 'Belarus', '(', 'co-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'co-official', '\\)']], 8, True)

['Russian', 'Official', 'language', 'in', 'Kazakhstan', '(', 'co-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'co-official', '\\)']], 8, True)

['Russian', 'Official', 'language', 'in', 'Kyrgyzstan', '(', 'co-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'co-official', '\\)']], 8, True)

['Russian', 'Official', 'language', 'in', 'Abkhazia', '(', 'co-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'co-official', '\\)']], 8, True)

['Russian', 'Official', 'language', 'in', 'Gagauzia', '(', 'co-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'co-official', '\\)']], 8, True)

['Russian', 'Official', 'language', 'in', 'Transnistria', '(', 'co-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'co-official', '\\)']], 8, True)

```
['Russian', 'Official', 'language', 'in', 'IAEA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Russian', 'Official', 'language', 'in', 'United', 'Nations']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Russian', 'Recognised', 'minority', 'language', 'in', 'Ukraine', '(',
'regional', 'language', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'Ukraine', '\\(', 'regional', 'language', '\\)']]], 10,
True)
```

```
['Russian', 'Language', 'codes', 'ISO', '639-2', 'rus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Russian', 'Language', 'codes', 'ISO', '639-3', 'rus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Russian', 'Language', 'codes', 'ISO', '639-1', 'ru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Romansh', 'Spoken', 'in', 'Switzerland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Romansh', 'Region', 'Graub', '\\xc3', '\\xbc', 'nden']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Graub',
'\\xc3', '\\xbc', 'nden']]], 6, True)
```

```
['Romansh', 'Official', 'language', 'in', 'Switzerland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Romansh', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Romansh', 'Language', 'codes', 'ISO', '639-2', 'roh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Romansh', 'Language', 'codes', 'ISO', '639-3', 'roh']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Romansh', 'Language', 'codes', 'Linguasphere', '51-AAA-k']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]]], 5, True)
```

```
['Romansh', 'Language', 'codes', 'ISO', '639-1', 'rm']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Romanian,', 'Daco-Romanian', 'Total', 'speakers', 'First', 'language',
':', '24', 'million']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Daco-Romanian', 'Total',
'speakers', 'First', 'language', ':', 'NUM', 'million']], ['NUM', '24']],
9, True)
```

```
['Romanian,', 'Total', 'speakers', 'Second', 'language', ':', '4',
'million']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'Second', 'language', ':', 'NUM', 'million']], ['NUM', '4']], 8, True)
```

```
['Romanian,', 'Spoken', 'in', 'By', 'a', 'majority']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'By',
'a', 'majority']]]], 6, True)
```

```
['Romanian,', 'Spoken', 'in', ':', 'Romania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',
'Type: ANY']]]], 5, True)
```

```
['Romanian,', 'Spoken', 'in', 'Republic', 'of', 'Moldova']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Republic', 'of', 'Moldova']]]], 6, True)
```

```
['Romanian,', 'Spoken', 'in', 'Minority', 'speakers', 'in']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Minority', 'speakers', 'in']]]], 6, True)
```

```
['Romanian,', 'Spoken', 'in', ':', 'Albania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',
'Type: ANY']]]], 5, True)
```

```
['Romanian,', 'Spoken', 'in', 'Bulgaria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

['Romanian,', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romanian,', 'Spoken', 'in', 'Migrant', 'speakers', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Migrant', 'speakers', 'in']], 6, True)

['Romanian,', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'New', 'Zealand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Romanian,', 'Pronunciation', 'ro', '\xcb', '\x88', 'm', '\xc9', '\xa8', 'n', '\xc9', '\x99']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ro', '\xcb', '\x88', 'm', '\xc9', '\xa8', 'n', '\xc9', '\x99']]], 11, True)

['Romanian,', 'Official', 'language', 'in', 'Romania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romanian,', 'Official', 'language', 'in', 'Moldova']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romanian,', 'Official', 'language', 'in', 'Vojvodina', '(', 'Serbia', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Vojvodina', '\\(', 'Serbia', '\\)']]], 8, True)

['Romanian,', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Romanian,', 'Official', 'language', 'in', 'Mount', 'Athos']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Romanian,', 'Official', 'language', 'in', 'Latin', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Romanian,', 'Regulated', 'by', 'Academia', 'Rom', '\xc3', '\xa2', 'n', '\xc4', '\x83']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Academia', 'Rom', '\xc3', '\xa2', 'n', '\xc4', '\x83']]], 10, True)

['Romanian,', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-2', 'rum', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rum', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-2', 'ron', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ron', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-3', 'ron']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-1', 'ro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Swedish', 'Writing', 'system', 'Latin', '(', 'Swedish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Swedish', 'Pronunciation', '\xcb', '\x88', 'sv', '\xc9', '\x9b', '\xcc', '\x82', 'nsk', '\xc3', '\xa2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'sv', '\xc9', '\x9b', '\xcc', '\x82', 'nsk', '\xc3', '\xa2']]], 12, True)

['Swedish', 'Official', 'language', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

```
['Swedish', 'Official', 'language', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Swedish', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Swedish', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Swedish', 'Regulated', 'by', 'Research', 'Institute', 'for', 'the',  
'Languages', 'of', 'Finland', '(', 'in', 'Finland', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Research', 'Institute', 'for', 'the', 'Languages', 'of', 'Finland', '\\\\(',  
'in', 'Finland', '\\\\)']]], 14, True)
```

```
['Swedish', 'Language', 'codes', 'ISO', '639-2', 'swe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Swedish', 'Language', 'codes', 'ISO', '639-3', 'swe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Swedish', 'Language', 'codes', 'ISO', '639-1', 'sv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Spanish,', 'Castilian', 'Total', 'speakers', 'First', 'language', '329',  
'million', 'to', '400', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Castilian', 'Total',  
'speakers', 'First', 'language', 'NUM', 'million', 'to', 'NUM',  
'million']]], ['NUM', '329'], ['NUM', '400']], 11, True)
```

```
['Spanish,', 'Total', 'speakers', '500', 'million', 'as', 'first', 'or',  
'second', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', 'as', 'first', 'or', 'second', 'language']]], ['NUM',  
'500']], 10, True)
```

```
['Spanish,', 'Spoken', 'in', '(', 'see', 'below', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\\\(',  
'see', 'below', '\\\\)']]], 7, True)
```

```
['Spanish,', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]]], 8, True)
```

```
['Spanish,', 'Pronunciation', 'espa', '\xcb', '\x88', '\xc9', '\xb2', 'ol',  
,', 'kaste', '\xcb', '\x88', '\xca', '\x8e', 'ano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'espa',  
'\xcb', '\x88', '\xc9', '\xb2', 'ol', ',', 'kaste', '\xcb', '\x88', '\xca',  
'\x8e', 'ano']]]], 15, True)
```

```
['Spanish,', 'Official', 'language', 'in', '21', 'countries']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'United', 'Nations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'European', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'Organization', 'of', 'Ibero-  
American', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Organization', 'of', 'Ibero-American', 'States']]]], 8, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'African', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'Caricom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'World', 'Trade',  
'Organization']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'World', 'Trade', 'Organization']]]], 7, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'Andean', 'Community', 'of',  
'Nations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Andean', 'Community', 'of', 'Nations']]]], 8, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'Mercosur']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Spanish,', 'Official', 'language', 'in', 'Inter-American', 'Development', 'Bank']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Inter-American', 'Development', 'Bank']]], 7, True)

['Spanish,', 'Official', 'language', 'in', 'Latin', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Spanish,', 'Official', 'language', 'in', 'Antarctic', 'Treaty']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Spanish,', 'Language', 'codes', 'ISO', '639-2', 'spa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Spanish,', 'Language', 'codes', 'ISO', '639-3', 'spa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Spanish,', 'Language', 'codes', 'Linguasphere', '51-AAA-b']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Spanish,', 'Language', 'codes', 'ISO', '639-1', 'es']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Logudorese', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Logudorese', 'Region', 'Sardinia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Logudorese', 'Language', 'codes', 'ISO', '639-3', 'src']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sardinian', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Sardinian', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sardinian', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sardinian', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sardinian', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sardinian', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sardinian', 'Region', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sardinian', 'Official', 'language', 'in', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-2', 'srd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'srd', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'srd', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', ':', 'sro', '\xe2', '\x80', '\x93', 'Campidanese']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'sro', '\xe2', '\x80', '\x93', 'Campidanese']], ['NUM', '639-3']], 11, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'sdn', '\xe2', '\x80', '\x93', 'Gallurese']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sdn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'src', '\xe2', '\x80', '\x93', 'Logudorese']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'src', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'sdc', '\xe2', '\x80', '\x93', 'Sassarese']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sdc', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-1', 'sc']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Swiss', 'Region', 'Europe']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Swiss', 'Spoken', 'in', 'Switzerland', 'Italy']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Swiss', 'Pronunciation', '\xcb', '\x88', '\xca', '\x83', '\xca', '\x8b', 'its', '\xc9', '\x99', 'r', '\xcb', '\x8c', 'd', '\xcc', '\xa5', 'yt', '\xca', '\x83']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', '\xca', '\x83', '\xca', '\x8b', 'its', '\xc9', '\x99', 'r', '\xcb', '\x8c', 'd', '\xcc', '\xa5', 'yt', '\xca', '\x83']], 20, True)

['Swiss', 'Language', 'codes', 'ISO', '639-2', 'gsw']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Swiss', 'Language', 'codes', 'ISO', '639-3', 'gsw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sanskrit', 'Spoken', 'in', 'Greater', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Sanskrit', 'Writing', 'system', 'Devan', '\xc4', '\x81', 'gar', '\xc4', '\xab', '(', 'de', 'facto', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devan', '\xc4', '\x81', 'gar', '\xc4', '\xab', '\\(', 'de', 'facto', '\\)']], 13, True)

['Sanskrit', 'Writing', 'system', 'various', 'Br', '\xc4', '\x81', 'hm', '\xc4', '\xab', '\xe2', '\x80', '\x93', 'based', 'scripts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'various', 'Br', '\xc4', '\x81', 'hm', '\xc4', '\xab', '\xe2', '\x80', '\x93', 'based', 'scripts']], 15, True)

['Sanskrit', 'Writing', 'system', 'and', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Sanskrit', 'Pronunciation', 's', '\xc9', '\x99', '\xcc', '\x83', 'sk', '\xc9', '\xb9', '\xcc', '\xa9', 't', '\xcc', '\xaa', '\xc9', '\x99', 'm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's', '\xc9', '\x99', '\xcc', '\x83', 'sk', '\xc9', '\xb9', '\xcc', '\xa9', 't', '\xcc', '\xaa', '\xc9', '\x99', 'm']], 18, True)

['Sanskrit', 'Official', 'language', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Sanskrit', 'Official', 'language', 'in', 'one', 'of', 'the', '22', 'scheduled', 'languages', 'of', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'one', 'of', 'the', 'NUM', 'scheduled', 'languages', 'of', 'India']], ['NUM', '22']], 12, True)

['Sanskrit', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Sanskrit', 'Language', 'codes', 'ISO', '639-2', 'san']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sanskrit', 'Language', 'codes', 'ISO', '639-3', 'san']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sanskrit', 'Language', 'codes', 'ISO', '639-1', 'sa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Serbo-Croatian', 'Dialects', '\xc4', '\x8c', 'akavian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xc4',  
'\x8c', 'akavian']], 5, True)

['Serbo-Croatian', 'Dialects', 'Kajkavian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Serbo-Croatian', 'Dialects', '\xc5', '\xa0', 'tokavian', '(', 'standard',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xc5',  
'\xa0', 'tokavian', '\\(', 'standard', '\\)']], 8, True)

['Serbo-Croatian', 'Dialects', 'Torlakian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Serbo-Croatian', 'Region', 'Serb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Serbo-Croatian', 'Region', 'Croat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Serbo-Croatian', 'Region', 'Bosniak']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Serbo-Croatian', 'Region', 'Montenegrin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Serbo-Croatian', 'Spoken', 'in', 'Serbia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)



```
['Serbo-Croatian', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)
```

```
['Serbo-Croatian', 'Spoken', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Spoken', 'in', 'Kosovoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Serbo-Croatian', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Serbo-Croatian', 'Standard', 'forms', 'Serbian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Standard', 'forms', 'Croatian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Standard', 'forms', 'Bosnian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Standard', 'forms', 'Montenegrin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-2', 'scr', ',', 'scc',  
'(', 'deprecated', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'scr', ',', 'scc', '\\\\(', 'deprecated', '\\\\)']], ['NUM',  
'639-2']], 11, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'hbs', '\xe2',
'\x80', '\x93', 'Macrolanguage']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'hbs', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'individual',
'codes']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', ':', 'srp', '\xe2',
'\x80', '\x93', 'Serbian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'srp', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 11, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'hrv', '\xe2',
'\x80', '\x93', 'Croatian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'hrv', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'bos', '\xe2',
'\x80', '\x93', 'Bosnian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'bos', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'Linguasphere', '53-AAA-g']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-1', 'sh', '(',
'deprecated', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'sh', '\\(', 'deprecated', '\\)']], ['NUM', '639-1']], 9,
True)
```

```
['Slovene/Slovenian', 'Dialects', 'Prekmurje', 'dialect']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Slovene/Slovenian', 'Dialects', 'Resian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

['Slovene/Slovenian', 'Dialects', 'approx', '.', '32', 'unstandardised', 'dialects']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'approx', '.', 'NUM', 'unstandardised', 'dialects']], ['NUM', '32']], 7, True)

['Slovene/Slovenian', 'Spoken', 'in', 'Slovenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Slovene/Slovenian', 'Spoken', 'in', 'Italy', '(', 'in', 'Friuli', 'Venezia', 'Giulia', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy', '\\(', 'in', 'Friuli', 'Venezia', 'Giulia', '\\)']], 10, True)

['Slovene/Slovenian', 'Spoken', 'in', 'Austria', '(', 'in', 'Carinthia', 'and', 'Styria', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Austria', '\\(', 'in', 'Carinthia', 'and', 'Styria', '\\)']], 10, True)

['Slovene/Slovenian', 'Spoken', 'in', 'Hungary', '(', 'in', 'Vas', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Hungary', '\\(', 'in', 'Vas', '\\)']], 8, True)

['Slovene/Slovenian', 'Spoken', 'in', 'Croatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Slovene/Slovenian', 'Spoken', 'in', 'and', 'emigrant', 'groups', 'in', 'various', 'countries']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'emigrant', 'groups', 'in', 'various', 'countries']], 9, True)

['Slovene/Slovenian', 'Pronunciation', 'slo', '\\xcb', '\\x88', '\\xca', '\\x8b', 'enski', '\\xcb', '\\x88', 'j', '\\xc9', '\\x9b', 'zik', ',', 'slo', '\\xcb', '\\x88', '\\xca', '\\x8b', 'en', '\\xca', '\\x83', 't', '\\xca', '\\x83', 'ina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'slo', '\\xcb', '\\x88', '\\xca', '\\x8b', 'enski', '\\xcb', '\\x88', 'j', '\\xc9', '\\x9b', 'zik', ',', 'slo', '\\xcb', '\\x88', '\\xca', '\\x8b', 'en', '\\xca', '\\x83', 't', '\\xca', '\\x83', 'ina']], 27, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'Slovenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Slovene/Slovenian', 'Official', 'language', 'in', ':', 'Austria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Slovene/Slovenian', 'Regulated', 'by', 'Slovenian', 'Academy', 'of',  
'Sciences', 'and', 'Arts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Slovenian', 'Academy', 'of', 'Sciences', 'and', 'Arts']]], 9, True)

['Slovene/Slovenian', 'Language', 'codes', 'ISO', '639-2', 'slv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Slovene/Slovenian', 'Language', 'codes', 'ISO', '639-3', 'slv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Slovene/Slovenian', 'Language', 'codes', 'Linguasphere', 'Slovene-  
speaking', 'areas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Slovene-speaking', 'areas']]], 6, True)

['Slovene/Slovenian', 'Language', 'codes', 'ISO', '639-1', 'sl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Slovak', 'Total', 'speakers', 'over', '7', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'over', 'NUM', 'million']], ['NUM', '7']], 6, True)

['Slovak', 'Region', 'Central', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Slovak', 'Spoken', 'in', 'Slovakia', 'and', 'as', 'a', 'minority',  
'language', 'also', 'in', 'the']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Slovakia', 'and', 'as', 'a', 'minority', 'language', 'also', 'in',  
'the']], 12, True)

['Slovak', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Slovak', 'Spoken', 'in', 'Canada', ',', '']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', '']], 5, True)

['Slovak', 'Spoken', 'in', 'Serbia', ',', '']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', '']], 5, True)

['Slovak', 'Spoken', 'in', 'Hungary']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Slovak', 'Spoken', 'in', 'Bulgaria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Slovak', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Slovak', 'Official', 'language', 'in', 'Slovakia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Slovak', 'Official', 'language', 'in', 'Vojvodina', 'in', 'Serbia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Vojvodina', 'in', 'Serbia']], 7, True)

['Slovak', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']], 6, True)

['Slovak', 'Language', 'codes', 'ISO', '639-2', 'slo', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'slo', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

```
['Slovak', 'Language', 'codes', 'ISO', '639-2', 'slk', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'slk', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)
```

```
['Slovak', 'Language', 'codes', 'ISO', '639-3', 'slk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Slovak', 'Language', 'codes', 'ISO', '639-1', 'sk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Scottish', 'Gaelic', 'Total', 'speakers', '58', ',', '552', 'in',  
'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Gaelic', 'Total',  
'speakers', 'NUM', ',', 'NUM', 'in', 'Scotland']], ['NUM', '58'], ['NUM',  
'552']], 9, True)
```

```
['Scottish', 'Region', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Scottish', 'Region', 'Cape', 'Breton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Scottish', 'Region', 'Nova', 'Scotia', 'and', 'Glengarry', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nova',  
'Scotia', 'and', 'Glengarry', 'County']], 7, True)
```

```
['Scottish', 'Region', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Scottish', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Scottish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Scottish', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Scottish', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Scottish', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Scottish', 'Writing', 'system', 'Latin', '(', 'Gaelic', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '\\(', 'Gaelic', 'variant', '\\)']]], 8, True)
```

```
['Scottish', 'Pronunciation', '\\xcb', '\\x88', 'ka', '\\xcb', '\\x90', 'lik',  
'\\xca', '\\xb2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb',  
'\\x88', 'ka', '\\xcb', '\\x90', 'lik', '\\xca', '\\xb2']]], 10, True)
```

```
['Scottish', 'Official', 'language', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Scottish', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Scottish', 'Language', 'codes', 'ISO', '639-2', 'gla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Scottish', 'Language', 'codes', 'ISO', '639-3', 'gla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Scottish', 'Language', 'codes', 'Linguasphere', '50-AAA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Scottish', 'Language', 'codes', 'ISO', '639-1', 'gd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Swahili', 'Spoken', 'in', 'Burundi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'DR', 'Congo']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Swahili', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Mauritius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swahili', 'Total', 'speakers', 'First', 'language', ':', '5', '\xe2', '\x80', '\x93', '10', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'First', 'language', ':', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']]], ['NUM', '5'], ['NUM', '10']], 12, True)

['Swahili', 'Writing', 'system', 'Latin', ',', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swahili', 'Official', 'language', 'in', 'African', 'Union']:



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Swahili', 'Official', 'language', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Swahili', 'Official', 'language', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Swahili', 'Official', 'language', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Swahili', 'Regulated', 'by', 'Baraza', 'la', 'Kiswahili', 'la', 'Taifa',  
'(', 'Tanzania', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Baraza', 'la', 'Kiswahili', 'la', 'Taifa', '\\(', 'Tanzania', '\\)']]],  
11, True)
```

```
['Swahili', 'Language', 'codes', 'ISO', '639-2', 'swa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Swahili', 'Language', 'codes', 'ISO', '639-3', 'swa', '\\xe2', '\\x80',  
'\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'swa', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)
```

```
['Swahili', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Swahili', 'Language', 'codes', 'ISO', '639-3', ':', 'swc', '\\xe2',  
'\\x80', '\\x93', 'Congo', 'Swahili']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'swc', '\\xe2', '\\x80', '\\x93', 'Congo', 'Type: ANY']],  
['NUM', '639-3']], 12, True)
```

```
['Swahili', 'Language', 'codes', 'ISO', '639-3', 'swh', '\\xe2', '\\x80',  
'\\x93', 'Coastal', 'Swahili']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'swh', '\\xe2', '\\x80', '\\x93', 'Coastal', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

['Swahili', 'Language', 'codes', 'Linguasphere', '99-AUS-m']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Swahili', 'Language', 'codes', 'ISO', '639-1', 'sw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Tamil', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tamil', 'Spoken', 'in', 'Sri', 'Lanka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tamil', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tamil', 'Spoken', 'in', 'where', 'it', 'has', 'official', 'status', ';',  
'with', 'significant', 'minorities', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'where',  
'it', 'has', 'official', 'status', ';', 'with', 'significant',  
'minorities', 'in', 'Canada']]], 14, True)

['Tamil', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tamil', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tamil', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tamil', 'Spoken', 'in', 'Mauritius', 'and', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Tamil', 'Spoken', 'in', 'and', 'emigrant', 'communities', 'around',  
'the', 'world']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',

'emigrant', 'communities', 'around', 'the', 'world']]], 9, True)

['Tamil', 'Writing', 'system', 'Tamil', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Tamil', 'Pronunciation', 't', '\xcc', '\xaa', '\xc9', '\x90', 'm',  
'\xc9', '\xa8', '\xc9', '\xbb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't',  
'\xcc', '\xaa', '\xc9', '\x90', 'm', '\xc9', '\xa8', '\xc9', '\xbb']]], 12,  
True)

['Tamil', 'Official', 'language', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Tamil', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tamil', 'Language', 'codes', 'ISO', '639-2', 'tam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tamil', 'Language', 'codes', 'ISO', '639-3', 'tam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tamil', 'Language', 'codes', 'ISO', '639-1', 'ta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Turkish', 'Dialects', 'Karamanli', 'Turkish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Dialects', 'Cypriot', 'Turkish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Turkish', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Northern', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Turkish', 'Spoken', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Macedonia']]], 6, True)

['Turkish', 'Spoken', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'and', 'by', 'immigrant', 'communities', 'in', 'Austria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'by', 'immigrant', 'communities', 'in', 'Austria']]], 9, True)

['Turkish', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'the', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Turkish', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'Switzerland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turkish', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Turkish', 'Spoken', 'in', 'and', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'the', 'United', 'States']]], 7, True)

['Turkish', 'Pronunciation', '\xcb', '\x88', 'ty', '\xc9', '\xbe', 'kt', '\xca', '\x83', 'e', '(', 'listen', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'ty', '\xc9', '\xbe', 'kt', '\xca', '\x83', 'e', '\\(', 'listen', '\\)']]], 14, True)

['Turkish', 'Official', 'language', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Turkish', 'Official', 'language', 'in', 'Northern', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Turkish', 'Official', 'language', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Turkish', 'Recognised', 'minority', 'language', 'in', 'Kosovo', '(', 'regional', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Kosovo', '\\(', 'regional', '\\)']]], 9, True)

['Turkish', 'Recognised', 'minority', 'language', 'in', 'Macedonia', '(', 'regional', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Macedonia', '\\(', 'regional', '\\)']]], 9, True)

['Turkish', 'Recognised', 'minority', 'language', 'in', 'Romania', '(', 'recognized', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Romania', '\\(', 'recognized', '\\)']]], 9, True)

['Turkish', 'Language', 'codes', 'ISO', '639-2', 'tur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Turkish', 'Language', 'codes', 'ISO', '639-3', 'tur']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turkish', 'Language', 'codes', 'Linguasphere', 'Countries', 'with',  
'significant', 'Turkish-speaking', 'populations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Countries', 'with', 'significant', 'Turkish-speaking',  
'populations']]], 9, True)

['Turkish', 'Language', 'codes', 'Linguasphere', '(', 'Click', 'on',  
'image', 'for', 'the', 'legend', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', '\\\\(', 'Click', 'on', 'image', 'for', 'the', 'legend', '\\  
)']]], 12, True)

['Turkish', 'Language', 'codes', 'ISO', '639-1', 'tr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Setswana', 'Spoken', 'in', 'Botswana']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Setswana', 'Spoken', 'in', 'South', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Setswana', 'Spoken', 'in', 'Zimbabwe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Setswana', 'Spoken', 'in', 'Namibia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Setswana', 'Region', 'Southern', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Setswana', 'Official', 'language', 'in', 'Botswana']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Setswana', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Setswana', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Setswana', 'Language', 'codes', 'ISO', '639-2', 'tsn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Setswana', 'Language', 'codes', 'ISO', '639-3', 'tsn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Setswana', 'Language', 'codes', 'ISO', '639-1', 'tn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tocharian', 'Writing', 'system', 'Tocharian', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Tocharian', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tagalog', 'Total', 'speakers', 'First', 'language', '(', 'in', 'the',  
'Philippines', ')', ':', '25', 'million', 'Second', 'language', ':',  
'over', '60', 'million', 'Overall', ':', '90', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'First', 'language', '\\\\(', 'in', 'the', 'Philippines', '\\\\)', ':', 'NUM',  
'million', 'Second', 'language', ':', 'over', 'NUM', 'million', 'Overall',  
True)

['Tagalog', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tagalog', 'Writing', 'system', 'Historically', 'written', 'in',  
'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Historically', 'written', 'in', 'Type: ANY']]], 7, True)



```
['Tagalog', 'Standard', 'forms', 'Filipino']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tagalog', 'Official', 'language', 'in', 'Alaska', '(', 'minority',  
'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'minority', 'language', '\\)']]], 9, True)
```

```
['Tagalog', 'Official', 'language', 'in', 'California', '(', 'minority',  
'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'minority', 'language', '\\)']]], 9, True)
```

```
['Tagalog', 'Official', 'language', 'in', 'Nevada', '(', 'minority',  
'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'minority', 'language', '\\)']]], 9, True)
```

```
['Tagalog', 'Language', 'codes', 'ISO', '639-2', 'tgl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Tagalog', 'Language', 'codes', 'ISO', '639-3', 'tgl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tagalog', 'Language', 'codes', 'Linguasphere', '31-CKA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Tagalog', 'Language', 'codes', 'ISO', '639-1', 'tl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Urdu', 'Region', 'South', 'Asia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Urdu', 'Spoken', 'in', 'Pakistan', ',', 'India', 'and', 'Bangladesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Pakistan', ',', 'India', 'and', 'Bangladesh']]], 8, True)
```

```
['Urdu', 'Pronunciation', 'Hindustani', 'pronunciation', ':', '\\xcb',  
'\\x88', '\\xca', '\\x8a', 'rd', '\\xcc', '\\xaa', 'u']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation',
```

'Hindustani', 'pronunciation', ':', '\xcb', '\x88', '\xca', '\x8a', 'rd', '\xcc', '\xaa', 'u']]], 13, True)

['Urdu', 'Official', 'language', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Urdu', 'Language', 'codes', 'ISO', '639-2', 'urd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Urdu', 'Language', 'codes', 'ISO', '639-3', 'urd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Urdu', 'Language', 'codes', 'ISO', '639-1', 'ur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Vietnamese', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vietnamese', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vietnamese', 'Spoken', 'in', 'Vietnamese', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Vietnamese', 'Writing', 'system', 'Vietnamese', 'variant', 'of', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Vietnamese', 'Official', 'language', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Vietnamese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-2', 'vie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-3', 'vie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vietnamese', 'Language', 'codes', 'Linguasphere', '46-EBA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Vietnamese', 'Language', 'codes', 'Linguasphere', 'Extent', 'of',  
'Vietnamese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Extent', 'of', 'Type: ANY']]], 7, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-1', 'vi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Volap\xc3\xbck', 'Total', 'speakers', '20']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Volap\xc3\xbck', 'Writing', 'system', 'German']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Volap\xc3\xbck', 'Writing', 'system', 'and', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Volap\xc3\xbck', 'Category', '(', 'purpose', ')', 'constructed',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Volap\xc3\xbck', 'Category', '(', 'purpose', ')', 'Volap', '\\xc3',  
'\\xbc', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Volap', '\\xc3', '\\xbc', 'k']]], 9, True)

['Volap\xc3\xbck', 'Created', 'by', 'Johann', 'Martin', 'Schleyer']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
'Johann', 'Martin', 'Schleyer']]], 6, True)

['Volap\xc3\xbck', 'Language', 'codes', 'ISO', '639-2', 'vol']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Volap\xc3\xbck', 'Language', 'codes', 'ISO', '639-3', 'vol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Volap\xc3\xbck', 'Language', 'codes', 'ISO', '639-1', 'vo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Welsh', 'Region', 'Spoken', 'throughout', 'Wales', 'and', 'some', 'nearby', 'parts', 'of', 'England', ',', 'the', 'Chubut', 'region', 'of', 'Patagonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Spoken', 'throughout', 'Wales', 'and', 'some', 'nearby', 'parts', 'of', 'England', ',', 'the', 'Chubut', 'region', 'of', 'Patagonia']], 17, True)

['Welsh', 'Spoken', 'in', 'Wales', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Welsh', 'Pronunciation', 'k', '\xc9', '\x99', 'm', '\xcb', '\x88', 'r', '\xc9', '\x91', '\xcb', '\x90', '\xc9', '\xa8', '\xc9', '\xa1']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k', '\xc9', '\x99', 'm', '\xcb', '\x88', 'r', '\xc9', '\x91', '\xcb', '\x90', '\xc9', '\xa8', '\xc9', '\xa1']], 17, True)

['Welsh', 'Official', 'language', 'in', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Welsh', 'Language', 'codes', 'ISO', '639-2', 'wel', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wel', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Welsh', 'Language', 'codes', 'ISO', '639-2', 'cym', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cym', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Welsh', 'Language', 'codes', 'ISO', '639-3', 'cym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Welsh', 'Language', 'codes', 'ISO', '639-1', 'cy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Yiddish', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Yiddish', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Yiddish', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Belarus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'France', 'and', 'elsewhere']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Yiddish', 'Writing', 'system', 'uses', 'a', 'Hebrew-based', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'uses', 'a', 'Hebrew-based', 'alphabet']]], 7, True)

['Yiddish', 'Pronunciation', '\xcb', '\x88', 'j', '\xc9', '\xaa', 'd', '\xc9', '\xaa', '\xca', '\x83', ',', '\xcb', '\x88', 'jid', '\xc9', '\xaa', '\xca', '\x83']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'j', '\xc9', '\xaa', 'd', '\xc9', '\xaa', '\xca', '\x83', ',', '\xcb', '\x88', 'jid', '\xc9', '\xaa', '\xca', '\x83']]], 20, True)

['Yiddish', 'Official', 'language', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Yiddish', 'Official', 'language', 'in', 'Recognized', 'as', 'a', 'minority', 'language', 'in', 'the', 'Russian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

```
'in', 'Recognized', 'as', 'a', 'minority', 'language', 'in', 'the', 'Type:
ANY']]], 12, True)
```

```
['Yiddish', 'Regulated', 'by', 'no', 'formal', 'bodies', ';']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'no',
'formal', 'bodies', ';']]], 7, True)
```

```
['Yiddish', 'Regulated', 'by', 'YIVO', 'de', 'facto']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'YIVO', 'de', 'facto']]], 6, True)
```

```
['Yiddish', 'Language', 'codes', 'ISO', '639-2', 'yid']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Yiddish', 'Language', 'codes', 'ISO', '639-3', 'yid', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'yid', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Yiddish', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Yiddish', 'Language', 'codes', 'ISO', '639-1', 'yi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Indonesian', 'Total', 'speakers', '23', 'million', 'native', ';',
'approx', '.', '165', 'million', 'total']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'native', ';', 'approx', '.', 'NUM', 'million',
'total']], ['NUM', '23'], ['NUM', '165']], 12, True)
```

```
['Indonesian', 'Region', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Indonesian', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Indonesian', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

['Indonesian', 'Official', 'language', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Indonesian', 'Regulated', 'by', 'Pusat', 'Bahasa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Indonesian', 'Language', 'codes', 'ISO', '639-2', 'ind']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Indonesian', 'Language', 'codes', 'ISO', '639-3', 'ind']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Indonesian', 'Language', 'codes', 'ISO', '639-1', 'id']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Malayalam', 'Total', 'speakers', ':', '\xe2', '\x80', '\xa2', '773', ',',  
'624', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', ':',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'UAE']]], ['NUM', '773'],  
['NUM', '624']], 12, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '134', ',',  
'728', 'in', 'Kuwait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'134'], ['NUM', '728']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '134', ',',  
'019', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'134'], ['NUM', '019']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '105', ',',  
'655', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'105'], ['NUM', '655']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '94', ',',



```

'310', 'in', 'Qatar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM',
'94'], ['NUM', '310']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '58', ',', '146', 'in', 'Bahrain']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM',
'58'], ['NUM', '146']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '26', ',', '237', 'in', 'UK']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM',
'26'], ['NUM', '237']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '15', ',', '600', 'in', 'other', 'Europe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'other', 'Europe']], ['NUM',
'15'], ['NUM', '600']], 12, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '11', ',', '346', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM',
'11'], ['NUM', '346']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '10', ',', '636', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM',
'10'], ['NUM', '636']], 11, True)

['Malayalam', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Malayalam', 'Region', 'Kerala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Malayalam', 'Region', 'Lakshadweep']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

```

['Malayalam', 'Region', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Malayalam', 'Region', 'Tamil', 'Nadu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malayalam', 'Region', 'Mah', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Malayalam', 'Region', 'Andaman', 'and', 'Nicobar', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Andaman', 'and', 'Nicobar', 'Islands']], 6, True)

['Malayalam', 'Official', 'language', 'in', 'India', '(', 'Kerala', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\(', 'Type: ANY', '\\)']], 8, True)

['Malayalam', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Malayalam', 'Language', 'codes', 'ISO', '639-2', 'mal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Malayalam', 'Language', 'codes', 'ISO', '639-3', 'mal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Malayalam', 'Language', 'codes', 'ISO', '639-1', 'ml']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Language', 'Spoken', 'in', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Language', 'Spoken', 'in', 'Kerala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Language', 'Spoken', 'in', 'Maharashtra']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Language', 'Spoken', 'in', 'Andhra', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Language', 'Spoken', 'in', 'Goa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Language', 'Spoken', 'in', 'Tamil', 'Nadu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Language', 'Spoken', 'in', 'USA', 'and', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Language', 'Official', 'language', 'in', 'India', '(', 'Karnataka', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Language', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Karnataka']]], 10, True)

['Language', 'Language', 'codes', 'ISO', '639-2', 'kan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Language', 'Language', 'codes', 'ISO', '639-3', 'kan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Language', 'Language', 'codes', 'ISO', '639-1', 'kn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Telugu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Telugu', 'Writing', 'system', 'Telugu', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Telugu', 'Official', 'language', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Telugu', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Telugu', 'Language', 'codes', 'ISO', '639-2', 'tel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Telugu', 'Language', 'codes', 'ISO', '639-3', 'tel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Telugu', 'Language', 'codes', 'ISO', '639-1', 'te']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dari', 'Dialects', 'Kaboli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Mazari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Herati']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Badakhshi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Panjshiri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Laghmani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Sistani']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Aimaqi']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Hazaragi']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Region', 'Central', 'Asia']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Region', 'Middle', 'East']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Region', 'South', 'Asia']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Spoken', 'in', 'Afghanistan']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Official', 'language', 'in', 'Afghanistan']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Dari', 'Regulated', 'by', 'Academy', 'of', 'Sciences', 'of', 'Afghanistan']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Academy', 'of', 'Sciences', 'of', 'Afghanistan']]], 8, True)

['Dari', 'Language', 'codes', 'ISO', '639-2', 'per', '(', 'B', ')']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'per', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Dari', 'Language', 'codes', 'ISO', '639-2', 'fas', '(', 'T', ')']:  
 ([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fas', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

```
['Dari', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dari', 'Language', 'codes', 'ISO', '639-3', 'aiq', '\xe2', '\x80',
'\x93', 'Aimaq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'aiq', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)
```

```
['Dari', 'Language', 'codes', 'ISO', '639-3', 'haz', '\xe2', '\x80',
'\x93', 'Hazaragi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'haz', '\xe2', '\x80', '\x93', 'Hazaragi']], ['NUM',
'639-3']], 10, True)
```

```
['Dari', 'Language', 'codes', 'ISO', '639-1', 'fa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Kurdish', 'Total', 'speakers', '16', 'million', 'to', '35', 'million']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'to', 'NUM', 'million']], ['NUM', '16'], ['NUM', '35']],
8, True)
```

```
['Kurdish', 'Region', 'West', 'Asia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kurdish', 'Spoken', 'in', 'Turkey']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kurdish', 'Spoken', 'in', 'Iran']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kurdish', 'Spoken', 'in', 'Iraq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kurdish', 'Spoken', 'in', 'Syria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kurdish', 'Spoken', 'in', 'Lebanon']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Georgia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Turkmenistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', '(', 'see', 'article', 'for', 'full', 'list', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\(', 'see', 'article', 'for', 'full', 'list', '\\)']]], 10, True)

['Kurdish', 'Official', 'language', 'in', 'Iran', ':', 'constitutional', 'status', 'as', 'a', 'regional', 'language', ',', 'but', 'never', 'implemented']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Iran', ':', 'constitutional', 'status', 'as', 'a', 'regional', 'language', ',', 'but', 'never', 'implemented']]], 16, True)

['Kurdish', 'Official', 'language', 'in', 'Armenia', ':', 'minority', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Armenia', ':', 'minority', 'language']]], 8, True)

['Kurdish', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-2', 'kur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'kur', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kur', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', ':', 'ckb', '\xe2', '\x80', '\x93', 'Sorani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ckb', '\xe2', '\x80', '\x93', 'Sorani']], ['NUM', '639-3']], 11, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'kmr', '\xe2', '\x80', '\x93', 'Kurmanji']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kmr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'lki', '\xe2', '\x80', '\x93', 'Laki']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lki', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-1', 'ku']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sudovian', 'Spoken', 'in', 'Yotvingia', 'and', 'S', '\xc5', '\xab', 'duva']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Yotvingia', 'and', 'S', '\xc5', '\xab', 'duva']], 9, True)

['Sudovian', 'Region', 'Europe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sudovian', 'Language', 'codes', 'ISO', '639-3', 'xsv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Old', 'Region', 'Eastern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Writing', 'system', 'Glagolitic', 'alphabet', ';', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'chu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'chu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Language', 'codes', 'Linguasphere', '53-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-1', 'cu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Latvian', 'Total', 'speakers', '110', ',', '000', '(', 'Abroad', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'Abroad', '\\\\)']], ['NUM', '110'], ['NUM', '000']], 9, True)

['Latvian', 'Region', 'Baltic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Latvian', 'Spoken', 'in', 'Latvia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Latvian', 'Writing', 'system', 'Latvian', 'variant', 'of', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Latvian', 'Official', 'language', 'in', 'Latvia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['Latvian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Latvian', 'Language', 'codes', 'ISO', '639-2', 'lav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Latvian', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Latvian', 'Language', 'codes', 'ISO', '639-3', ':', 'lav', '\xe2',  
'\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'lav', '\xe2', '\x80', '\x93', 'Macrolanguage']]],  
['NUM', '639-3']], 11, True)

['Latvian', 'Language', 'codes', 'Linguasphere', '54-AAB-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Latvian', 'Language', 'codes', 'ISO', '639-1', 'lv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Mon', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mon', 'Pronunciation', 'p', '\xca', '\xb0', 'esa', 'm', '\xc9', '\x91',  
'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p',  
'\xca', '\xb0', 'esa', 'm', '\xc9', '\x91', 'n']]], 10, True)

['Mon', 'Language', 'codes', 'ISO', '639-3', 'mnw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Michif', 'Total', 'speakers', '500', '\xe2', '\x80', '\x93', '1', ',',  
'000', 'endangered']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\xe2', '\x80', '\x93', 'NUM', ',', 'NUM', 'endangered']]], ['NUM',  
'500'], ['NUM', '1'], ['NUM', '000']], 11, True)

['Michif', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Michif', 'Region', 'M', '\xc3', '\xa9', 'tis', 'communities', 'in', 'the', 'Prairies', ';', 'mostly', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'M', '\xc3', '\xa9', 'tis', 'communities', 'in', 'the', 'Prairies', ';', 'mostly', 'Manitoba']], 13, True)

['Michif', 'Region', 'Saskatchewan', 'and', 'northwestern', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Saskatchewan', 'and', 'northwestern', 'Ontario']], 6, True)

['Michif', 'Official', 'language', 'in', 'M', '\xc3', '\xa9', 'tis', 'Nation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'M', '\xc3', '\xa9', 'tis', 'Nation']], 9, True)

['Michif', 'Regulated', 'by', 'M', '\xc3', '\xa9', 'tis', 'National', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'M', '\xc3', '\xa9', 'tis', 'National', 'Council']], 9, True)

['Michif', 'Language', 'codes', 'ISO', '639-3', 'crg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Belarusian', 'Spoken', 'in', 'Belarus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Belarusian', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Belarusian', 'Spoken', 'in', 'in', '14', 'other', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'in', 'NUM', 'other', 'countries']], ['NUM', '14']], 7, True)

['Belarusian', 'Total', 'speakers', '4', 'to', '9', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'to', 'NUM', 'million']], ['NUM', '4'], ['NUM', '9']], 7, True)

['Belarusian', 'Writing', 'system', 'Cyrillic', ',', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Belarusian', 'Official', 'language', 'in', 'Belarus']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Belarusian', 'Official', 'language', 'in', 'Poland', '(', 'in', 'Gmina', 'Orla', ',', 'Gmina', 'Narewka', ',', 'and', 'Hajn', '\\xc3', '\\xb3', 'wka', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Poland', '\\\\(', 'in', 'Gmina', 'Orla', ',', 'Gmina', 'Narewka', ',', 'and', 'Hajn', '\\xc3', '\\xb3', 'wka', '\\\\)']]], 19, True)

['Belarusian', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences', 'of', 'Belarus']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences', 'of', 'Belarus']]], 9, True)

['Belarusian', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Belarusian', 'Language', 'codes', 'ISO', '639-2', 'bel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Belarusian', 'Language', 'codes', 'ISO', '639-3', 'bel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Belarusian', 'Language', 'codes', 'ISO', '639-1', 'be']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ladin', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ladin', 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Ladin', 'Language', 'codes', 'ISO', '639-3', 'lld']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ladin', 'Language', 'codes', 'Linguasphere', '51-AAA-l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Ukrainian', 'Total', 'speakers', 'approximately', '42', 'up', 'to', '47',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'approximately', 'NUM', 'up', 'to', 'NUM', 'million']], ['NUM', '42'],  
['NUM', '47']], 9, True)

['Ukrainian', 'Spoken', 'in', 'See', 'article']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ukrainian', 'Writing', 'system', 'Cyrillic', '(', 'Ukrainian', 'variant',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Ukrainian', 'Pronunciation', 'ukr', '\xc9', '\x91', '\xcb', '\x88',  
'jin', '\xca', '\xb2', 's', '\xca', '\xb2', 'k', '\xc9', '\x91', '\xcb',  
'\x88', 'm', '\xc9', '\x94', 'w', '\xc9', '\x91']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ukr',  
'\xc9', '\x91', '\xcb', '\x88', 'jin', '\xca', '\xb2', 's', '\xca', '\xb2',  
'k', '\xc9', '\x91', '\xcb', '\x88', 'm', '\xc9', '\x94', 'w', '\xc9',  
'\x91']]], 24, True)

['Ukrainian', 'Official', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Ukrainian', 'Official', 'language', 'in', 'Transnistria', '(',  
'unrecognized', 'de', 'facto', 'state', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Transnistria', '\\(', 'unrecognized', 'de', 'facto', 'state', '\\  
)']]], 11, True)

['Ukrainian', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences',  
'of', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'National', 'Academy', 'of', 'Sciences', 'of', 'Ukraine']]], 9, True)

['Ukrainian', 'Recognised', 'minority', 'language', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Ukrainian', 'Recognised', 'minority', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Ukrainian', 'Recognised', 'minority', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Ukrainian', 'Recognised', 'minority', 'language', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Ukrainian', 'Recognised', 'minority', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Ukrainian', 'Recognised', 'minority', 'language', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Ukrainian', 'Language', 'codes', 'ISO', '639-2', 'ukr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ukrainian', 'Language', 'codes', 'ISO', '639-3', 'ukr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ukrainian', 'Language', 'codes', 'ISO', '639-1', 'uk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['International', 'Region', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['International', 'Signed', 'in', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', '\xe2', '\x80', '\x93']]], 6, True)

['International', 'Total', 'signers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\xe2', '\x80', '\x93']]], 6, True)

['British', 'Signed', 'in', 'United', 'Kingdom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'United', 'Kingdom']]], 5, True)

['British', 'Total', 'signers', 'About', '125', ',', '000', 'first-language', 'signers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'About', 'NUM', ',', 'NUM', 'first-language', 'signers']]], ['NUM', '125'], ['NUM', '000']], 9, True)

['Coptic', 'Spoken', 'in', 'Egypt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coptic', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coptic', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coptic', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Coptic', 'Total', 'speakers', '300']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coptic', 'Writing', 'system', 'Coptic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Coptic', 'Language', 'codes', 'ISO', '639-2', 'cop']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Coptic', 'Language', 'codes', 'ISO', '639-3', 'cop']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Thai', 'Total', 'speakers', 'over', '60', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'over', 'NUM', 'million']]], ['NUM', '60']], 6, True)

['Thai', 'Spoken', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Spoken', 'in', 'Northern', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Thai', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Spoken', 'in', 'Southern', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Thai', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Spoken', 'in', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thai', 'Writing', 'system', 'Thai', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Thai', 'Pronunciation', 'p', '\xca', '\xb0', '\xc4', '\x81', '\xcb', '\x90', 's', '\xc7', '\x8e', '\xcb', '\x90', 't', '\xca', '\xb0', '\xc4', '\x81', 'j']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\xca', '\xb0', '\xc4', '\x81', '\xcb', '\x90', 's', '\xc7', '\x8e', '\xcb', '\x90', 't', '\xca', '\xb0', '\xc4', '\x81', 'j']], 20, True)



['Thai', 'Official', 'language', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Thai', 'Regulated', 'by', 'The', 'Royal', 'Institute']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'The', 'Royal', 'Institute']]], 6, True)

['Thai', 'Language', 'codes', 'ISO', '639-2', 'tha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Thai', 'Language', 'codes', 'ISO', '639-3', 'tha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thai', 'Language', 'codes', 'Linguasphere', '47-AAA-b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Thai', 'Language', 'codes', 'ISO', '639-1', 'th']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Polari', 'Spoken', 'in', 'United', 'Kingdom', 'and', 'Republic', 'of', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'Kingdom', 'and', 'Republic', 'of', 'Ireland']]], 9, True)

['Polari', 'Total', 'speakers', 'no', 'estimate', 'available']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'no', 'estimate', 'available']]], 6, True)

['Polari', 'Language', 'codes', 'ISO', '639-3', 'pld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Duala', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Duala', 'Language', 'codes', 'ISO', '639-3', 'dua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Akkadian', 'Spoken', 'in', 'Assyria', 'and', 'Babylonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Akkadian', 'Region', 'Mesopotamia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Akkadian', 'Writing', 'system', 'Sumero-Akkadian', 'cuneiform']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Akkadian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Akkadian', 'Language', 'codes', 'ISO', '639-2', 'akk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Akkadian', 'Language', 'codes', 'ISO', '639-3', 'akk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Middle', 'Spoken', 'in', 'to', 'some', 'extent', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'to', 'some', 'extent', 'in', 'Ireland']]], 8, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'enm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'enm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['M\\xc4\\x81ori', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['M\\xc4\\x81ori', 'Region', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['M\\xc4\\x81ori', 'Official', 'language', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-2', 'mao', '(', 'B',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mao', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-2', 'mri', '(', 'T',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mri', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-3', 'mri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-1', 'mi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Vulgar', 'Spoken', 'in', 'Roman', 'Republic', ',', 'Roman', 'Empire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Roman',  
'Republic', ',', 'Roman', 'Empire']]], 8, True)

['Vulgar', 'Writing', 'system', 'Unwritten', 'Deduced', 'from', 'Romance',  
'languages', 'and', 'literary', 'references']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Unwritten', 'Deduced', 'from', 'Romance', 'languages', 'and', 'literary',  
'references']]], 11, True)

['Vulgar', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Vulgar', 'Language', 'codes', 'ISO', '639-3', 'lat', 'The', 'Roman',  
'Empire', 'in', '60', 'AD']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lat', 'The', 'Roman', 'Empire', 'in', 'NUM', 'AD']]], ['NUM',  
'639-3'], ['NUM', '60']], 12, True)

['Vulgar', 'Language', 'codes', 'ISO', '639-1', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Cree', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cree', 'Official', 'language', 'in', 'Northwest', 'Territories', '(', 'Canada', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Northwest', 'Territories', '\\(', 'Canada', '\\)']], 9, True)

['Cree', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Cree', 'Language', 'codes', 'ISO', '639-2', 'cre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'cre', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cre', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', ':', 'nsk', '\\xe2', '\\x80', '\\x93', 'Naskapi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'nsk', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'moe', '\\xe2', '\\x80', '\\x93', 'Montagnais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'moe', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'atj', '\\xe2', '\\x80', '\\x93', 'Atikamekw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'atj', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

```

['Cree', 'Language', 'codes', 'ISO', '639-1', 'cr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tetum', 'Spoken', 'in', 'East', 'Timor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Tetum', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tetum', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tetum', 'Spoken', 'in', 'Portugal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tetum', 'Region', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tetum', 'Official', 'language', 'in', 'East', 'Timor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Tetum', 'Regulated', 'by', 'National', 'Institute', 'of', 'Linguistics']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'National', 'Institute', 'of', 'Linguistics']]], 7, True)

['Tetum', 'Language', 'codes', 'ISO', '639-2', 'tet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tetum', 'Language', 'codes', 'ISO', '639-3', 'tet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Igbo', 'Total', 'speakers', '18', '\xe2', '\x80', '\x93', '25',
'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '18'], ['NUM',
'25']], 9, True)

```

['Igbo', 'Dialects', 'Aro']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Bonny']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', '\xe1', '\xba', '\xb8', 'kpeye']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Dialects', '\xe1', '\xba', '\xb8', 'kpeye' ] ]], 6, True)

['Igbo', 'Dialects', 'Enuan', '\xe1', '\xbb', '\x8b']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Dialects', 'Enuan', '\xe1', '\xbb', '\x8b' ] ]], 6, True)

['Igbo', 'Dialects', 'Ika']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Ikwerre']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Isuama']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Izzi']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Ngwana']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Ogbah']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Ohafia']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Igbo', 'Dialects', 'Ohuhu']:  
 ([[ 'Start', [ 'Phrase' ]], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

ANY'']], 3, True)

['Igbo', 'Dialects', '\xe1', '\xbb', '\x8c', 'n', '\xe1', '\xbb', '\x8b',  
'ch', '\xe1', '\xba', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xe1',  
'\xbb', '\x8c', 'n', '\xe1', '\xbb', '\x8b', 'ch', '\xe1', '\xba',  
'\xa1']]], 13, True)

['Igbo', 'Dialects', 'Opobo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Igbo', 'Dialects', '\xe1', '\xbb', '\x8c', 'l', '\xe1', '\xbb', '\xa5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xe1',  
'\xbb', '\x8c', 'l', '\xe1', '\xbb', '\xa5']]], 9, True)

['Igbo', 'Dialects', 'Owerre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Igbo', 'Dialects', '\xe1', '\xbb', '\xa4', 'kw', '\xe1', '\xbb', '\xa5',  
'\xc3', '\xa0', 'n', '\xc3', '\xac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xe1',  
'\xbb', '\xa4', 'kw', '\xe1', '\xbb', '\xa5', '\xc3', '\xa0', 'n', '\xc3',  
'\xac']]], 14, True)

['Igbo', 'Spoken', 'in', 'Nigeria', 'where', 'it', 'is', 'a', 'national',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Nigeria', 'where', 'it', 'is', 'a', 'national', 'language']]], 10, True)

['Igbo', 'Spoken', 'in', 'Equatorial', 'Guinea', 'where', 'it', 'is', 'a',  
'recognised', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Equatorial', 'Guinea', 'where', 'it', 'is', 'a', 'recognised',  
'language']]], 11, True)

['Igbo', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Igbo', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Igbo', 'Pronunciation', 'i', '\xc9', '\xa1', '\xcd', '\xa1', 'bo',

```

'\xcb', '\x90']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'i',
'\xc9', '\xa1', '\xcd', '\xa1', 'bo', '\xcb', '\x90']]], 10, True)

['Igbo', 'Official', 'language', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Igbo', 'Regulated', 'by', 'SPILC']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Igbo', 'Recognised', 'minority', 'language', 'in', 'Equatorial',
'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'Equatorial', 'Guinea']]], 7, True)

['Igbo', 'Language', 'codes', 'ISO', '639-2', 'ibo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Igbo', 'Language', 'codes', 'ISO', '639-1', 'ig']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Hakka', 'Total', 'speakers', '34', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '34']], 5, True)

['Hakka', 'Spoken', 'in', 'Mainland', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Hakka', 'Spoken', 'in', 'Thailand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Taiwan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Singapore']:

```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Mauritius']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Suriname']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'South', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hakka', 'Writing', 'system', 'hanzi', ',', 'romanization']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Hakka', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Hakka', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Hakka', 'Language', 'codes', 'ISO', '639-3', 'hak']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nyabwa', 'Spoken', 'in', 'Ivory', 'Coast']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nyabwa', 'Language', 'codes', 'ISO', '639-3', 'nwb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Quebec', 'Spoken', 'in', 'Quebec', '(', 'mainly', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Quebec', 'Spoken', 'in', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quebec', 'Spoken', 'in', 'Western', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Quebec', 'Spoken', 'in', 'New', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Quebec', 'Regulated', 'by', 'Office', 'qu', '\\xc3', '\\xa9', 'b', '\\xc3',  
'\\xa9', 'cois', 'de', 'la', 'langue', 'fran', '\\xc3', '\\xa7', 'aise']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Office', 'qu', '\\xc3', '\\xa9', 'b', '\\xc3', '\\xa9', 'cois', 'de', 'la',  
'langue', 'fran', '\\xc3', '\\xa7', 'aise']], 18, True)

['Quebec', 'Language', 'codes', 'ISO', '639-2', 'fre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Quebec', 'Language', 'codes', 'ISO', '639-3', 'fra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quebec', 'Language', 'codes', 'ISO', '639-1', 'fr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dida', 'Spoken', 'in', 'Ivory', 'Coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Dida', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dida', 'Language', 'codes', 'ISO', '639-3', ':', 'gud', '\\xe2', '\\x80',  
'\\x93', 'Yocobou', '\\xc3', '\\xa9', 'Dida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', ':', 'gud', '\xe2', '\x80', '\x93', 'Yocobou', '\xc3',  
'\xa9', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Taiwanese', 'Region', 'Taiwan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Taiwanese', 'Spoken', 'in', 'Taiwan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Taiwanese', 'Official', 'language', 'in', 'None', ',', 'de', 'facto',  
'status', 'in', 'Taiwan', 'since', 'it', 'is', 'one', 'of', 'the',  
'statutory', 'languages', 'for', 'public', 'transport', 'announcements',  
'in', 'Taiwan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'None', ',', 'de', 'facto', 'status', 'in', 'Taiwan', 'since', 'it',  
'is', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public',  
'transport', 'announcements', 'in', 'Taiwan']]], 25, True)

['Taiwanese', 'Regulated', 'by', 'National', 'Languages', 'Committee', '(',  
'Ministry', 'of', 'Education', ',', 'ROC', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'National', 'Languages', 'Committee', '\\(', 'Ministry', 'of', 'Education',  
,', 'ROC', '\\)']]], 13, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-3', 'nan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Syriac', 'Spoken', 'in', 'Mesopotamia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Syriac', 'Spoken', 'in', 'Assyria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Syriac', 'Spoken', 'in', 'Roman', 'Syria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Syriac', 'Writing', 'system', 'Syriac', 'abjad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Syriac', 'Pronunciation', '/', 'l', '\xc9', '\x9b', '\xca', '\x83',
'\xcb', '\x90', '\xc9', '\x91', 'n', '\xc9', '\x91', '\xca', '\x94',
'surj', '\xc9', '\x91', 'j', '\xc9', '\x91', '\xca', '\x94', '/']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',
'l', '\xc9', '\x9b', '\xca', '\x83', '\xcb', '\x90', '\xc9', '\x91', 'n',
'\xc9', '\x91', '\xca', '\x94', 'surj', '\xc9', '\x91', 'j', '\xc9',
'\x91', '\xca', '\x94', '/']]], 26, True)
```

```
['Syriac', 'Language', 'codes', 'ISO', '639-2', 'syc', '(', 'classical',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'syc', '\\(', 'classical', '\\)']]], ['NUM', '639-2']], 9,
True)
```

```
['Balochi', 'Spoken', 'in', 'Pakistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Turkmenistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'UAE']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Language', 'codes', 'ISO', '639-2', 'bal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bal', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bal', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)
```

```
['Balochi', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Zulu', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Zulu', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Lesotho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Region', 'KwaZulu-Natal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Zulu', 'Region', 'eastern', 'Gauteng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Region', 'eastern', 'Free', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Zulu', 'Region', 'southern', 'Mpumalanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Zulu', 'Language', 'codes', 'ISO', '639-2', 'zul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Zulu', 'Language', 'codes', 'ISO', '639-3', 'zul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zulu', 'Language', 'codes', 'ISO', '639-1', 'zu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Nepali', 'Total', 'speakers', 'native', '\xe2', '\x80', '\x93', '14',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'native', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '14']], 9,  
True)
```

```
['Nepali', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'Myanmar', '(', 'Burma', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)
```

```
['Nepali', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepali', 'Official', 'language', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Nepali', 'Official', 'language', 'in', 'Sikkim', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'India', '\\)']]], 8, True)
```

```
['Nepali', 'Language', 'codes', 'ISO', '639-2', 'nep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Nepali', 'Language', 'codes', 'ISO', '639-3', 'nep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nepali', 'Language', 'codes', 'ISO', '639-1', 'ne']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Scots', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Scots', 'Spoken', 'in', 'Northern', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Scots', 'Spoken', 'in', 'Republic', 'of', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Ireland']]], 6, True)
```

```
['Scots', 'Region', 'Northern', 'Isles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Scots', 'Region', 'Caithness']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Scots', 'Region', 'Ulster', ':', 'Counties', 'Down']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ulster', ':', 'Counties', 'Down']]], 6, True)

['Scots', 'Region', 'Antrim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Scots', 'Region', 'Londonderry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Scots', 'Region', 'Donegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Scots', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Scots', 'Regulated', 'by', '\xe2', '\x80', '\x94', 'Ireland', ':', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', '\xe2', '\x80', '\x94', 'Ireland', ':', 'None']]], 9, True)

['Scots', 'Regulated', 'by', 'although', 'the', 'cross-border', 'Ulster-Scots', 'Agency']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'although', 'the', 'cross-border', 'Ulster-Scots', 'Agency']]], 8, True)

['Scots', 'Regulated', 'by', 'established', 'by', 'the', 'Implementation', 'Agreement', 'following', 'the', 'Good', 'Friday', 'Agreement', 'promotes', 'usage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'established', 'by', 'the', 'Implementation', 'Agreement', 'following', 'the', 'Good', 'Friday', 'Agreement', 'promotes', 'usage']]], 15, True)

['Scots', 'Language', 'codes', 'ISO', '639-2', 'sco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)



```

['Scots', 'Language', 'codes', 'ISO', '639-3', 'sco']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sumerian', 'Spoken', 'in', 'Sumer', 'and', 'Akkad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']]], 6, True)

['Sumerian', 'Region', 'Iraq', '(', 'Mesopotamia', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Iraq', '\\(',
'Mesopotamia', '\\)']]], 6, True)

['Sumerian', 'Language', 'codes', 'ISO', '639-2', 'sux']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sumerian', 'Language', 'codes', 'ISO', '639-3', 'sux']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hindi-Urdu', 'Dialects', 'Dakhni']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Hindi-Urdu', 'Region', 'South', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Hindi-Urdu', 'Writing', 'system', 'Devanagari', 'script', ',', '']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Hindi-Urdu', 'Writing', 'system', 'Perso-Arabic', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-2', 'hin', ',', 'ur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'hi', ',', 'ur']], ['NUM', '639-2']], 8, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-3', ':', 'hin', '\\xe2',

```

'\x80', '\x93', 'Hindi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'hin', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-3', 'urd', '\xe2', '\x80', '\x93', 'Urdu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'urd', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-1', 'hi', ',', 'ur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hi', ',', 'ur']], ['NUM', '639-1']], 8, True)

["Ge'ez,", 'Ethiopic', 'Spoken', 'in', 'Ethiopia', ',', 'Eritrea', 'and', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Ethiopic', 'Spoken', 'in', 'Ethiopia', ',', 'Eritrea', 'and', 'Israel']]], 9, True)

["Ge'ez,", 'Writing', 'system', 'Ge'ez', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

["Ge'ez,", 'Pronunciation', '\xc9', '\xa1', '\xc9', '\xa8', '\xca', '\x95', '\xc9', '\xa8', 'z']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9', '\xa1', '\xc9', '\xa8', '\xca', '\x95', '\xc9', '\xa8', 'z']]], 11, True)

["Ge'ez,", 'Official', 'language', 'in', 'Liturgical', 'language', 'of', 'the', 'Ethiopian', 'Orthodox', 'Tewahedo', 'Church']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Liturgical', 'language', 'of', 'the', 'Ethiopian', 'Orthodox', 'Tewahedo', 'Type: ANY']]], 12, True)

["Ge'ez,", 'Official', 'language', 'in', 'Eritrean', 'Orthodox', 'Tewahedo', 'Church']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Eritrean', 'Orthodox', 'Tewahedo', 'Type: ANY']]], 8, True)

["Ge'ez,", 'Official', 'language', 'in', 'Ethiopic', 'Catholic', 'Church']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Ethiopic', 'Catholic', 'Type: ANY']]], 7, True)

["Ge'ez,", 'Official', 'language', 'in', 'and', 'Beta', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'and', 'Beta', 'Israel']]], 7, True)

["Ge'ez,", 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

["Ge'ez,", 'Language', 'codes', 'ISO', '639-2', 'gez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

["Ge'ez,", 'Language', 'codes', 'ISO', '639-3', 'gez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ad\xc3\xbbnaic', 'Category', '(', 'sources', ')', 'a', 'priori', 'language', ',', 'but', 'related', 'to', 'other', 'languages', 'of', 'Arda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'a', 'priori', 'language', ',', 'but', 'related', 'to', 'other', 'languages', 'of', 'Arda']]], 16, True)

['Ad\xc3\xbbnaic', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Ad\xc3\xbbnaic', 'Setting', 'and', 'usage', 'Fantasy', 'world', 'of', 'Middle-earth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Fantasy', 'world', 'of', 'Middle-earth']]], 8, True)

['Ad\xc3\xbbnaic', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Ad\xc3\xbbnaic', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Ad\xc3\xbbnaic', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Ad\xc3\xbbnaic', 'Category', '(', 'purpose', ')', 'fictional',

'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Ad\\xc3\\xbbnaic', 'Category', '(', 'purpose', ')', 'Ad', '\\xc3', '\\xbb', 'naic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Ad', '\\xc3', '\\xbb', 'naic']]], 9, True)

['Ad\\xc3\\xbbnaic', 'Language', 'codes', 'ISO', '639-2', 'art']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ad\\xc3\\xbbnaic', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Ainu', 'Region', 'Hokkaid', '\\xc5', '\\x8d']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Ainu', 'Spoken', 'in', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ainu', 'Writing', 'system', 'Katakana', ',', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', ',', 'Type: ANY', 'alphabet']]], 7, True)

['Ainu', 'Pronunciation', '\\xcb', '\\x88', 'ainu', 'i', '\\xcb', '\\x88', 'tak']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb', '\\x88', 'ainu', 'i', '\\xcb', '\\x88', 'tak']]], 9, True)

['Ainu', 'Language', 'codes', 'ISO', '639-3', 'ain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kazakh', 'Region', 'Central', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Kazakh', 'Writing', 'system', 'Latin', 'alphabet']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kazakh', 'Writing', 'system', 'Arabic', 'alphabet']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kazakh', 'Pronunciation', 'q', '\xc9', '\x91', 'z', '\xc9', '\x91', 'q', '\xcb', '\x88', 'tl', '\xc9', '\x99']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'q', '\xc9', '\x91', 'z', '\xc9', '\x91', 'q', '\xcb', '\x88', 'tl', '\xc9', '\x99']]], 14, True)
```

```
['Kazakh', 'Official', 'language', 'in', 'Kazakhstan']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)
```

```
['Kazakh', 'Regulated', 'by', 'No', 'official', 'regulation']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Kazakh', 'Language', 'codes', 'ISO', '639-2', 'kaz']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Kazakh', 'Language', 'codes', 'ISO', '639-3', 'kaz']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kazakh', 'Language', 'codes', 'ISO', '639-3', 'Area', 'with', 'significant', 'Kazakh-speaking', 'population']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Area', 'with', 'significant', 'Kazakh-speaking', 'population']], ['NUM', '639-3']], 10, True)
```

```
['Kazakh', 'Language', 'codes', 'ISO', '639-1', 'kk']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Uzbek', 'Spoken', 'in', 'Uzbekistan']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Uzbek', 'Spoken', 'in', 'Kyrgyzstan']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

['Uzbek', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uzbek', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uzbek', 'Spoken', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uzbek', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uzbek', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uzbek', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uzbek', 'Official', 'language', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Uzbek', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-2', 'uzb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-3', 'uzb', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'uzb', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

```

['Uzbek', 'Language', 'codes', 'ISO', '639-1', 'uz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tok', 'Total', 'speakers', '5-6', 'million', ';', 'approx', '.', '1',
'million', 'native', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', ';', 'approx', '.', 'NUM', 'million', 'native',
'speakers']], ['NUM', '5-6'], ['NUM', '1']], 12, True)

['Tok', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']], 6, True)

['Tok', 'Language', 'codes', 'ISO', '639-2', 'tpi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tok', 'Language', 'codes', 'ISO', '639-3', 'tpi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tok', 'Language', 'codes', 'Linguasphere', '52-ABB-cc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

['Serbian', 'Pronunciation', 'sr', '\xcc', '\xa9', '\xcc', '\x82', 'pski',
'\xcb', '\x90']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sr',
'\xcc', '\xa9', '\xcc', '\x82', 'pski', '\xcb', '\x90']], 10, True)

['Serbian', 'Official', 'language', 'in', 'Serbia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']], 5, True)

['Serbian', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Bosnia', 'and', 'Herzegovina']], 7, True)

['Serbian', 'Official', 'language', 'in', 'Kosovo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']], 5, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Croatia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',

```



'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Greece', '(',  
'in', 'Mount', 'Athos', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Greece', '\\\\(', 'in', 'Mount', 'Athos', '\\\\)']]], 11,  
True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Language', 'codes', 'ISO', '639-2', 'srp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Serbian', 'Language', 'codes', 'ISO', '639-3', 'srp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Serbian', 'Language', 'codes', 'ISO', '639-1', 'sr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['(Paraguayan)', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['(Paraguayan)', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['(Paraguayan)', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['(Paraguayan)', 'Pronunciation', '\\xca', '\\x94', '\\xc3', '\\xa3', '\\xca',  
'\\x8b', '\\xcc', '\\x83', '\\xc3', '\\xa3', '\\xc9', '\\xb2', '\\xe1', '\\xba',

```
'\xbd', '\xcb', '\x88', '\xca', '\x94', '\xe1', '\xba', '\xbd']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',  
'\x94', '\xc3', '\xa3', '\xca', '\x8b', '\xcc', '\x83', '\xc3', '\xa3',  
'\xc9', '\xb2', '\xe1', '\xba', '\xbd', '\xcb', '\x88', '\xca', '\x94',  
'\xe1', '\xba', '\xbd']]], 24, True)
```

```
['(Paraguayan)', 'Official', 'language', 'in', 'Paraguay']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['(Paraguayan)', 'Official', 'language', 'in', 'Corrientes', '(',  
'Argentina', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Corrientes', '\\(', 'Argentina', '\\)']]], 8, True)
```

```
['(Paraguayan)', 'Official', 'language', 'in', 'Brazil', '(' , 'regional',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Brazil', '\\(', 'regional', '\\)']]], 8, True)
```

```
['(Paraguayan)', 'Official', 'language', 'in', 'Mercosur']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['(Paraguayan)', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['(Paraguayan)', 'Language', 'codes', 'ISO', '639-2', 'grn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['(Paraguayan)', 'Language', 'codes', 'ISO', '639-3', 'gug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['(Paraguayan)', 'Language', 'codes', 'ISO', '639-1', 'gn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Bislama', 'Spoken', 'in', 'Vanuatu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bislama', 'Total', 'speakers', '6', ',', '200', '(' , 'first', 'language',  
)', ',', '']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'first', 'language', '\\)'], ',']], ['NUM', '6'],  
['NUM', '200']], 11, True)
```

```
['Bislama', 'Total', 'speakers', '200', ',', '000', '(', 'additional',  
'language', ')']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'additional', 'language', '\\)']], ['NUM',  
'200'], ['NUM', '000']], 10, True)
```

```
['Bislama', 'Official', 'language', 'in', 'Vanuatu']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Bislama', 'Regulated', 'by', 'No', 'official', 'regulation']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Bislama', 'Language', 'codes', 'ISO', '639-2', 'bis']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Bislama', 'Language', 'codes', 'ISO', '639-3', 'bis']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bislama', 'Language', 'codes', 'Linguasphere', '52-ABB-ce']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Bislama', 'Language', 'codes', 'ISO', '639-1', 'bi']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Lingala', 'Total', 'speakers', 'ca', '.', '2', 'million', 'native',  
'speakers', ',', 'between', '8', 'and', '30', 'million', 'second-language',  
'speakers']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'ca', '.', 'NUM', 'million', 'native', 'speakers', ',', 'between', 'NUM',  
'and', 'NUM', 'million', 'second-language', 'speakers']]], ['NUM', '2'],  
['NUM', '8'], ['NUM', '30']], 17, True)
```

```
['Lingala', 'Dialects', 'Bangala']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

['Lingala', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Lingala', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']]], 7, True)

['Lingala', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo', ',', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo', ',', 'Republic', 'of', 'the', 'Congo']]], 14, True)

['Lingala', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Lingala', 'Language', 'codes', 'ISO', '639-2', 'lin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Lingala', 'Language', 'codes', 'ISO', '639-3', 'lin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lingala', 'Language', 'codes', 'Linguasphere', '99-AUI-f']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Lingala', 'Language', 'codes', 'ISO', '639-1', 'ln']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Lithuanian', 'Total', 'speakers', '500', ',', '000', '(', 'Abroad', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'Abroad', '\\\\)']]], ['NUM', '500'], ['NUM', '000']], 9, True)

['Lithuanian', 'Total', 'speakers', '4', 'million', '(', 'Worldwide', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', '\\\\(', 'Worldwide', '\\\\)']]], ['NUM', '4']], 8, True)

['Lithuanian', 'Dialects', 'Samogitian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY']]], 3, True)

['Lithuanian', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lithuanian', 'Spoken', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lithuanian', 'Writing', 'system', 'Latin', '(', 'Lithuanian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Lithuanian', 'Official', 'language', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Lithuanian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Lithuanian', 'Official', 'language', 'in', 'Gmina', 'Pu', '\xc5', '\x84', 'sk', 'in', 'the', 'Podlaskie', 'Voivodeship', 'of', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Gmina', 'Pu', '\xc5', '\x84', 'sk', 'in', 'the', 'Podlaskie', 'Voivodeship', 'of', 'Poland']]], 15, True)

['Lithuanian', 'Language', 'codes', 'ISO', '639-2', 'lit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lithuanian', 'Language', 'codes', 'ISO', '639-3', 'lit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lithuanian', 'Language', 'codes', 'Linguasphere', '54-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Lithuanian', 'Language', 'codes', 'ISO', '639-1', 'lt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Astur-Leonese', 'Spoken', 'in', 'Spain', '(', 'in', 'the', 'autonomous',

'communities', 'of', 'Asturias', ',', 'northwestern', 'Extremadura', ',', 'northwestern', 'Castile', 'and', 'Le', '\xc3', '\xb3', 'n', 'and', 'western', 'Cantabria', ')', 'and', 'small', 'pockets', 'in', 'northeastern', 'Portugal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Spain', '\\(', 'in', 'the', 'autonomous', 'communities', 'of', 'Asturias', ',', 'northwestern', 'Extremadura', ',', 'northwestern', 'Castile', 'and', 'Le', '\xc3', '\xb3', 'n', 'and', 'western', 'Cantabria', '\\)', 'and', 'small', 'pockets', 'in', 'northeastern', 'Portugal']]], 32, True)

['Astur-Leonese', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-2', 'ast', '\xe2', '\x80', '\x93', 'Asturian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ast', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-2', 'ext', '\xe2', '\x80', '\x93', 'Extremaduran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ext', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-2', 'mwl', '\xe2', '\x80', '\x93', 'Mirandese']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mwl', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-3', ':', 'ast', '\xe2', '\x80', '\x93', 'Asturian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ast', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-3', 'ext', '\xe2', '\x80', '\x93', 'Extremaduran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ext', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',

'639-3']], 10, True)

['Astur-Leonese', 'Language', 'codes', 'ISO', '639-3', 'mwl', '\xe2',  
\x80', '\x93', 'Mirandese']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mwl', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Alsatian', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Alsatian', 'Region', 'Alsace']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Alsatian', 'Language', 'codes', 'ISO', '639-3', 'gsw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kyrgyz', 'Spoken', 'in', 'Kyrgyzstan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Afghanistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Pakistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Tajikistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Xinjiang', '(', 'China', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Kyrgyz', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kyrgyz', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Kyrgyz', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'UK']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Japan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Official', 'language', 'in', 'Kyrgyzstan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Kyrgyz', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kyrgyz', 'Language', 'codes', 'ISO', '639-2', 'kir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kyrgyz', 'Language', 'codes', 'ISO', '639-3', 'kir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kyrgyz', 'Language', 'codes', 'ISO', '639-1', 'ky']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)



```
['Ganda', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ganda', 'Region', 'Mainly', 'Buganda', 'region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Ganda', 'Language', 'codes', 'ISO', '639-2', 'lug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Ganda', 'Language', 'codes', 'ISO', '639-3', 'lug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ganda', 'Language', 'codes', 'ISO', '639-1', 'lg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Sabine', 'Spoken', 'in', 'Sabinium']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sabine', 'Region', 'Central', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sabine', 'Writing', 'system', 'Not', 'written', 'except', 'as',  
'Latinized', 'words']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Not', 'written', 'except', 'as', 'Latinized', 'words']]], 9, True)
```

```
['Sabine', 'Language', 'codes', 'ISO', '639-3', 'sbv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sabine', 'Language', 'codes', 'ISO', '639-3', 'Sabinium', '.', 'The',  
'border', 'with', 'Latium', 'to', 'the', 'south', 'was', 'the', 'Aniene',  
'river', ';', 'however', ',', 'it', 'is', 'possible', 'that', 'Sabines',  
'extended', 'to', 'Lake', 'Regillus', 'slightly', 'to', 'the', 'south',  
'of', 'it', 'near', 'Gabii']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Sabinium', '.', 'The', 'border', 'with', 'Latium', 'to',  
'the', 'south', 'was', 'the', 'Aniene', 'river', ';', 'however', ',', 'it',  
'is', 'possible', 'that', 'Sabines', 'extended', 'to', 'Lake', 'Regillus',
```

'slightly', 'to', 'the', 'south', 'of', 'it', 'near', 'Gabii']], ['NUM', '639-3']], 38, True)

['Tigrinya', 'Region', 'Eritrea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tigrinya', 'Region', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tigrinya', 'Region', 'especially', 'in', 'Tigray']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tigrinya', 'Spoken', 'in', 'Eritrea', ',', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Tigrinya', 'Writing', 'system', 'Ge'ez', 'alphabet', 'abugida']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Tigrinya', 'Pronunciation', '/', 't', '\xc9', '\xa8', '\xc9', '\xa1', 'r', '\xc9', '\xa8', '\xc9', '\xb2', 'a', '/']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', 't', '\xc9', '\xa8', '\xc9', '\xa1', 'r', '\xc9', '\xa8', '\xc9', '\xb2', 'a', '/']]], 15, True)

['Tigrinya', 'Official', 'language', 'in', 'Eritrea', '(', 'working', 'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Eritrea', '\\(', 'working', 'language', '\\)']]], 9, True)

['Tigrinya', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tigrinya', 'Language', 'codes', 'ISO', '639-2', 'tir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Tigrinya', 'Language', 'codes', 'ISO', '639-3', 'tir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```

['Tigrinya', 'Language', 'codes', 'ISO', '639-1', 'ti']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Amharic', 'Total', 'speakers', '17', ',', '528', ',', '500', ';', '14',
',', '743', ',', '556', 'monolingual', '(', 'Ethnologue', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', ',', 'NUM', ';', 'NUM', ',', 'NUM', ',', 'NUM',
'monolingual', '\\(', 'Ethnologue', '\\)']], ['NUM', '17'], ['NUM', '528'],
['NUM', '500'], ['NUM', '14'], ['NUM', '743'], ['NUM', '556']], 18, True)

['Amharic', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Amharic', 'Spoken', 'in', 'Eritrea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Amharic', 'Writing', 'system', "Ge'ez", 'alphabet', 'abugida']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Amharic', 'Pronunciation', 'amar', '\\xc9', '\\xa8', '\\xc9', '\\xb2',
'\\xc9', '\\xb2', 'a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'amar',
'\\xc9', '\\xa8', '\\xc9', '\\xb2', '\\xc9', '\\xb2', 'a']]], 10, True)

['Amharic', 'Official', 'language', 'in', 'Ethiopia', 'and', 'the',
'following', 'specific', 'regions', ':', 'Addis', 'Ababa', 'City',
'Council']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Ethiopia', 'and', 'the', 'following', 'specific', 'regions', ':',
'Addis', 'Ababa', 'City', 'Council']]], 15, True)

['Amharic', 'Official', 'language', 'in', 'Amhara', 'Region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Amharic', 'Official', 'language', 'in', 'Benishangul-Gumuz', 'Region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Amharic', 'Official', 'language', 'in', 'Dire', 'Dawa', 'Administrative',
'council']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

```

'in', 'Dire', 'Dawa', 'Administrative', 'council']]], 8, True)

['Amharic', 'Official', 'language', 'in', 'Gambela', 'Region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)

['Amharic', 'Official', 'language', 'in', 'SNNPR']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Amharic', 'Regulated', 'by', 'no', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]]], 6, True)

['Amharic', 'Language', 'codes', 'ISO', '639-2', 'amh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Amharic', 'Language', 'codes', 'ISO', '639-3', 'amh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Amharic', 'Language', 'codes', 'Linguasphere', '12-ACB-a']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]]], 5, True)

['Amharic', 'Language', 'codes', 'ISO', '639-1', 'am']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Wolof', 'Spoken', 'in', 'Senegal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Wolof', 'Spoken', 'in', 'Gambia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Wolof', 'Spoken', 'in', 'Mauritania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Wolof', 'Region', 'West', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

```
['Wolof', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Wolof', 'Regulated', 'by', 'CLAD', '(', 'Centre', 'de', 'linguistique',  
'appliqu', '\xc3', '\xa9', 'e', 'de', 'Dakar', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'CLAD', '\\(', 'Centre', 'de', 'linguistique', 'appliqu', '\xc3', '\xa9',  
'e', 'de', 'Dakar', '\\)']]], 15, True)
```

```
['Wolof', 'Language', 'codes', 'ISO', '639-2', 'wol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Wolof', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wolof', 'Language', 'codes', 'ISO', '639-3', ':', 'wol', '\xe2', '\x80',  
'\x93', 'Wolof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'wol', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Wolof', 'Language', 'codes', 'ISO', '639-3', 'wof', '\xe2', '\x80',  
'\x93', 'Gambian', 'Wolof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'wof', '\xe2', '\x80', '\x93', 'Gambian', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Wolof', 'Language', 'codes', 'ISO', '639-1', 'wo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Xhosa', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Xhosa', 'Spoken', 'in', 'Lesotho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Xhosa', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

['Xhosa', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Xhosa', 'Language', 'codes', 'ISO', '639-2', 'xho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Xhosa', 'Language', 'codes', 'ISO', '639-3', 'xho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xhosa', 'Language', 'codes', 'ISO', '639-1', 'xh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Istro-Romanian', 'Total', 'speakers', '1000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Istro-Romanian', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Istro-Romanian', 'Region', 'Istria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Istro-Romanian', 'Language', 'codes', 'ISO', '639-3', 'ruo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Istro-Romanian', 'Language', 'codes', 'Linguasphere', '51-AAD-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Oscan', 'Spoken', 'in', 'Samnium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oscan', 'Spoken', 'in', 'Campania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oscan', 'Spoken', 'in', 'Lucania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Oscan', 'Spoken', 'in', 'Calabria', 'and', 'Abruzzo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Oscan', 'Region', 'south', 'and', 'south-central', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'south', 'and',  
'south-central', 'Italy']]], 6, True)

['Oscan', 'Writing', 'system', 'Old', 'Italic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Oscan', 'Language', 'codes', 'ISO', '639-3', 'osc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Umbrian', 'Spoken', 'in', 'Umbria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Umbrian', 'Region', 'central', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Umbrian', 'Writing', 'system', 'Old', 'Italic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Tuvaluan', 'Spoken', 'in', 'Tuvalu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tuvaluan', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tuvaluan', 'Spoken', 'in', 'Kiribati']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tuvaluan', 'Spoken', 'in', 'Nauru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```

['Tuvaluan', 'Spoken', 'in', 'New', 'Zealand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Tuvaluan', 'Total', 'speakers', '10', ',', '670', 'in', 'Tuvalu', '.',
'13', ',', '051', 'total']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'in', 'Tuvalu', '.', 'NUM', ',', 'NUM', 'total']],
['NUM', '10'], ['NUM', '670'], ['NUM', '13'], ['NUM', '051']], 13, True)

['Tuvaluan', 'Official', 'language', 'in', 'Tuvalu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Tuvaluan', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Tuvaluan', 'Language', 'codes', 'ISO', '639-2', 'tvl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tuvaluan', 'Language', 'codes', 'ISO', '639-3', 'tvl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ribagor\xc3\xa7an', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)

['Ribagor\xc3\xa7an', 'Region', 'Ribagorza']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Ribagor\xc3\xa7an', 'Region', 'Campo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Ribagor\xc3\xa7an', 'Region', 'Estadilla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Ribagor\xc3\xa7an', 'Region', 'Graus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

```



```
['Ribagor\xc3\xa7an', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)
```

```
['Ribagor\xc3\xa7an', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Ribagor\xc3\xa7an', 'Language', 'codes', 'ISO', '639-3', 'cat-rib']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Classical', 'Region', 'mare', 'nostrum', '(', 'Mediterranean', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'mare',  
'nostrum', '\\(', 'Mediterranean', '\\)']]]], 7, True)
```

```
['Classical', 'Writing', 'system', 'Latin', 'alphabet', 'Literature']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Classical', 'Writing', 'system', 'documents']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Classical', 'Writing', 'system', 'inscriptions']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Classical', 'Pronunciation', 'la', '\xcb', '\x88', 'ti', '\xcb', '\x90',  
'nita', '\xcb', '\x90', 's']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'la',  
'\xcb', '\x88', 'ti', '\xcb', '\x90', 'nita', '\xcb', '\x90', 's']]]], 12,  
True)
```

```
['Classical', 'Official', 'language', 'in', 'Roman', 'republic', ',',  
'Roman', 'empire']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Roman', 'republic', ',', 'Roman', 'empire']]]], 9, True)
```

```
['Classical', 'Regulated', 'by', 'Schools', 'of', 'grammar', 'and',  
'rhetoric']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Schools', 'of', 'grammar', 'and', 'rhetoric']]]], 8, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-2', 'lat']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-3', 'lat-cla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-1', 'la']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Pashto', 'Total', 'speakers', 'estimated', '50', 'to', '60', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'estimated', 'NUM', 'to', 'NUM', 'million']], ['NUM', '50'], ['NUM',
'60']], 8, True)
```

```
['Pashto', 'Region', 'South-Central', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pashto', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pashto', 'Spoken', 'in', 'Pakistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pashto', 'Spoken', 'in', 'Iran', '(', 'minor', ')', 'and', 'by', 'the',
'Pashtun', 'diaspora', 'around', 'the', 'world']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iran',
'\\(', 'minor', '\\)', 'and', 'by', 'the', 'Pashtun', 'diaspora', 'around',
'the', 'world']]], 15, True)
```

```
['Pashto', 'Writing', 'system', 'Pashto', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Pashto', 'Pronunciation', 'pa', '\xca', '\x82', '\xcb', '\x88', 'to']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pa',  
'\xca', '\x82', '\xcb', '\x88', 'to']]], 8, True)

['Pashto', 'Pronunciation', 'pa', '\xc3', '\xa7', '\xcb', '\x88', 'to']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pa',  
'\xc3', '\xa7', '\xcb', '\x88', 'to']]], 8, True)

['Pashto', 'Pronunciation', 'pax', '\xcb', '\x88', 'to']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pax',  
'\xcb', '\x88', 'to']]], 6, True)

['Pashto', 'Official', 'language', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Pashto', 'Official', 'language', 'in', 'Pakistan', '(', 'provincial',  
'language', 'in', 'Khyber', 'Pakhtunkhwa', '.', 'and', 'FATA', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Pakistan', '\\(', 'provincial', 'language', 'in', 'Khyber',  
'Pakhtunkhwa', '.', 'and', 'FATA', '\\)']]], 15, True)

['Pashto', 'Regulated', 'by', 'Academy', 'of', 'Sciences', 'of',  
'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Academy', 'of', 'Sciences', 'of', 'Afghanistan']]], 8, True)

['Pashto', 'Language', 'codes', 'ISO', '639-2', 'pus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pashto', 'Language', 'codes', 'ISO', '639-3', 'pus', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'pus', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Pashto', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Pashto', 'Language', 'codes', 'Linguasphere', '58-ABD-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Pashto', 'Language', 'codes', 'ISO', '639-1', 'ps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Classical', 'Spoken', 'in', 'mainland', 'China', ';', 'Taiwan', ';', 'Japan', ';', 'Korea', 'and', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'mainland', 'China', ';', 'Taiwan', ';', 'Japan', ';', 'Korea', 'and', 'Vietnam']]], 13, True)

['Classical', 'Total', 'speakers', 'Not', 'a', 'spoken', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Not', 'a', 'spoken', 'language']]], 7, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'lzh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Language', 'codes', 'ISO', '639-1', 'zh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Kanuri', 'Spoken', 'in', 'Niger']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kanuri', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kanuri', 'Spoken', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kanuri', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kanuri', 'Region', 'West', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-2', 'kau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'kau', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kau', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'kby', '\xe2', '\x80', '\x93', 'Manga', 'Kanuri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kby', '\xe2', '\x80', '\x93', 'Manga', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'krt', '\xe2', '\x80', '\x93', 'Tumari', 'Kanuri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'krt', '\xe2', '\x80', '\x93', 'Tumari', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'bms', '\xe2', '\x80', '\x93', 'Bilma', 'Kanuri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bms', '\xe2', '\x80', '\x93', 'Bilma', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'kbl', '\xe2', '\x80', '\x93', 'Kanembu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kbl', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'KBL', 'Kanembu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'KBL', 'Type: ANY']], ['NUM', '639-3']], 7, True)

['Kanuri', 'Language', 'codes', 'ISO', '639-1', 'kr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Malay', 'Spoken', 'in', 'Brunei']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malay', 'Spoken', 'in', 'Singapore']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malay', 'Spoken', 'in', 'Philippines', '(', 'mostly', 'in', 'Tawi', 'Tawi', 'and', 'Sulu', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Philippines', '\\\\(', 'mostly', 'in', 'Tawi', 'Tawi', 'and', 'Sulu', '\\\\)']], 12, True)

['Malay', 'Spoken', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malay', 'Spoken', 'in', 'Myanmar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malay', 'Spoken', 'in', 'Sri', 'Lanka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Malay', 'Spoken', 'in', 'Cocos', '(', 'Keeling', ')', 'Islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Cocos', '\\\\(', 'Keeling', '\\\\)', 'Islands']], 8, True)

['Malay', 'Spoken', 'in', 'Christmas', 'Island']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Malay', 'Dialects', 'Malay', 'dialects']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malay', 'Standard', 'forms', 'Malaysian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malay', 'Standard', 'forms', 'Indonesian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Malay', 'Official', 'language', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Malay', 'Official', 'language', 'in', 'Brunei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Malay', 'Official', 'language', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Malay', 'Language', 'codes', 'ISO', '639-2', 'may', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'may', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Malay', 'Language', 'codes', 'ISO', '639-2', 'msa', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'msa', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'msa', '\\xe2', '\\x80',  
'\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'msa', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']]], ['NUM',  
'639-3']], 10, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', ':', 'zsm', '\\xe2', '\\x80',  
'\\x93', 'Malaysian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'zsm', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 11, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'ind', '\\xe2', '\\x80',  
'\\x93', 'Indonesian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ind', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'coa', '\\xe2', '\\x80',  
'\\x93', 'Cocos', 'Islands', 'Malay']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'coa', '\xe2', '\x80', '\x93', 'Cocos', 'Islands', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Malay', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Valencian', 'Region', 'of', 'Murcia', '(', 'Carche', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'of', 'Murcia', '\\(', 'Carche', '\\)']], 7, True)

['Valencian', 'Region', 'See', 'also', 'geographic', 'distribution', 'of', 'Catalan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'See', 'also', 'geographic', 'distribution', 'of', 'Type: ANY']], 8, True)

['Valencian', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Valencian', 'Spoken', 'in', 'Valencian', 'Community']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Valencian', 'Spoken', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']], 3, True)

['Valencian', 'Writing', 'system', 'Latin', '(', 'Catalan', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Valencian', 'Pronunciation', 'valensi', '\xcb', '\x88', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'valensi', '\xcb', '\x88', 'a']], 6, True)

['Valencian', 'Official', 'language', 'in', 'In', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Valencian', 'Regulated', 'by', 'Acad', '\xc3', '\xa8', 'mia', 'Valenciana', 'de', 'la', 'Llengua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Acad', '\xc3', '\xa8', 'mia', 'Valenciana', 'de', 'la', 'Llengua']], 11,



True)

```
['Valencian', 'Language', 'codes', 'ISO', '639-2', 'cat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Valencian', 'Language', 'codes', 'ISO', '639-3', 'cat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Valencian', 'Language', 'codes', 'Linguasphere', '51-AAA-e']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Valencian', 'Language', 'codes', 'ISO', '639-1', 'ca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Tutchone', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tutchone', 'Region', 'Yukon', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tutchone', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tutchone', 'Language', 'codes', 'ISO', '639-3', ':', 'tce', '\xe2',  
'\x80', '\x93', 'Southern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'tce', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Tutchone', 'Language', 'codes', 'ISO', '639-3', 'ttm', '\xe2', '\x80',  
'\x93', 'Northern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ttm', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Lao', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Lao', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lao', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lao', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lao', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lao', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lao', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lao', 'Spoken', 'in', 'Argentina', '(', 'Misiones', 'Province', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Argentina', '\\(', 'Misiones', 'Province', '\\)']]], 8, True)

['Lao', 'Writing', 'system', 'Lao', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lao', 'Pronunciation', 'p', '\\ca', '\\b0', '\\c3', '\\a1', '\\cb', '\\x90', 's', '\\c7', '\\x8e', '\\xcb', '\\x90', 'l', '\\c3', '\\a1', '\\cb', '\\x90', 'w']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\\ca', '\\b0', '\\c3', '\\a1', '\\cb', '\\x90', 's', '\\c7', '\\x8e', '\\xcb', '\\x90', 'l', '\\c3', '\\a1', '\\cb', '\\x90', 'w']]], 20, True)

['Lao', 'Official', 'language', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Lao', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Lao', 'Language', 'codes', 'ISO', '639-2', 'lao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lao', 'Language', 'codes', 'ISO', '639-3', 'lao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lao', 'Language', 'codes', 'ISO', '639-1', 'lo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Gujarati', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Gujarati', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gujarati', 'Spoken', 'in', 'New', 'Zealand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gujarati', 'Spoken', 'in', 'Mauritius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Panama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Writing', 'system', 'Gujarati', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gujarati', 'Writing', 'system', 'mainly', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Gujarati', 'Pronunciation', '/', '\xc9', '\xa1', 'ud', '\xca', '\x92', '(', '\xc9', '\x99', ')', '\xcb', '\x88', '\xc9', '\xbe', 'at', '\xcc', '\xaa', 'i', '/']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'\xc9', '\xa1', 'ud', '\xca', '\x92', '\\(', '\xc9', '\x99', '\\)', '\xcb',  
'\x88', '\xc9', '\xbe', 'at', '\xcc', '\xaa', 'i', '/']]], 21, True)

['Gujarati', 'Official', 'language', 'in', 'Gujarat', '(', 'India', ')']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'India', '\\)']]], 8, True)

['Gujarati', 'Official', 'language', 'in', 'Dadra', 'and', 'Nagar',  
'Haveli', '(', 'India', ')']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Dadra', 'and', 'Nagar', 'Haveli', '\\(', 'India', '\\)']]], 11,  
True)

['Gujarati', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Gujarati', 'Language', 'codes', 'ISO', '639-2', 'guj']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gujarati', 'Language', 'codes', 'ISO', '639-3', 'guj']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gujarati', 'Language', 'codes', 'ISO', '639-1', 'gu']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Southern', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Southern', 'Region', 'Mpumalanga']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Southern', 'Region', 'Limpopo']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Southern', 'Region', 'Gauteng']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Southern', 'Region', 'North', 'West']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Southern', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Southern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'nbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'nbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-1', 'nr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dalmatian', 'Spoken', 'in', 'Croatia', ',', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Dalmatian', 'Region', 'Adriatic', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dalmatian', 'Language', 'codes', 'ISO', '639-3', 'dlm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dalmatian', 'Language', 'codes', 'Linguasphere', '51-AAA-t']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Vandalic', 'Language', 'codes', 'ISO', '639-3', 'xvn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romani', 'Official', 'language', 'in', ':', 'Republic', 'of',  
'Macedonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Macedonia']]], 8, True)

['Romani', 'Official', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Slovenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Romani', 'Language', 'codes', 'ISO', '639-2', 'rom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rom', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rom', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rml', '\xe2', '\x80', '\x93', 'Baltic', 'Romani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rml', '\xe2', '\x80', '\x93', 'Baltic', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmc', '\xe2', '\x80', '\x93', 'Carpathian', 'Romani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rmc', '\xe2', '\x80', '\x93', 'Carpathian', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmo', '\xe2', '\x80', '\x93', 'Sinte', 'Romani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rmo', '\xe2', '\x80', '\x93', 'Sinte', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Yue', 'Dialects', 'Yuehai', '(', 'Cantonese', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Yue', 'Dialects', 'Siyi', '(', 'Taishanese', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Yue', 'Dialects', 'Gao-Yang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Yue', 'Dialects', 'Yong-Xun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Yue', 'Dialects', 'Goulou']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yue', 'Dialects', 'Luo-Guang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yue', 'Dialects', 'Qin-Lian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yue', 'Dialects', 'Wu-Hua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yue', 'Dialects', 'Hainan', 'Yue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yue', 'Writing', 'system', 'Traditional', 'Chinese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Yue', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Yue', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Yue', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Yue', 'Language', 'codes', 'ISO', '639-3', 'yue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yue', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Siwi', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Siwi', 'Region', 'Siwa', 'Oasis', ',', 'Gara', 'Oasis']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Siwa',
'Oasis', ',', 'Gara', 'Oasis']]], 7, True)

['Siwi', 'Language', 'codes', 'ISO', '639-3', 'siz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zenaga', 'Spoken', 'in', 'Mauritania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Zenaga', 'Region', 'Mederdra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Zenaga', 'Language', 'codes', 'ISO', '639-2', 'zen']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zenaga', 'Language', 'codes', 'ISO', '639-3', 'zen']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tuareg', 'Spoken', 'in', 'Algeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tuareg', 'Spoken', 'in', 'Burkina', 'Faso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Tuareg', 'Spoken', 'in', 'Libya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tuareg', 'Spoken', 'in', 'Mali']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tuareg', 'Spoken', 'in', 'Niger']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tuareg', 'Region', 'Sahara']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-2', 'tmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'tmh', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tmh', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-3', ':', 'thv', '\xe2', '\x80', '\x93', 'Tahaggart', 'Tamahaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'thv', '\xe2', '\x80', '\x93', 'Tahaggart', 'Type: ANY']]], ['NUM', '639-3']], 12, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'taq', '\xe2', '\x80', '\x93', 'Tamasheq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'taq', '\xe2', '\x80', '\x93', 'Tamasheq']]], ['NUM', '639-3']], 10, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'ttq', '\xe2', '\x80', '\x93', 'Tawallammat', 'Tamajaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ttq', '\xe2', '\x80', '\x93', 'Tawallammat', 'Tamajaq']]], ['NUM', '639-3']], 11, True)

['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'thz', '\xe2', '\x80', '\x93', 'Tayart', 'Tamajeq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'thz', '\xe2', '\x80', '\x93', 'Tayart', 'Tamajeq']]], ['NUM', '639-3']], 11, True)

['Ancient', 'Spoken', 'in', 'eastern', 'Mediterranean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ancient', 'Writing', 'system', 'Greek', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ancient', 'Language', 'codes', 'ISO', '639-2', 'grc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Ancient', 'Language', 'codes', 'ISO', '639-3', 'grc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Black', 'Speech', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Speech', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']]], 17, True)

['Black', 'Setting', 'and', 'usage', 'Mordor', 'in', 'Middle-earth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Mordor', 'in', 'Middle-earth']]], 7, True)

['Black', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Black', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Black', 'Category', '(', 'purpose', ')', 'artistic', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Black', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Polabian', 'Spoken', 'in', 'Poland', ',', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Polabian', 'Language', 'codes', 'ISO', '639-3', 'pox']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Polabian', 'Language', 'codes', 'Linguasphere', '53-AAA-bc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Slovincian', 'Spoken', 'in', 'Poland', ',', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Slovincian', 'Region', 'Pomerania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Slovincian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Slovincian', 'Language', 'codes', 'ISO', '639-3', 'csb-slo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Slovincian', 'Language', 'codes', 'Linguasphere', '53-AAA-ca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Knaanic', 'Spoken', 'in', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Knaanic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Knaanic', 'Language', 'codes', 'ISO', '639-3', 'czk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rusyn', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Poland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Hungary']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Romania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Serbia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Croatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rusyn', 'Official', 'language', 'in', 'Vojvodina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Rusyn', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Rusyn', 'Language', 'codes', 'ISO', '639-3', 'rue']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romano-Serbian', 'Spoken', 'in', 'Serbia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romano-Serbian', 'Language', 'codes', 'ISO', '639-3', 'rsb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Manchu', 'Region', 'Heilongjiang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Manchu', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Manchu', 'Writing', 'system', 'Manchu', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Manchu', 'Language', 'codes', 'ISO', '639-2', 'mnc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Manchu', 'Language', 'codes', 'ISO', '639-3', 'mnc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cimbrian', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cimbrian', 'Language', 'codes', 'ISO', '639-3', 'cim']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Austro-Bavarian', 'Spoken', 'in', 'Austria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Austro-Bavarian', 'Spoken', 'in', 'Germany', ',', 'Bavaria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Austro-Bavarian', 'Total', 'speakers', '12', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']]], ['NUM', '12']], 5, True)

['Austro-Bavarian', 'Language', 'codes', 'ISO', '639-3', 'bar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Galindan', 'Region', 'Europe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Galindan', 'Language', 'codes', 'ISO', '639-3', 'xgl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Georgian', 'Spoken', 'in', 'Georgia']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Total', 'speakers', '6-7', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '6-7']], 5, True)

['Georgian', 'Writing', 'system', 'Georgian', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Georgian', 'Official', 'language', 'in', 'Georgia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Georgian', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)



['Georgian', 'Language', 'codes', 'ISO', '639-2', 'geo', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'geo', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Georgian', 'Language', 'codes', 'ISO', '639-2', 'kat', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kat', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Georgian', 'Language', 'codes', 'ISO', '639-3', 'kat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Georgian', 'Language', 'codes', 'ISO', '639-1', 'ka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Mongolian', 'Region', 'All', 'of', 'Mongolia', 'and', 'Inner',  
'Mongolia', ';', 'parts', 'of', 'Liaoning']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'All', 'of',  
'Mongolia', 'and', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Liaoning']]],  
12, True)

['Mongolian', 'Region', 'Jilin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mongolian', 'Region', 'and', 'Heilongjiang', 'provinces', 'in', 'China']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Heilongjiang', 'provinces', 'in', 'China']]], 7, True)

['Mongolian', 'Pronunciation', '/', 'm', '\\xc9', '\\x94', '\\xc5', '\\x8b',  
'\\xc9', '\\xa2', '\\xc9', '\\x94', '\\xcc', '\\x86', '\\xc9', '\\xae', 'xi',  
'\\xc9', '\\xae', '/']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'm', '\\xc9', '\\x94', '\\xc5', '\\x8b', '\\xc9', '\\xa2', '\\xc9', '\\x94',  
'\\xcc', '\\x86', '\\xc9', '\\xae', 'xi', '\\xc9', '\\xae', '/']]], 20, True)

['Mongolian', 'Official', 'language', 'in', 'Mongolia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-2', 'mon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'mon', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mon', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', ':', 'khk', '\xe2', '\x80', '\x93', 'Khalkha', 'Mongolian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'khk', '\xe2', '\x80', '\x93', 'Khalkha', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'mvf', '\xe2', '\x80', '\x93', 'Peripheral', 'Mongolian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mvf', '\xe2', '\x80', '\x93', 'Peripheral', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-1', 'mn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Manglish', 'Spoken', 'in', 'Malaysia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Manglish', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Manglish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Marsian', 'Spoken', 'in', 'Marruvium']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marsian', 'Region', 'Marsica', 'in', 'Abruzzo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

```

['Marsian', 'Writing', 'system', 'Inscriptions', 'in', 'votive',
'offerings']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Inscriptions', 'in', 'votive', 'offerings']]], 7, True)

['Marsian', 'Language', 'codes', 'ISO', '639-3', 'ims']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marsian', 'Language', 'codes', 'ISO', '639-3', 'Sabinium', '.', 'The',
'Marsi', 'were', 'to', 'the', 'south', 'of', 'Lake', 'Fucinus', 'shown',
'in', 'the', 'lower', 'right']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Sabinium', '.', 'The', 'Marsi', 'were', 'to', 'the',
'south', 'of', 'Lake', 'Fucinus', 'shown', 'in', 'the', 'lower', 'right']],
['NUM', '639-3']], 21, True)

['Rotokas', 'Region', 'Bougainville']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Rotokas', 'Language', 'codes', 'ISO', '639-3', 'roo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Luba', 'Region', 'Kasai-Occidental', 'and', 'Kasai-Oriental',
'provinces']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kasai-
Occidental', 'and', 'Kasai-Oriental', 'provinces']]], 6, True)

['Luba', 'Language', 'codes', 'ISO', '639-2', 'lua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Luba', 'Language', 'codes', 'ISO', '639-3', 'lua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Burushaski', 'Spoken', 'in', 'Gilgit', '\xe2', '\x80', '\x93',
'Baltistan', ',', 'Pakistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Gilgit',
'\xe2', '\x80', '\x93', 'Baltistan', ',', 'Pakistan']]], 10, True)

```

['Burushaski', 'Spoken', 'in', 'Jammu', 'and', 'Kashmir', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Jammu',  
'and', 'Kashmir', ',', 'India']]], 8, True)

['Burushaski', 'Language', 'codes', 'ISO', '639-3', 'bsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Acadian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Acadian', 'Spoken', 'in', ':', 'New', 'Brunswick']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',  
'Type: ANY', 'Type: ANY']]], 6, True)

['Acadian', 'Spoken', 'in', 'Nova', 'Scotia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Acadian', 'Spoken', 'in', 'Prince', 'Edward', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Prince',  
'Edward', 'Island']]], 6, True)

['Acadian', 'Spoken', 'in', 'Magdalen', 'Islands', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Magdalen', 'Islands', ',', '']]], 6, True)

['Acadian', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Acadian', 'Spoken', 'in', ':', 'Maine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',  
'Type: ANY']]], 5, True)

['Acadian', 'Language', 'codes', 'ISO', '639-3', 'fra-aca', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'fra-aca', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,  
True)

['Acadian', 'Language', 'codes', 'Linguasphere', '51-AAA-ho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Norwegian', 'Total', 'speakers', '5', 'million', 'Norwegians']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', 'Norwegians']], ['NUM', '5']], 6, True)

['Norwegian', 'Writing', 'system', 'Latin', '(', 'Norwegian', 'variant',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Norwegian', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n',  
'\xc9', '\x94', '\xca', '\x82', 'k']], 8, True)

['Norwegian', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Norwegian', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nor', '\xe2', '\x80',  
'\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nob', '\xe2', '\x80',  
'\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']],  
['NUM', '639-2']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nno', '\xe2', '\x80',  
'\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-2']],  
10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nor', '\xe2',  
'\x80', '\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'nor', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-3']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-3']], 10, True)

['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-cd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'no', '\xe2', '\x80', '\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-1']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-1']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-1']], 10, True)

['Dungan', 'Spoken', 'in', 'Kyrgyzstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dungan', 'Spoken', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dungan', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dungan', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dungan', 'Spoken', 'in', 'and', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dungan', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dungan', 'Pronunciation', 'xw', '\xc9', '\x9b', 'itsu', 'jyj', '\xc9', '\x91', 'nneed', 'tone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'xw', '\xc9', '\x9b', 'itsu', 'jyj', '\xc9', '\x91', 'nneed', 'tone']]], 11, True)

['Dungan', 'Language', 'codes', 'ISO', '639-3', 'dng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Norwegian', 'Total', 'speakers', '5', 'million', 'Norwegians']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'Norwegians']]], ['NUM', '5']], 6, True)

['Norwegian', 'Writing', 'system', 'Latin', '(', 'Norwegian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Norwegian', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']]], 8, True)

['Norwegian', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Norwegian', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-2']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'nor', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-3']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-3']], 10, True)

['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-be']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'no', '\xe2', '\x80',



'\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-1']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-1']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-1']], 10, True)

['Brandenburgisch', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Brandenburgisch', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Brandenburgisch', 'Region', 'Brandenburg', ',', 'Saxony-Anhalt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Brandenburgisch', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Brandenburgisch', 'Language', 'codes', 'ISO', '639-3', 'nds-mar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Haz\xc4\x81rag\xc4\xab', 'Spoken', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Haz\xc4\x81rag\xc4\xab', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Region', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'haz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Avestan', 'Spoken', 'in', 'Liturgical', 'language', 'of',  
'Zoroastrianism']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Liturgical', 'language', 'of', 'Zoroastrianism']]], 7, True)

['Avestan', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']]], 6, True)

['Avestan', 'Language', 'codes', 'ISO', '639-2', 'ave']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Avestan', 'Language', 'codes', 'ISO', '639-3', 'ave']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Avestan', 'Language', 'codes', 'Linguasphere', '58-ABA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Avestan', 'Language', 'codes', 'ISO', '639-1', 'ae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Navajo', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Navajo', 'Region', 'Arizona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Navajo', 'Region', 'New', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Navajo', 'Region', 'Utah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Navajo', 'Region', 'Colorado']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Navajo', 'Language', 'codes', 'ISO', '639-2', 'nav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Navajo', 'Language', 'codes', 'ISO', '639-3', 'nav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Navajo', 'Language', 'codes', 'ISO', '639-1', 'nv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Bosnian', 'Total', 'speakers', '3', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '3']], 5, True)

['Bosnian', 'Region', 'Southern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bosnian', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Bosnian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Bosnian', 'Spoken', 'in', 'Montenegro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bosnian', 'Spoken', 'in', 'Kosovo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bosnian', 'Spoken', 'in', 'Croatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bosnian', 'Spoken', 'in', 'Macedonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bosnian', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bosnian', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bosnian', 'Pronunciation', 'bosanski', '\xcb', '\x90']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'bosanski', '\xcb', '\x90']]], 5, True)

['Bosnian', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina']]], 7, True)

['Bosnian', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Albania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Montenegro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Language', 'codes', 'ISO', '639-2', 'bos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bosnian', 'Language', 'codes', 'ISO', '639-3', 'bos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bosnian', 'Language', 'codes', 'ISO', '639-1', 'bs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Eastphalian', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eastphalian', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eastphalian', 'Language', 'codes', 'ISO', '639-2', 'gem']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Eastphalian', 'Language', 'codes', 'ISO', '639-3', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Okinawan', 'Region', 'The', 'Okinawa', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Okinawan', 'Spoken', 'in', 'Okinawa', 'Prefecture', ',', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Okinawa', 'Prefecture', ',', 'Japan']]], 7, True)

['Okinawan', 'Writing', 'system', 'Okinawan', 'writing', 'system']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Okinawan', 'Writing', 'system', 'Japanese', 'writing', 'system']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Okinawan', 'Pronunciation', '\xca', '\x94', 'ut', '\xcd', '\xa1', '\xc9',  
'\x95', 'ina', '\xcb', '\x90', '\xc9', '\xa1', 'ut', '\xcd', '\xa1',  
'\xc9', '\x95', 'i']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',  
'\x94', 'ut', '\xcd', '\xa1', '\xc9', '\x95', 'ina', '\xcb', '\x90',  
'\xc9', '\xa1', 'ut', '\xcd', '\xa1', '\xc9', '\x95', 'i']]], 20, True)

['Okinawan', 'Language', 'codes', 'ISO', '639-3', 'ryu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Okinawan', 'Language', 'codes', 'Linguasphere', '45-CAC-ai']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Okinawan', 'Language', 'codes', 'Linguasphere', '45-CAC-aj']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Okinawan', 'Language', 'codes', 'Linguasphere', '45-CAC-ak']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Texas', 'Spoken', 'in', 'Texas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Texas', 'Region', 'Texas', 'Hill', 'Country']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Texas', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Hutterite', 'Spoken', 'in', 'Alberta']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hutterite', 'Spoken', 'in', 'Saskatchewan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hutterite', 'Spoken', 'in', 'and', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hutterite', 'Spoken', 'in', 'Canada', ';', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada', ';', 'Washington']]], 6, True)

['Hutterite', 'Spoken', 'in', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hutterite', 'Language', 'codes', 'ISO', '639-3', 'geh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pennsylvania', 'Spoken', 'in', 'USA', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Pennsylvania', 'Language', 'codes', 'ISO', '639-3', 'pdc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pennsylvania', 'Language', 'codes', 'Linguasphere', '52-ACB-he']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Plautdietsch', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Belize']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Plautdietsch', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Uruguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Plautdietsch', 'Language', 'codes', 'ISO', '639-3', 'pdt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mohawk', 'Total', 'speakers', '3', ',', '350', '(', 'Ethnologue', ')',  
';', '900', 'in', 'Canada', ';', '2', ',', '017', 'in', 'the', 'United',  
'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)', ';', 'NUM', 'in', 'Canada',



[';', 'NUM', ',', 'NUM', 'in', 'the', 'United', 'States']], ['NUM', '3'],  
['NUM', '350'], ['NUM', '900'], ['NUM', '2'], ['NUM', '017']], 21, True)

['Mohawk', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Mohawk', 'Pronunciation', 'kanj', '\xca', '\x8c', '\xcc', '\x83', '\xca',  
\x94', '\xcb', '\x88', '\xc9', '\xa1', '\xc9', '\x9b', 'ha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kanj',  
\xca', '\x8c', '\xcc', '\x83', '\xca', '\x94', '\xcb', '\x88', '\xc9',  
\xa1', '\xc9', '\x9b', 'ha']]], 16, True)

['Mohawk', 'Language', 'codes', 'ISO', '639-2', 'moh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Mohawk', 'Language', 'codes', 'ISO', '639-3', 'moh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Soran\x3\xae', 'Total', 'speakers', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
NUM', 'million']]], ['NUM', '5']], 5, True)

['Soran\x3\xae', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Soran\x3\xae', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Soran\x3\xae', 'Region', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Soran\x3\xae', 'Language', 'codes', 'ISO', '639-3', 'ckb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Soran\x3\xae', 'Language', 'codes', 'Linguasphere', '58-AAA-cae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
Linguasphere', 'Type: ANY']]], 5, True)

['Quechua', 'Total', 'speakers', '\xe2', '\x89', '\x88', '10', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x89', '\x88', 'NUM', 'million']], ['NUM', '10']], 8, True)

['Quechua', 'Region', 'Central', 'Andes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quechua', 'Spoken', 'in', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quechua', 'Spoken', 'in', 'Bolivia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quechua', 'Spoken', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quechua', 'Spoken', 'in', 'Ecuador']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quechua', 'Spoken', 'in', 'Chile', 'and', 'Argentina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']], 6, True)

['Quechua', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Quechua', 'Official', 'language', 'in', 'Peru', 'and', 'Bolivia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Peru', 'and', 'Bolivia']], 7, True)

['Quechua', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Quechua', 'Language', 'codes', 'ISO', '639-2', 'que']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Quechua', 'Language', 'codes', 'ISO', '639-3', 'que']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quechua', 'Language', 'codes', 'ISO', '639-1', 'qu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Chinook', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Chinook', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chinook', 'Language', 'codes', 'ISO', '639-2', 'chn', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chn', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Chinook', 'Language', 'codes', 'ISO', '639-2', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 8, True)

['Chinook', 'Language', 'codes', 'ISO', '639-3', 'chn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chechen', 'Region', 'Republic', 'of', 'Chechnya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chechen', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chechen', 'Writing', 'system', 'Cyrillic', 'alphabet', ',', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Chechen', 'Official', 'language', 'in', 'Chechnya', ',', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', ',', 'Type: ANY']]], 7, True)

['Chechen', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

```

['Chechen', 'Language', 'codes', 'ISO', '639-2', 'che']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chechen', 'Language', 'codes', 'ISO', '639-3', 'che']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chechen', 'Language', 'codes', 'ISO', '639-1', 'ce']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['East', 'Spoken', 'in', 'Germany']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['East', 'Region', 'East', 'Frisia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['East', 'Language', 'codes', 'ISO', '639-2', 'frs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['East', 'Language', 'codes', 'ISO', '639-3', 'frs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taiwanese', 'Signed', 'in', 'Taiwan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Taiwanese', 'Total', 'signers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Gullah', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Gullah', 'Language', 'codes', 'ISO', '639-2', 'cpe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gullah', 'Language', 'codes', 'ISO', '639-3', 'gul']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gullah', 'Language', 'codes', 'Linguasphere', '52-ABB-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Arvanitika', 'Region', 'Attica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Arvanitika', 'Region', 'Boeotia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Arvanitika', 'Region', 'South', 'Euboea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Arvanitika', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Arvanitika', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Arvanitika', 'Pronunciation', 'a', '\xc9', '\xbe', 'b', '\xc9', '\x99',  
'\xcb', '\x88', '\xc9', '\xbe', 'i', '\xca', '\x83', 't']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'a',  
'\xc9', '\xbe', 'b', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xbe', 'i',  
'\xca', '\x83', 't']]], 16, True)

['Arvanitika', 'Language', 'codes', 'ISO', '639-2', 'alb', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'alb', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Arvanitika', 'Language', 'codes', 'ISO', '639-2', 'sqi', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sqi', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Arvanitika', 'Language', 'codes', 'ISO', '639-3', 'aat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Modern', 'Total', 'speakers', 'Second', 'language', ':', '199', '\xe2',  
'\x80', '\x93', '600', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Second', 'language', ':', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '199'], ['NUM', '600']], 12, True)

['Modern', 'Spoken', 'in', 'Listed', 'in', 'the', 'article']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Listed', 'in', 'the', 'article']]], 7, True)

['Modern', 'Writing', 'system', 'Latin', '(', 'English', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Modern', 'Official', 'language', 'in', '53', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Modern', 'Official', 'language', 'in', 'United', 'Nations']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Modern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', 'eng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-3', 'eng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-1', 'en']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Assamese', 'Region', 'Assam', ',', 'Arunachal', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Assam', ',', 'Arunachal', 'Pradesh']]], 6, True)

['Assamese', 'Writing', 'system', 'Assamese', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Assamese', 'Official', 'language', 'in', 'India', '(', 'Assam', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'India', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Assamese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Assamese', 'Language', 'codes', 'ISO', '639-2', 'asm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Assamese', 'Language', 'codes', 'ISO', '639-3', 'asm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Assamese', 'Language', 'codes', 'Linguasphere', '59-AAF-w']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Assamese', 'Language', 'codes', 'ISO', '639-1', 'as']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Spanish', 'Total', 'speakers', '35', ',', '468', ',', '501', '(',  
'United', 'States', 'only', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'United', 'States', 'only', '\\)']]], ['NUM', '35'], ['NUM', '468'], ['NUM', '501']], 13, True)

['Spanish', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Spanish', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Spanish', 'Pronunciation', 'espa', '\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'espa', '\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol']]], 8, True)

['Spanish', 'Language', 'codes', 'ISO', '639-2', 'spa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Spanish', 'Language', 'codes', 'ISO', '639-3', 'spa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Spanish', 'Language', 'codes', 'ISO', '639-1', 'es']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Lydian', 'Spoken', 'in', 'Lydia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lydian', 'Language', 'codes', 'ISO', '639-3', 'xld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lycian', 'Spoken', 'in', 'Lycia', ',', 'Lycaonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Lycian', 'Region', 'Southwestern', 'Anatolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lycian', 'Language', 'codes', 'ISO', '639-3', 'xlc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tatar', 'Spoken', 'in', 'Russia', ',', 'other', 'former', 'Soviet',  
'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Russia',  
, ',', 'other', 'former', 'Soviet', 'Union']], 9, True)

['Tatar', 'Official', 'language', 'in', 'Tatarstan', '(', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Tatarstan', '\\\\(', 'Russia', '\\\\)']], 8, True)

['Tatar', 'Language', 'codes', 'ISO', '639-2', 'tat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tatar', 'Language', 'codes', 'ISO', '639-3', 'tat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tatar', 'Language', 'codes', 'ISO', '639-1', 'tt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)



['Azerbaijani', 'Total', 'speakers', '20', '\xe2', '\x80', '\x93', '30', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '20'], ['NUM', '30']], 9, True)

['Azerbaijani', 'Spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Georgia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Uzbekistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Turkmenistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Pronunciation', 'az', '\xc3', '\xa6', 'rbajd', '\xca', '\x92', 'an', 'dili']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'az',

'\xc3', '\xa6', 'rbajd', '\xca', '\x92', 'an', 'dili']]], 10, True)

['Azerbaijani', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-2', 'aze']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', 'aze', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'aze', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-1', 'az']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Alem\xc3\xa1n', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Alem\xc3\xa1n', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Alem\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'gct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ket', 'Total', 'speakers', '550']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ket', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ket', 'Region', 'Krasnoyarsk', 'Krai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ket', 'Language', 'codes', 'ISO', '639-3', 'ket']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shelta', 'Spoken', 'in', 'Ireland', ',', 'Irish', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'diaspora']], 7, True)

['Shelta', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shelta', 'Language', 'codes', 'ISO', '639-3', 'sth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Min', 'Total', 'speakers', '49', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', 'NUM', 'million']], ['NUM', '49']], 6, True)

['Southern', 'Dialects', 'Hokkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Southern', 'Dialects', 'Teochew']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Southern', 'Dialects', 'Hainanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Southern', 'Spoken', 'in', 'Republic', 'of', 'China', '(', 'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'China', '\\(', 'Taiwan', '\\)']], 9, True)

['Southern', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Southern', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Southern', 'Spoken', 'in', 'Singapore']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Spoken', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Regulated', 'by', 'None', '(', 'The', 'Republic', 'of', 'China', 'Ministry', 'of', 'Education', 'and', 'some', 'NGOs', 'are', 'influential', 'in', 'Taiwan', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'None', '\\\\(', 'The', 'Republic', 'of', 'China', 'Ministry', 'of', 'Education', 'and', 'some', 'NGOs', 'are', 'influential', 'in', 'Taiwan', '\\\\)']]], 20, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'nan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-1', 'zh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ubykh', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Ubykh', 'Region', 'Manyas', ',', 'Bal', '\xc4', '\xb1', 'kesir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Manyas', ',',  
'Bal', '\xc4', '\xb1', 'kesir']]], 8, True)
```

```
['Ubykh', 'Language', 'codes', 'ISO', '639-3', 'uby']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Westphalian', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Westphalian', 'Spoken', 'in', 'Germany', ',', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Westphalian', 'Region', 'Northwestern', ',', 'Westphalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Westphalian', 'Language', 'codes', 'ISO', '639-3', 'wep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chagatai', 'Language', 'extinction', '1990s']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Chagatai', 'Language', 'codes', 'ISO', '639-2', 'chg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Chagatai', 'Language', 'codes', 'ISO', '639-3', 'chg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Proto-Germanic', 'Language', 'extinction', 'Gothic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Proto-Germanic', 'Spoken', 'in', 'Northern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Proto-Germanic', 'Writing', 'system', 'Elder', 'Futhark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['Proto-Germanic', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Abkhaz', 'Dialects', 'Abzhywa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abkhaz', 'Dialects', 'Bzyb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abkhaz', 'Dialects', 'Sadz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abkhaz', 'Region', 'Caucasus']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abkhaz', 'Spoken', 'in', 'Abkhazia', '(', 'Recognized', 'as', 'independent', 'by', 'Russia', ',', 'Nicaragua', ',', 'Nauru', 'and', 'Venezuela', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Abkhazia', '\\(', 'Recognized', 'as', 'independent', 'by', 'Russia', ',', 'Nicaragua', ',', 'Nauru', 'and', 'Venezuela', '\\)']]], 17, True)

['Abkhaz', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Georgia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Jordan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Abkhaz', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Abkhaz', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Abkhaz', 'Spoken', 'in', 'Iran', 'and', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)
```

```
['Abkhaz', 'Official', 'language', 'in', 'Abkhazia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Abkhaz', 'Official', 'language', 'in', 'Georgia', '(', 'on', 'the',  
'territory', 'of', 'the', 'Autonomous', 'Republic', 'of', 'Abkhazia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Georgia', '\\\\(', 'on', 'the', 'territory', 'of', 'the',  
'Autonomous', 'Republic', 'of', 'Abkhazia', '\\\\)']]], 16, True)
```

```
['Abkhaz', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Abkhaz', 'Language', 'codes', 'ISO', '639-2', 'abk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Abkhaz', 'Language', 'codes', 'ISO', '639-3', 'abk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Abkhaz', 'Language', 'codes', 'ISO', '639-1', 'ab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Northern', 'Region', 'Northern', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Northern', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Northern', 'Spoken', 'in', 'Burma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Northern', 'Spoken', 'in', 'Laos']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Northern', 'Spoken', 'in', 'Cambodia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Northern', 'Pronunciation', 'kam', '\xcb', '\xa7', 'm', '\xc9', '\xaf',
'\xcb', '\x90', '\xc9', '\x99', '\xc5', '\x8b', '\xcb', '\xa7']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kam',
'\xcb', '\xa7', 'm', '\xc9', '\xaf', '\xcb', '\x90', '\xc9', '\x99',
'\xc5', '\x8b', '\xcb', '\xa7']]], 16, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-2', 'tai']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-3', 'nod']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dyirbal', 'Spoken', 'in', 'Northeast', 'Queensland']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Dyirbal', 'Language', 'codes', 'ISO', '639-3', 'dbl']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cheyenne', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Cheyenne', 'Region', 'Montana', 'and', 'Oklahoma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Cheyenne', 'Language', 'codes', 'ISO', '639-2', 'chy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```



```

['Cheyenne', 'Language', 'codes', 'ISO', '639-3', 'chy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ugaritic', 'Spoken', 'in', 'ancient', 'Ugarit']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Ugaritic', 'Language', 'codes', 'ISO', '639-2', 'uga']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ugaritic', 'Language', 'codes', 'ISO', '639-3', 'uga']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pirah\u3\xa3', 'Total', 'speakers', '250-380']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pirah\u3\xa3', 'Region', 'Amazon', 'River']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pirah\u3\xa3', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pirah\u3\xa3', 'Pronunciation', '\xca', '\x94', '\xc3', '\xa0', 'p',
'\xc3', '\xa0', 'i', '\xcc', '\xaf', '\xcb', '\x88', 't', '\xcd', '\xa1',
'\xca', '\x83', '\xc3', '\xae', '\xcb', '\x90', 's', '\xc3', '\xb2']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',
'\x94', '\xc3', '\xa0', 'p', '\xc3', '\xa0', 'i', '\xcc', '\xaf', '\xcb',
'\x88', 't', '\xcd', '\xa1', '\xca', '\x83', '\xc3', '\xae', '\xcb',
'\x90', 's', '\xc3', '\xb2']]], 26, True)

['Pirah\u3\xa3', 'Language', 'codes', 'ISO', '639-3', 'myp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tanacross', 'Region', 'Alaska', '(', 'middle', 'Tanana', 'River', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\
(', 'middle', 'Tanana', 'River', '\)']]], 8, True)

['Tanacross', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

```

ANY', 'Type: ANY']]], 5, True)

['Tanacross', 'Language', 'codes', 'ISO', '639-3', 'tcb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nux\xc3\xa1lk', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nux\xc3\xa1lk', 'Region', 'Bella', 'Coola', 'area']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nux\xc3\xa1lk', 'Region', 'Central', 'Coast', 'region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nux\xc3\xa1lk', 'Region', 'British', 'Columbia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nux\xc3\xa1lk', 'Language', 'codes', 'ISO', '639-3', 'blc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Westron', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Westron', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Westron', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Westron', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Westron', 'Category', '(', 'purpose', ')', 'languages', 'of', 'Arda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'languages', 'of', 'Arda']]], 8, True)

```

['Westron', 'Category', '(', 'purpose', ')', 'Taliskan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Westron', 'Category', '(', 'purpose', ')', 'Ad', '\\xc3', '\\xbb', 'naic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'Ad', '\\xc3', '\\xbb', 'naic']]], 9, True)

['Westron', 'Category', '(', 'purpose', ')', 'Westron']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Westron', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',
',', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Westron', 'Category', '(', 'sources', ')', 'a', 'priori', 'language',
',', 'but', 'relative', 'to', 'other', 'languages', 'of', 'Arda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'sources', '\\\\)', 'a', 'priori', 'language', ',', 'but', 'relative', 'to',
'other', 'languages', 'of', 'Arda']]], 16, True)

['Westron', 'Setting', 'and', 'usage', 'Fantasy', 'world', 'of', 'Middle-
earth']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',
'usage', 'Fantasy', 'world', 'of', 'Middle-earth']]], 8, True)

['Westron', 'Language', 'codes', 'ISO', '639-2', 'art']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Westron', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Filipino', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Filipino', 'Total', 'speakers', 'First', 'language', ':', '25',
'million', '.', 'Second', 'language', ':', 'over', '60', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'First', 'language', ':', 'NUM', 'million', '.', 'Second', 'language', ':',
'over', 'NUM', 'million']], ['NUM', '25'], ['NUM', '60']], 15, True)

['Filipino', 'Total', 'speakers', 'Overall', ':', '90', 'million']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Overall', ':', 'NUM', 'million']], ['NUM', '90']], 7, True)

['Filipino', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Filipino', 'Official', 'language', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Filipino', 'Language', 'codes', 'ISO', '639-2', 'fil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Filipino', 'Language', 'codes', 'ISO', '639-3', 'fil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Filipino', 'Language', 'codes', 'ISO', '639-1', 'tl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Votic', 'Total', 'speakers', '20', '(', 'or', 'fewer', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\(', 'or', 'fewer', '\\)']], ['NUM', '20']], 8, True)

['Votic', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Votic', 'Region', 'Ingria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Votic', 'Language', 'codes', 'ISO', '639-2', 'vot']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Votic', 'Language', 'codes', 'ISO', '639-3', 'vot']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Region', 'Medieval', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Middle', 'Writing', 'system', 'Seal', 'Script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Middle', 'Writing', 'system', 'Clerical', 'Script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Middle', 'Writing', 'system', 'Kaishu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Middle', 'Writing', 'system', 'Semi-cursive', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Middle', 'Writing', 'system', 'Grass', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Middle', 'Writing', 'system', 'Phagspa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Middle', 'Writing', 'system', 'Hangul']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Middle', 'Official', 'language', 'in', 'Sui', 'Dynasty']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Middle', 'Official', 'language', 'in', 'Tang', 'Dynasty']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Middle', 'Official', 'language', 'in', 'Song', 'Dynasty']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Middle', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'ltc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sona', 'language', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Sona', 'Date', 'founded', '1935']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sona', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sona', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sona', 'Category', '(', 'purpose', ')', 'Sona', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sona', 'Created', 'by', 'Kenneth', 'Searight']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Sona', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Sona', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Livonian', 'Spoken', 'in', 'Latvia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Livonian', 'Region', 'Livonian', 'Coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Livonian', 'Language', 'codes', 'ISO', '639-3', 'liv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Latvia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Spoken', 'in', 'Lithuania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'xcu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nivkh', 'Spoken', 'in', 'Russia', ',', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Nivkh', 'Region', 'Sakhalin', 'Island', ',', 'and', 'along', 'the', 'Amur', 'River']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sakhalin', 'Island', ',', 'and', 'along', 'the', 'Amur', 'River']], 10, True)

['Nivkh', 'Language', 'codes', 'ISO', '639-3', 'niv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ga', 'Spoken', 'in', 'South-eastern', 'Ghana', ',', 'around', 'Accra']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'South-eastern', 'Ghana', ',', 'around', 'Accra']], 8, True)

['Ga', 'Pronunciation', '\xc9', '\xa1', '\xc3', '\xa3']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9', '\xa1', '\xc3', '\xa3']], 6, True)

['Ga', 'Official', 'language', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Ga', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:

ANY', 'of', 'Type: ANY']]], 6, True)

['Ga', 'Language', 'codes', 'ISO', '639-2', 'gaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ga', 'Language', 'codes', 'ISO', '639-3', 'gaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tulu', 'Region', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tulu', 'Region', 'Gulf', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tulu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tulu', 'Writing', 'system', 'Tulu', 'script', '(', 'originally', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Tulu', 'Writing', 'system', 'Kannada', 'script', '(', 'presently', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Tulu', 'Language', 'codes', 'ISO', '639-3', 'tcy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chukchi', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chukchi', 'Region', 'Chukotka', 'Autonomous', 'Okrug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Chukchi', 'Language', 'codes', 'ISO', '639-3', 'ckt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Karelian', 'Region', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Karelian', 'Region', 'Tver', 'Oblast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karelian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karelian', 'Spoken', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karelian', 'Official', 'language', 'in', 'recognised', 'as', 'minority',  
'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'recognised', 'as', 'minority', 'language', 'in']]], 9, True)

['Karelian', 'Official', 'language', 'in', ':', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Karelian', 'Official', 'language', 'in', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Republic', 'of', 'Karelia']]], 7, True)

['Karelian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Karelian', 'Language', 'codes', 'ISO', '639-2', 'krl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Karelian', 'Language', 'codes', 'ISO', '639-3', 'krl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Veps', 'Region', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Veps', 'Region', 'Vologda', 'Oblast']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Veps', 'Region', 'Veps', 'National', 'Volost']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Veps', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Veps', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']]], 9, True)

['Veps', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Karelia']]], 8, True)

['Veps', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Veps', 'Language', 'codes', 'ISO', '639-3', 'vep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luwian', 'Spoken', 'in', 'Hittite', 'Empire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Luwian', 'Spoken', 'in', 'Arzawa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Luwian', 'Spoken', 'in', 'Neo-Hittite', 'kingdoms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Luwian', 'Category', '(', 'sources', ')', 'Inscriptions', 'and', 'names']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'sources', '\\\\)', 'Inscriptions', 'and', 'names']]], 8, True)

['Luwian', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luwian', 'Language', 'codes', 'ISO', '639-3', ':', 'xlu', '\xe2', '\x80', '\x93', 'Cuneiform', 'Luwian']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'xlu', '\xe2', '\x80', '\x93', 'Cuneiform', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Turkmen', 'Spoken', 'in', 'Turkmenistan']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Turkmen', 'Spoken', 'in', 'Iran']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Turkmen', 'Spoken', 'in', 'Afghanistan']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Turkmen', 'Spoken', 'in', 'Stavropol', 'krai', '(', 'Russia', ')']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Stavropol', 'krai', '\\(', 'Russia', '\\)']], 8, True)

['Turkmen', 'Total', 'speakers', 'ca', '.', '4', 'million']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c', '.', 'NUM', 'million']], ['NUM', '4']], 7, True)

['Turkmen', 'Official', 'language', 'in', 'Turkmenistan']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Turkmen', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Turkmen', 'Language', 'codes', 'ISO', '639-2', 'tuk']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Turkmen', 'Language', 'codes', 'ISO', '639-3', 'tuk']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turkmen', 'Language', 'codes', 'ISO', '639-1', 'tk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Proven\xc3\xa7al', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proven\xc3\xa7al', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proven\xc3\xa7al', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proven\xc3\xa7al', 'Spoken', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proven\xc3\xa7al', 'Spoken', 'in', 'small', 'community', 'in', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'small', 'community', 'in', 'California']], 7, True)

['Proven\xc3\xa7al', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Proven\xc3\xa7al', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Proven\xc3\xa7al', 'Language', 'codes', 'ISO', '639-3', 'oci-prv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kodava', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kodava', 'Region', 'Kodagu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

```

['Kodava', 'Language', 'codes', 'ISO', '639-3', 'kfa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mirandese', 'Spoken', 'in', 'Portugal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mirandese', 'Official', 'language', 'in', 'Co-official', 'recognition',
'.', 'Special', 'protection', 'status', 'in', 'Miranda', 'do', 'Douro',
',', 'Portugal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Co-official', 'recognition', '.', 'Special', 'protection', 'status',
'in', 'Miranda', 'do', 'Douro', ',', 'Portugal']]], 16, True)

['Mirandese', 'Regulated', 'by', 'Anstituto', 'de', 'la', 'Lh', '\xc3',
'\xa9', 'ngua', 'Mirandesa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Anstituto', 'de', 'la', 'Lh', '\xc3', '\xa9', 'ngua', 'Mirandesa']]], 11,
True)

['Mirandese', 'Language', 'codes', 'ISO', '639-2', 'mwl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mirandese', 'Language', 'codes', 'ISO', '639-3', 'mwl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mirandese', 'Language', 'codes', 'Linguasphere', '51-AAA-cb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Modern', 'Spoken', 'in', 'Greece']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Modern', 'Spoken', 'in', 'Cyprus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Modern', 'Spoken', 'in', 'Greek', 'diaspora']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Modern', 'Total', 'speakers', '12', 'million']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '12']], 5, True)

['Modern', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Modern', 'Official', 'language', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Modern', 'Official', 'language', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Modern', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Modern', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in', 'parts', 'of']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in', 'parts', 'of']], 11, True)

['Modern', 'Official', 'language', 'in', ':', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Modern', 'Official', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Modern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', 'gre', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gre', '\\\\(', 'B', '\\\)']], ['NUM', '639-2']], 9, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', 'ell', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ell', '\\\\(', 'T', '\\\)']], ['NUM', '639-2']], 9, True)

['Modern', 'Language', 'codes', 'ISO', '639-1', 'el']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Upper', 'Region', 'Saxony', ',', 'Brandenburg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Upper', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Upper', 'Writing', 'system', 'Latin', '(', 'Sorbian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Upper', 'Pronunciation', '\xcb', '\x88', 'h', '\xc9', '\x94', 'rnj', '\xc9', '\x94', '\xcb', '\x8c', 's', '\xc9', '\x9b', 'rb', '\xca', '\x83', 't', '\xca', '\x83', 'ina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'h', '\xc9', '\x94', 'rnj', '\xc9', '\x94', '\xcb', '\x8c', 's', '\xc9', '\x9b', 'rb', '\xca', '\x83', 't', '\xca', '\x83', 'ina']], 22, True)

['Upper', 'Official', 'language', 'in', 'regional', 'language', 'in', 'Germany', '(', 'Brandenburg', 'and', 'Saxony', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'regional', 'language', 'in', 'Germany', '\\(', 'Brandenburg', 'and', 'Saxony', '\\)']], 13, True)

['Upper', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Upper', 'Language', 'codes', 'ISO', '639-2', 'hsb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Upper', 'Language', 'codes', 'ISO', '639-3', 'hsb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lower', 'Region', 'Brandenburg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Lower', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lower', 'Writing', 'system', 'Latin', '(', 'Sorbian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Lower', 'Pronunciation', '\xcb', '\x88', 'd', '\xc9', '\x94', 'ln', '\xc9', '\x94', '\xcb', '\x8c', 's', '\xc9', '\x9b', 'rsk', '\xca', '\xb2', 'i']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'd', '\xc9', '\x94', 'ln', '\xc9', '\x94', '\xcb', '\x8c', 's', '\xc9', '\x9b', 'rsk', '\xca', '\xb2', 'i']], 19, True)

['Lower', 'Language', 'codes', 'ISO', '639-2', 'dsb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lower', 'Language', 'codes', 'ISO', '639-3', 'dsb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abaza', 'Spoken', 'in', 'Russia', ',', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Abaza', 'Region', 'Karachay-Cherkess', 'Republic', 'of', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Abaza', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Abaza', 'Language', 'codes', 'ISO', '639-3', 'abq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sundanese', 'Total', 'speakers', '33', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '33']], 5, True)

['Sundanese', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



```
['Sundanese', 'Region', 'West', 'Java']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Sundanese', 'Region', 'Banten']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Sundanese', 'Region', 'Jakarta']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Sundanese', 'Official', 'language', 'in', 'West', 'Java']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Sundanese', 'Regulated', 'by', 'No', 'official', 'regulation']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Sundanese', 'Language', 'codes', 'ISO', '639-2', 'sun']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Sundanese', 'Language', 'codes', 'ISO', '639-3', 'sun']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sundanese', 'Language', 'codes', 'ISO', '639-1', 'su']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Walloon', 'Spoken', 'in', 'Belgium']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Walloon', 'Spoken', 'in', 'France']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Walloon', 'Spoken', 'in', '(', 'also', ':', 'some', 'speakers', 'in',
'the', 'United', 'States', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\(',
'also', ':', 'some', 'speakers', 'in', 'the', 'United', 'States', '\\)']]],
13, True)
```

['Walloon', 'Language', 'codes', 'ISO', '639-2', 'wln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Walloon', 'Language', 'codes', 'ISO', '639-3', 'wln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Walloon', 'Language', 'codes', 'Linguasphere', '51-AAA-hf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Walloon', 'Language', 'codes', 'ISO', '639-1', 'wa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Orokolo', 'Region', 'Gulf', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Orokolo', 'Language', 'codes', 'ISO', '639-3', 'oro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tiwa', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tiwa', 'Region', 'Tiwa', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tiwa', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Proto-Esperanto', 'Date', 'founded', '1878-1881']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Esperanto', 'Language', 'extinction', '1887']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Esperanto', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Proto-Esperanto', 'Category', '(', 'purpose', ')', 'constructed', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Proto-Esperanto', 'Category', '(', 'purpose', ')', 'international', 'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'international', 'auxiliary', 'language']]], 8, True)

['Proto-Esperanto', 'Category', '(', 'purpose', ')', 'Proto-Esperanto']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Proto-Esperanto', 'Created', 'by', 'L', '.', 'L', '.', 'Zamenhof']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'L', '.', 'L', '.', 'Zamenhof']]], 8, True)

['Proto-Esperanto', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Proto-Esperanto', 'Language', 'codes', 'ISO', '639-2', 'art']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Proto-Esperanto', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Old', 'Spoken', 'in', 'Netherlands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Southern', 'Denmark']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Old', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Old', 'Writing', 'system', 'Anglo-Saxon', 'runes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Old', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'ofs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pictish', 'Language', 'extinction', 'by', '900', 'AD']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY', 'NUM', 'AD']], ['NUM', '900']], 6, True)
```

```
['Pictish', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pictish', 'Language', 'codes', 'ISO', '639-3', 'xpi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cumbric', 'Spoken', 'in', 'Southern', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Cumbric', 'Spoken', 'in', 'Cumberland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cumbric', 'Spoken', 'in', 'Westmorland', 'parts', 'of',  
'Northumberland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Westmorland', 'parts', 'of', 'Northumberland']]], 7, True)
```

```
['Cumbric', 'Language', 'codes', 'ISO', '639-3', 'xcb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Camuno', 'Region', 'Italian', 'Peninsula']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Camuno', 'Language', 'codes', 'ISO', '639-3', 'xcc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chemakum', 'Spoken', 'in', 'Olympic', 'Peninsula', ',', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Olympic', 'Peninsula', ',', 'Washington']], 7, True)

['Chemakum', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chemakum', 'Language', 'codes', 'ISO', '639-3', 'cmk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Western', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Western', 'Spoken', 'in', 'Uruguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Bulgaria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Hungary']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3, True)

['Western', 'Writing', 'system', 'Armenian', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'arm', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'arm', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'hye', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hye', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-1', 'hy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Friulan', 'Region', 'Europe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Friulan', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Friulan', 'Writing', 'system', 'Roman', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Friulan', 'Official', 'language', 'in', 'None']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Friulan', 'Regulated', 'by', 'Osservatori', 'Regjon', '\xc3', '\xa2',
'l', 'de', 'Lenghe', 'e', 'de', 'Culture', 'Furlanis']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Osservatori', 'Regjon', '\xc3', '\xa2', 'l', 'de', 'Lenghe', 'e', 'de',
'Culture', 'Furlanis']]], 14, True)
```

```
['Friulan', 'Language', 'codes', 'ISO', '639-2', 'fur']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Friulan', 'Language', 'codes', 'ISO', '639-3', 'fur']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Friulan', 'Language', 'codes', 'Linguasphere', '51-AAA-m']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Swabian', 'Spoken', 'in', 'Germany']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swabian', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Swabian', 'Language', 'codes', 'ISO', '639-3', 'swg']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'northern', 'France']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Old', 'Spoken', 'in', 'parts', 'of', 'Belgium', '(', 'Wallonia', ')',
'and', 'Switzerland']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',
'of', 'Belgium', '\\(', 'Wallonia', '\\)', 'and', 'Switzerland']]], 11,
True)
```

```
['Old', 'Spoken', 'in', 'England']:
```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Ireland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Kingdom', 'of', 'Sicily']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kingdom', 'of', 'Sicily']]], 6, True)

['Old', 'Spoken', 'in', 'Principality', 'of', 'Antioch']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Principality', 'of', 'Antioch']]], 6, True)

['Old', 'Spoken', 'in', 'Kingdom', 'of', 'Cyprus']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kingdom', 'of', 'Cyprus']]], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'fro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'fro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chamorro', 'Spoken', 'in', 'Guam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chamorro', 'Region', 'Western', 'Pacific', 'Ocean']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chamorro', 'Official', 'language', 'in', 'Guam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Chamorro', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Chamorro', 'Language', 'codes', 'ISO', '639-2', 'cha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

```

['Chamorro', 'Language', 'codes', 'ISO', '639-1', 'ch']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sotho', 'Total', 'speakers', 'at', 'least', '5', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'at', 'least', 'NUM', 'million']], ['NUM', '5']], 7, True)

['Sotho', 'Pronunciation', 's', '\xc9', '\xaa', '\xcc', '\x80', 's',
'\xca', '\x8a', '\xcc', '\x81', 't', '\xca', '\xb0', '\xca', '\x8a',
'\xcc', '\x80']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's',
'\xc9', '\xaa', '\xcc', '\x80', 's', '\xca', '\x8a', '\xcc', '\x81', 't',
'\xca', '\xb0', '\xca', '\x8a', '\xcc', '\x80']]], 19, True)

['Sotho', 'Spoken', 'in', 'Lesotho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sotho', 'Spoken', 'in', 'South', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Sotho', 'Official', 'language', 'in', 'Lesotho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Sotho', 'Official', 'language', 'in', 'South', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Sotho', 'Language', 'codes', 'ISO', '639-2', 'sot']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sotho', 'Language', 'codes', 'ISO', '639-3', 'sot']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sotho', 'Language', 'codes', 'ISO', '639-1', 'st']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Northern', 'Austro-Bavarian', 'Total', 'speakers', '\xe2', '\x80',
'\x93']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Northern', 'Region', 'Upper', 'Palatinate']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Region', 'Upper', 'Franconia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Region', 'Upper', 'Bavaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Region', 'Lower', 'Bavaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Sindhi', 'Total', 'speakers', '35', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '35']], 5, True)

['Sindhi', 'Dialects', 'Indo-European']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Dialects', 'Indo-Iranian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Dialects', 'Indo-Aryan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Dialects', 'Northwestern', 'Zone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Dialects', 'Sindhi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'India', '.', 'Also', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'India', '.', 'Also', 'Hong', 'Kong']], 7, True)

['Sindhi', 'Dialects', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Region', 'South', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'India', '.', 'Also', 'Hong', 'Kong']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', '.', 'Also', 'Hong', 'Kong']]], 8, True)

['Sindhi', 'Spoken', 'in', 'Oman']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Singapore']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'UAE']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'UK']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Sri', 'Lanka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Sindhi', 'Writing', 'system', 'Arabic']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Sindhi', 'Writing', 'system', 'Devanagari']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Sindhi', 'Writing', 'system', 'La', '\xe1', '\xb9', '\x87', '\xe1',
'\xb8', '\x8d', '\xc4', '\x81', 'scripts']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'La', '\xe1', '\xb9', '\x87', '\xe1', '\xb8', '\x8d', '\xc4', '\x81',
'scripts']]], 13, True)
```

```
['Sindhi', 'Writing', 'system', 'particularly', 'Gurumukhi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sindhi', 'Official', 'language', 'in', 'Pakistan', '(', 'Sindh', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Pakistan', '\\(', 'Sindh', '\\)']]], 8, True)
```

```
['Sindhi', 'Official', 'language', 'in', 'India']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Sindhi', 'Regulated', 'by', 'Indian', 'Institute', 'of', 'Sindhology',
'(', 'India', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Indian', 'Institute', 'of', 'Sindhology', '\\(', 'India', '\\)']]], 10,
True)
```

```
['Sindhi', 'Language', 'codes', 'ISO', '639-2', 'snd']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Sindhi', 'Language', 'codes', 'ISO', '639-3', 'snd']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sindhi', 'Language', 'codes', 'ISO', '639-1', 'sd']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Jamamad\\xc3\\xad', 'Spoken', 'in', 'Amazonas', 'State', ',', 'Brazil']:
```

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Jamamad\xc3\xad', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Jamamad\xc3\xad', 'Dialects', 'Jarawara']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Jamamad\xc3\xad', 'Dialects', 'Baniwa']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Jamamad\xc3\xad', 'Dialects', 'Kanamanti']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Jamamad\xc3\xad', 'Dialects', 'etc']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Jamamad\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'jaa']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tup\xc3\xad', 'Spoken', 'in', 'Brazil', 'Ethnicity']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tup\xc3\xad', 'Spoken', 'in', 'Tupinamb', '\xc3', '\xa1']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Tupinamb', '\xc3', '\xa1']]], 6, True)

['Tup\xc3\xad', 'Dialects', 'L', '\xc3', '\xad', 'ngua', 'geral', 'amaz', '\xc3', '\xb4', 'nica']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'L', '\xc3', '\xad', 'ngua', 'geral', 'amaz', '\xc3', '\xb4', 'nica']]], 11, True)

['Tup\xc3\xad', 'Writing', 'system', 'Latin', 'alphabet']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tup\xc3\xad', 'Language', 'codes', 'ISO', '639-2', 'sai', '(', 'Collective', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sai', '\\(', 'Collective', '\\)']], ['NUM', '639-2']], 9, True)

['Tup\\xc3\\xad', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tup\\xc3\\xad', 'Language', 'codes', 'ISO', '639-3', 'tpn', '\\xe2', '\\x80', '\\x93', 'Tupinamb', '\\xc3', '\\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tpn', '\\xe2', '\\x80', '\\x93', 'Tupinamb', '\\xc3', '\\xa1']], ['NUM', '639-3']], 12, True)

['Rwanda', 'Spoken', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Rwanda', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Rwanda', 'Spoken', 'in', 'DR', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Rwanda', 'Total', 'speakers', '12', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '12']], 5, True)

['Rwanda', 'Official', 'language', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Rwanda', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Rwanda', 'Language', 'codes', 'ISO', '639-2', 'kin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Rwanda', 'Language', 'codes', 'ISO', '639-3', 'kin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



```

['Rwanda', 'Language', 'codes', 'ISO', '639-1', 'rw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Myanmar', 'Total', 'speakers', 'First', 'language', ':', '32',
'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'First', 'language', ':', 'NUM', 'million']], ['NUM', '32']], 8, True)

['Myanmar', 'Total', 'speakers', 'Second', 'language', ':', '10',
'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'Second', 'language', ':', 'NUM', 'million']], ['NUM', '10']], 8, True)

['Myanmar', 'Spoken', 'in', 'Burma']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Myanmar', 'Spoken', 'in', 'Thailand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Myanmar', 'Spoken', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Myanmar', 'Spoken', 'in', 'Singapore']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Myanmar', 'Writing', 'system', 'Burmese', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Myanmar', 'Pronunciation', 'IPA', ':', 'mj', '\xc9', '\x99', 'm', '\xc3',
'\xa0', 'z', '\xc3', '\xa0', 'or', 'IPA', ':', 'mj', '\xc9', '\x99', 'm',
'\xc3', '\xa0', 'z', '\xc9', '\x99', '\xc9', '\xa1', '\xc3', '\xa1']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'IPA',
':', 'mj', '\xc9', '\x99', 'm', '\xc3', '\xa0', 'z', '\xc3', '\xa0', 'or',
'IPA', ':', 'mj', '\xc9', '\x99', 'm', '\xc3', '\xa0', 'z', '\xc9', '\x99',
'\xc9', '\xa1', '\xc3', '\xa1']]], 29, True)

['Myanmar', 'Official', 'language', 'in', 'Myanmar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

```

```
['Myanmar', 'Language', 'codes', 'ISO', '639-2', 'bur', '(', 'B', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'bur', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Myanmar', 'Language', 'codes', 'ISO', '639-2', 'mya', '(', 'T', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'mya', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Myanmar', 'Language', 'codes', 'ISO', '639-3', 'mya']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Myanmar', 'Language', 'codes', 'ISO', '639-1', 'my']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Wichita', 'Total', 'speakers', '1', 'fluent', ',', '4', 'conversant']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'fluent', ',', 'NUM', 'conversant']], ['NUM', '1'], ['NUM', '4']],
8, True)
```

```
['Wichita', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Wichita', 'Region', 'West-central', 'Oklahoma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Wichita', 'Language', 'codes', 'ISO', '639-3', 'wic']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tajik', 'Spoken', 'in', 'Tajikistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Tajik', 'Spoken', 'in', 'Uzbekistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Tajik', 'Spoken', 'in', 'Russia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Tajik', 'Writing', 'system', 'Cyrillic']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tajik', 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tajik', 'Writing', 'system', 'Perso-Arabic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tajik', 'Official', 'language', 'in', 'Tajikistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Tajik', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tajik', 'Language', 'codes', 'ISO', '639-2', 'tgk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Tajik', 'Language', 'codes', 'ISO', '639-3', 'tgk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tajik', 'Language', 'codes', 'ISO', '639-1', 'tg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Samoan', 'Spoken', 'in', 'Samoa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Samoan', 'Spoken', 'in', 'American', 'Samoa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Samoan', 'Region', 'Spoken', 'as', 'first', 'language', 'in', 'Samoa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Spoken', 'as', 'first', 'language', 'in', 'Samoa']]], 8, True)

['Samoan', 'Region', 'American', 'Samoa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Samoan', 'Region', 'Australia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'U', '.', 'S']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Samoan', 'Region', 'Canada']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'Tokelau']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'Tuvalu']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'Denmark']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'England']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'Japan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Samoan', 'Region', 'China', 'and', 'Germany']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Samoan', 'Regulated', 'by', 'No', 'official', 'regulation']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Samoan', 'Language', 'codes', 'ISO', '639-2', 'smo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Samoan', 'Language', 'codes', 'ISO', '639-3', 'smo']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Samoan', 'Language', 'codes', 'ISO', '639-1', 'sm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Uyghur', 'Total', 'speakers', '8-11', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '8-11']], 5, True)

['Uyghur', 'Region', 'Xinjiang', 'Uyghur', 'Autonomous']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Uyghur', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Kazakhstan', ';', 'also', 'spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kazakhstan', ';', 'also', 'spoken', 'in', 'Afghanistan']], 9, True)

['Uyghur', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Uyghur', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Uyghur', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Uyghur', 'Spoken', 'in', 'and', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```

['Uyghur', 'Writing', 'system', 'Uyghur', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Uyghur', 'Pronunciation', '\xca', '\x94', '\xca', '\x8a', 'j', '\xcb',
'\x88', '\xca', '\x81', '\xca', '\x8a', 'rt', '\xca', '\x83', '\xc9',
'\x9b']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',
'\x94', '\xca', '\x8a', 'j', '\xcb', '\x88', '\xca', '\x81', '\xca',
'\x8a', 'rt', '\xca', '\x83', '\xc9', '\x9b']]], 18, True)

['Uyghur', 'Language', 'codes', 'ISO', '639-2', 'uig']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Uyghur', 'Language', 'codes', 'ISO', '639-3', 'uig']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uyghur', 'Language', 'codes', 'ISO', '639-1', 'ug']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Oromo', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Oromo', 'Spoken', 'in', 'Kenya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Oromo', 'Spoken', 'in', 'Somalia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Oromo', 'Language', 'codes', 'ISO', '639-2', 'orm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', 'orm', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'orm', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

```

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', ':', 'gax', '\xe2', '\x80', '\x93', 'Borana-Arsi-Guji-Walagaa', 'Oromo']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'gax', '\xe2', '\x80', '\x93', 'Borana-Arsi-Guji-Walagaa', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', 'orc', '\xe2', '\x80', '\x93', 'Orma']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'orc', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Oromo', 'Language', 'codes', 'ISO', '639-1', 'om']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Aymara', 'Spoken', 'in', 'Bolivia', ',', 'Peru', 'and', 'Chile']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Bolivia', ',', 'Peru', 'and', 'Chile']], 8, True)

['Aymara', 'Total', 'speakers', '2', ',', '227', ',', '642', 'speakers', 'of', 'Aymara']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', ',', 'NUM', 'speakers', 'of', 'Type: ANY']], ['NUM', '2'], ['NUM', '227'], ['NUM', '642']], 11, True)

['Aymara', 'Language', 'codes', 'ISO', '639-2', 'aym']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Aymara', 'Language', 'codes', 'ISO', '639-3', 'aym', '\xe2', '\x80', '\x93', 'Macrolanguage']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'aym', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Aymara', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Aymara', 'Language', 'codes', 'ISO', '639-1', 'ay']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)



```

['Oriya', 'Region', 'Orissa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Oriya', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oriya', 'Writing', 'system', 'Oriya', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Oriya', 'Official', 'language', 'in', 'Orissa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Oriya', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-2', 'ori']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-3', 'ori']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oriya', 'Language', 'codes', 'Linguasphere', '59-AAF-x']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Oriya', 'Language', 'codes', 'ISO', '639-1', 'or']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Shona', 'Spoken', 'in', 'Zimbabwe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shona', 'Spoken', 'in', 'Mozambique']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shona', 'Spoken', 'in', 'Zambia']:

```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Shona', 'Region', 'Africa']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Shona', 'Language', 'codes', 'ISO', '639-2', 'sna']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Shona', 'Language', 'codes', 'ISO', '639-3', 'sna']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Shona', 'Language', 'codes', 'ISO', '639-1', 'sn']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Aleut', 'Spoken', 'in', 'Alaska', '(', 'Aleutian', 'and', 'Pribilof', 'Islands', ')', ',', 'Kamchatka', 'Krai', '(', 'Commander', 'Islands', ')']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Alaska', '\\\\(', 'Aleutian', 'and', 'Pribilof', 'Islands', '\\\\)', ',', 'Kamchatka', 'Krai', '\\\\(', 'Commander', 'Islands', '\\\\)']]], 17, True)
```

```
['Aleut', 'Writing', 'system', 'Latin', 'alphabet']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aleut', 'Writing', 'system', 'Cyrillic', 'alphabet', '(', 'historic', ')']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Aleut', 'Language', 'codes', 'ISO', '639-2', 'ale']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Aleut', 'Language', 'codes', 'ISO', '639-3', 'ale']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Hindko', 'Region', 'South', 'Asia']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Hindko', 'Spoken', 'in', 'Pakistan', '(', 'Khyber', 'Pakhtunkhwa', ',', 'Punjab', ',', 'Pakistan', 'Administered', 'Kashmir', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Pakistan', '\\\\(', 'Khyber', 'Pakhtunkhwa', ',', 'Punjab', ',', 'Pakistan', 'Administered', 'Kashmir', '\\\\)']]], 14, True)
```

```
['Hindko', 'Spoken', 'in', 'Afghanistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Hindko', 'Writing', 'system', 'Nasta'liq', 'script', ',', 'Devanagari', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Hindko', 'Pronunciation', 'hindkou', '\\xcc', '\\xaf']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'hindkou', '\\xcc', '\\xaf']]], 5, True)
```

```
['Hindko', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Hindko', 'Language', 'codes', 'ISO', '639-3', ':', 'hnd', '\\xe2', '\\x80', '\\x93', 'Southern']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'hnd', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Hindko', 'Language', 'codes', 'ISO', '639-3', 'hno', '\\xe2', '\\x80', '\\x93', 'Northern']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hno', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Lower', 'Spoken', 'in', 'Poland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Lower', 'Spoken', 'in', 'Czech', 'Republic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Lower', 'Spoken', 'in', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Lower', 'Language', 'codes', 'ISO', '639-3', 'sli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Potiguara', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Potiguara', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Potiguara', 'Language', 'codes', 'ISO', '639-3', 'pog']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Belgian', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Belgian', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the',  
'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Belgian', 'Spoken', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Belgian', 'Spoken', 'in', 'Burundi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Belgian', 'Spoken', 'in', 'northern', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Belgian', 'Total', 'speakers', '624', ',', '000', '(', 'Burundi', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\\\(', 'Burundi', '\\\\)']], ['NUM', '624'], ['NUM',  
'000']], 9, True)

['Belgian', 'Official', 'language', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['Belgian', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 9, True)

['Belgian', 'Official', 'language', 'in', 'Rwanda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Belgian', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Belgian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Mari', 'Spoken', 'in', 'Bashkortostan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Tatarstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Udmurtia', ';', 'oblasti', 'Nizhny', 'Novgorod']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Udmurtia', ';', 'oblasti', 'Nizhny', 'Novgorod']]], 8, True)

['Mari', 'Spoken', 'in', 'Kirov']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Sverdlovsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Orenburg', ';', 'Perm', 'Krai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Orenburg', ';', 'Perm', 'Krai']]], 7, True)

['Mari', 'Official', 'language', 'in', 'Mari', 'El']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

```
['Mari', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Mari', 'Language', 'codes', 'ISO', '639-2', 'chm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Mari', 'Language', 'codes', 'ISO', '639-3', 'chm', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chm', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM',  
'639-3']], 10, True)
```

```
['Mari', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)
```

```
['Yerkish', 'Total', 'speakers', '3', '(', 'ape', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'ape', '\\)']]], ['NUM', '3']], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'Animal', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'Yerkish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)
```

```
['Yerkish', 'Created', 'by', 'Ernst', 'von', 'Glasersfeld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Ernst',  
'von', 'Glasersfeld']]], 6, True)
```

```
['Yerkish', 'Category', '(', 'sources', ')', 'Use', 'a', 'keyboard', 'to',  
'punch', 'keys', 'with', 'so', 'called', 'lexigrams', ',', 'symbols',  
'corresponding', 'to', 'objects', 'or', 'ideas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Use', 'a', 'keyboard', 'to', 'punch', 'keys', 'with',  
'so', 'called', 'lexigrams', ',', 'symbols', 'corresponding', 'to',  
'objects', 'or', 'ideas']]], 22, True)
```

```
['Yerkish', 'Setting', 'and', 'usage', 'Animal', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Type: ANY', 'language']]], 6, True)
```

```
['Yerkish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Sassarese', 'Region', 'Sardinia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sassarese', 'Spoken', 'in', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sassarese', 'Pronunciation', 'sas', '\xcb', '\x90', 'a', '\xcb', '\x88',  
'rezu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sas',  
'\xcb', '\x90', 'a', '\xcb', '\x88', 'rezu']]], 9, True)
```

```
['Sassarese', 'Official', 'language', 'in', 'Sardinia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Sassarese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Sassarese', 'Language', 'codes', 'ISO', '639-2', 'srd', 'or', 'cos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'srd', 'or', 'cos']], ['NUM', '639-2']], 8, True)
```

```
['Sassarese', 'Language', 'codes', 'ISO', '639-3', 'sdc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sassarese', 'Language', 'codes', 'Linguasphere', '51-AAA-pe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Sassarese', 'Language', 'codes', 'ISO', '639-1', 'sc', 'or', 'co']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sc', 'or', 'co']], ['NUM', '639-1']], 8, True)
```

```
['Hittite', 'Spoken', 'in', 'Anatolia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Hittite', 'Language', 'codes', 'ISO', '639-2', 'hit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Hittite', 'Language', 'codes', 'ISO', '639-3', 'hit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Istriot', 'Spoken', 'in', 'Istria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Istriot', 'Language', 'codes', 'ISO', '639-3', 'ist']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Istriot', 'Language', 'codes', 'Linguasphere', '51-AAA-na']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Cebuano', 'Total', 'speakers', 'About', '20', 'million', 'people', 'and',  
'the', 'second', 'most', 'spoken', 'language', 'in', 'the', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'About', 'NUM', 'million', 'people', 'and', 'the', 'second', 'most',  
'spoken', 'language', 'in', 'the', 'Philippines']]], ['NUM', '20']], 16,  
True)
```

```
['Cebuano', 'Dialects', 'Cebuano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cebuano', 'Dialects', 'Boholano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cebuano', 'Dialects', 'Leytehanon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cebuano', 'Dialects', 'Mindanao', 'Visayan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```



['Cebuano', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cebuano', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cebuano', 'Writing', 'system', 'Latin', '(', 'Cebuano', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Cebuano', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Cebuano', 'Regulated', 'by', 'Visayan', 'Academy', 'of', 'Arts', 'and', 'Letters']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Visayan', 'Academy', 'of', 'Arts', 'and', 'Letters']], 9, True)

['Cebuano', 'Language', 'codes', 'ISO', '639-2', 'ceb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cebuano', 'Language', 'codes', 'ISO', '639-3', 'ceb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huilliche', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Huilliche', 'Language', 'codes', 'ISO', '639-3', 'huh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mapuche', 'Spoken', 'in', 'Chile', ',', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Mapuche', 'Language', 'codes', 'ISO', '639-2', 'arn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Mam', 'Spoken', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mam', 'Spoken', 'in', 'Mexico']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mam', 'Region', 'Quetzaltenango', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mam', 'Region', 'Huehuetenango', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mam', 'Region', 'San', 'Marcos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mam', 'Region', 'Chiapas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Mam', 'Language', 'codes', 'ISO', '639-3', 'mam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lingua', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Lingua', 'Created', 'by', 'C', '.', 'George', 'Boeree']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'C',
'.', 'George', 'Boeree']]], 7, True)
```

```
['Lingua', 'Category', '(', 'sources', ')', 'based', 'on', 'Romance',
'and', 'Creole', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'sources', '\\\\)', 'based', 'on', 'Romance', 'and', 'Creole',
'languages']]], 11, True)
```

```
['Lingua', 'Language', 'codes', 'ISO', '639-3', 'lfn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```

['Silesian', 'Total', 'speakers', 'no', 'data']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'no', 'data']]], 5, True)

['Silesian', 'Language', 'codes', 'ISO', '639-3', 'szl']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Silesian', 'Language', 'codes', 'Linguasphere', '53-AAA-cck']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Cherokee', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Cherokee', 'Writing', 'system', 'Cherokee', 'syllabary', ',', 'Latin',
'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Cherokee', 'Language', 'codes', 'ISO', '639-2', 'chr']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cherokee', 'Language', 'codes', 'ISO', '639-3', 'chr']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yor\u03b9\u0301\u0302\u0303', 'Spoken', 'in', 'Nigeria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yor\u03b9\u0301\u0302\u0303', 'Spoken', 'in', 'Togo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yor\u03b9\u0301\u0302\u0303', 'Spoken', 'in', 'Benin']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yor\u03b9\u0301\u0302\u0303', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

```

['Yor\xc3\xb9b\xc3\xa1', 'Official', 'language', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Yor\xc3\xb9b\xc3\xa1', 'Regulated', 'by', 'No', 'official',  
'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Yor\xc3\xb9b\xc3\xa1', 'Language', 'codes', 'ISO', '639-2', 'yor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Yor\xc3\xb9b\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'yor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yor\xc3\xb9b\xc3\xa1', 'Language', 'codes', 'ISO', '639-1', 'yo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Chinantec', 'Total', 'speakers', 'approx', '.', '93', ',', '000', '(',  
'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'approx', '.', 'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM',  
'93'], ['NUM', '000']], 11, True)

['Chinantec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chinantec', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', ':', 'cco', '\xe2',  
'\x80', '\x93', 'Comaltepec', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'cco', '\xe2', '\x80', '\x93', 'Comaltepec', 'Type:  
ANY']]], ['NUM', '639-3']], 12, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'chj', '\xe2', '\x80',

'\x93', 'Ojitl', '\xc3', '\xa1', 'n', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chj', '\xe2', '\x80', '\x93', 'Ojitl', '\xc3', '\xa1', 'n',  
'Type: ANY']], ['NUM', '639-3']], 14, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'chq', '\xe2', '\x80',  
'\x93', 'Quiotepec', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chq', '\xe2', '\x80', '\x93', 'Quiotepec', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'chz', '\xe2', '\x80',  
'\x93', 'Ozumac', '\xc3', '\xad', 'n', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chz', '\xe2', '\x80', '\x93', 'Ozumac', '\xc3', '\xad', 'n',  
'Type: ANY']], ['NUM', '639-3']], 14, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cle', '\xe2', '\x80',  
'\x93', 'Lealao', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cle', '\xe2', '\x80', '\x93', 'Lealao', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cnl', '\xe2', '\x80',  
'\x93', 'Lalana', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cnl', '\xe2', '\x80', '\x93', 'Lalana', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cnt', '\xe2', '\x80',  
'\x93', 'Tepetotutla', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cnt', '\xe2', '\x80', '\x93', 'Tepetotutla', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cpa', '\xe2', '\x80',  
'\x93', 'Palantla', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cpa', '\xe2', '\x80', '\x93', 'Palantla', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'csa', '\xe2', '\x80',  
'\x93', 'Chiltepec', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'csa', '\xe2', '\x80', '\x93', 'Chiltepec', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cso', '\xe2', '\x80', '\x93', 'Sochiapan', 'Chinantec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cso', '\xe2', '\x80', '\x93', 'Sochiapan', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cte', '\xe2', '\x80', '\x93', 'Tepinapa', 'Chinantec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cte', '\xe2', '\x80', '\x93', 'Tepinapa', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'ctl', '\xe2', '\x80', '\x93', 'Tlacoatzintepec', 'Chinantec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ctl', '\xe2', '\x80', '\x93', 'Tlacoatzintepec', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cuc', '\xe2', '\x80', '\x93', 'Usila', 'Chinantec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cuc', '\xe2', '\x80', '\x93', 'Usila', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cvn', '\xe2', '\x80', '\x93', 'Valle', 'Nacional', 'Chinantec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cvn', '\xe2', '\x80', '\x93', 'Valle', 'Nacional', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Taishanese', 'Spoken', 'in', 'Southern', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Taishanese', 'Spoken', 'in', 'Hong', 'Kong']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Taishanese', 'Spoken', 'in', 'Canada', 'and', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Taishanese', 'Region', 'western', 'and', 'southern', 'Guangdong', ';', 'the', 'Pearl', 'River', 'Delta', ';', 'parts', 'of', 'Hainan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western',

'and', 'southern', 'Guangdong', ';', 'the', 'Pearl', 'River', 'Delta', ';',  
'parts', 'of', 'Type: ANY']]], 15, True)

['Taishanese', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Taishanese', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Taishanese', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',  
'\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Taishanese', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Baoule', 'Language', 'codes', 'ISO', '639-3', 'bci']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['J\\xc3\\xa8rriais', 'Spoken', 'in', 'Jersey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['J\\xc3\\xa8rriais', 'Spoken', 'in', 'Sark']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['J\\xc3\\xa8rriais', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',  
'\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['J\\xc3\\xa8rriais', 'Language', 'codes', 'Linguasphere', '51-AAA-hc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Papiamentu', 'Spoken', 'in', 'Aruba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Papiamentu', 'Spoken', 'in', 'Cura', '\\xc3', '\\xa7', 'ao']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Cura',  
'\xc3', '\xa7', 'ao']]], 7, True)

['Papiamento', 'Spoken', 'in', 'Caribbean', 'Netherlands', '(' , 'Bonaire',  
)']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Caribbean', 'Netherlands', '\\(', 'Bonaire', '\\)']]], 8, True)

['Papiamento', 'Region', 'Caribbean', 'islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Papiamento', 'Official', 'language', 'in', 'Aruba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Papiamento', 'Official', 'language', 'in', 'Cura', '\xc3', '\xa7', 'ao']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Cura', '\xc3', '\xa7', 'ao']]], 8, True)

['Papiamento', 'Official', 'language', 'in', 'Caribbean', 'Netherlands',  
'(' , 'Bonaire', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Caribbean', 'Netherlands', '\\(', 'Bonaire', '\\)']]], 9, True)

['Papiamento', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Papiamento', 'Language', 'codes', 'ISO', '639-2', 'pap']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Papiamento', 'Language', 'codes', 'ISO', '639-3', 'pap']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papiamento', 'Language', 'codes', 'Linguasphere', '51-AAC-be']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Taa', 'Spoken', 'in', 'Botswana', ',,', 'Namibia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',,', 'Type: ANY']]], 6, True)

['Taa', 'Region', 'Southern', 'Ghanzi']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taa', 'Region', 'northern', 'Kgalagadi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taa', 'Language', 'codes', 'ISO', '639-3', 'nmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vedic', 'Spoken', 'in', 'Bronze', 'Age', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Bronze', 'Age', 'India']]], 6, True)

['Vedic', 'Spoken', 'in', 'Iron', 'Age', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iron', 'Age', 'India']]], 6, True)

['Vedic', 'Spoken', 'in', 'Greater', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Vedic', 'Language', 'codes', 'ISO', '639-3', 'san']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sinhala', 'Spoken', 'in', 'Sri', 'Lanka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sinhala', 'Official', 'language', 'in', 'Sri', 'Lanka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Sinhala', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Sinhala', 'Language', 'codes', 'ISO', '639-2', 'sin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sinhala', 'Language', 'codes', 'ISO', '639-3', 'sin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Sinhala', 'Language', 'codes', 'ISO', '639-1', 'si']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Balinese', 'Spoken', 'in', 'Bali']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Balinese', 'Spoken', 'in', 'Nusa', 'Penida']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Balinese', 'Spoken', 'in', 'Lombok', 'and', 'Java']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']]], 6, True)

['Balinese', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Balinese', 'Writing', 'system', 'Latin', ',', 'Balinese']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Balinese', 'Language', 'codes', 'ISO', '639-2', 'ban']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Balinese', 'Language', 'codes', 'ISO', '639-3', 'ban']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buginese', 'Spoken', 'in', 'Certain', 'Parts', 'of', 'Sumatra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Certain', 'Parts', 'of', 'Sumatra']]], 7, True)

['Buginese', 'Spoken', 'in', 'Riau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Buginese', 'Spoken', 'in', 'Kalimantan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Buginese', 'Spoken', 'in', 'Sabah']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Buginese', 'Spoken', 'in', 'Malay', 'peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Buginese', 'Writing', 'system', 'Lontara', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', ',', '']], 5, True)

['Buginese', 'Writing', 'system', 'Roman', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Buginese', 'Language', 'codes', 'ISO', '639-2', 'bug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Buginese', 'Language', 'codes', 'ISO', '639-3', 'bug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gayo', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gayo', 'Region', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Gayo', 'Language', 'codes', 'ISO', '639-2', 'gay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gayo', 'Language', 'codes', 'ISO', '639-3', 'gay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iban', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iban', 'Region', 'Borneo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Iban', 'Language', 'codes', 'ISO', '639-2', 'iba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Iban', 'Language', 'codes', 'ISO', '639-3', 'iba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Javanese', 'Spoken', 'in', 'Java', '(', 'Indonesia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']], 7, True)

['Javanese', 'Spoken', 'in', 'Peninsular', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Javanese', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Javanese', 'Spoken', 'in', 'New', 'Caledonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Javanese', 'Total', 'speakers', 'about', '80', 'million', 'total']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'about', 'NUM', 'million', 'total']], ['NUM', '80']], 7, True)

['Javanese', 'Writing', 'system', 'Javanese', 'script', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Javanese', 'Writing', 'system', 'Arabic', 'script', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Javanese', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Javanese', 'Language', 'codes', 'ISO', '639-2', 'jav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Javanese', 'Language', 'codes', 'ISO', '639-3', 'variously']:

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', ':', 'jav', '\xe2', '\x80', '\x93', 'Javanese']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'jav', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', 'jvn', '\xe2', '\x80', '\x93', 'Caribbean', 'Javanese']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'jvn', '\xe2', '\x80', '\x93', 'Caribbean', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', 'tes', '\xe2', '\x80', '\x93', 'Tenggerese']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tes', '\xe2', '\x80', '\x93', 'Tenggerese']], ['NUM', '639-3']], 10, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-1', 'jv']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Kawi', 'Spoken', 'in', 'Java']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kawi', 'Spoken', 'in', 'Bali']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kawi', 'Spoken', 'in', 'Madura']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kawi', 'Spoken', 'in', 'Lombok']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kawi', 'Spoken', 'in', 'Indonesia', ';', 'Philippines']:  
(['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Indonesia', ';', 'Philippines']], 6, True)
```

```
['Kawi', 'Region', 'Malay', 'Archipelago']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kawi', 'Language', 'codes', 'ISO', '639-2', 'kaw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kawi', 'Language', 'codes', 'ISO', '639-3', 'kaw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Madurese', 'Spoken', 'in', 'Island', 'of', 'Madura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Island', 'of', 'Madura']]], 6, True)

['Madurese', 'Spoken', 'in', 'Sapudi', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Madurese', 'Spoken', 'in', 'northern', 'coastal', 'area', 'of', 'eastern', 'Java']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'northern', 'coastal', 'area', 'of', 'eastern', 'Java']]], 9, True)

['Madurese', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Madurese', 'Language', 'codes', 'ISO', '639-2', 'mad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Madurese', 'Language', 'codes', 'ISO', '639-3', 'mad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Makassar', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Makassar', 'Writing', 'system', 'Lontara', ',', 'Roman', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', ',', 'Type: ANY', 'alphabet']]], 7, True)

['Makassar', 'Language', 'codes', 'ISO', '639-2', 'mak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Makassar', 'Language', 'codes', 'ISO', '639-3', 'mak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minangkabau', 'Spoken', 'in', 'Indonesia', ',', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Minangkabau', 'Region', 'West', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Minangkabau', 'Region', 'Riau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Minangkabau', 'Region', 'Jambi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Minangkabau', 'Region', 'Bengkulu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Minangkabau', 'Region', 'North', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Minangkabau', 'Region', 'Aceh', '(', 'Indonesia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Aceh', '\\(',  
'Indonesia', '\\)']]], 6, True)

['Minangkabau', 'Language', 'codes', 'ISO', '639-2', 'min']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Minangkabau', 'Language', 'codes', 'ISO', '639-3', 'min']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nicaraguan', 'Region', 'The', 'Managua', 'region', ',', 'but',  
'spreading', 'throughout', 'the', 'country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'The',  
'Managua', 'region', ',', 'but', 'spreading', 'throughout', 'the',

'country']]], 11, True)

['Nicaraguan', 'Signed', 'in', 'Nicaragua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bashkir', 'Spoken', 'in', 'Russia', ',', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Bashkir', 'Total', 'speakers', 'approx', '.', '1', ',', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'approx', '.', 'NUM', ',', 'NUM', 'million']], ['NUM', '1'], ['NUM', '5']], 9, True)

['Bashkir', 'Official', 'language', 'in', 'Bashkortostan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Bashkir', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Bashkir', 'Language', 'codes', 'ISO', '639-2', 'bak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bashkir', 'Language', 'codes', 'ISO', '639-3', 'bak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bashkir', 'Language', 'codes', 'ISO', '639-1', 'ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Afar', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afar', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afar', 'Spoken', 'in', 'Djibouti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



['Afar', 'Writing', 'system', 'Ethiopic', '(', 'in', 'Ethiopia', ')', '/', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Ethiopic', '\\\\(', 'in', 'Ethiopia', '\\\\)'], '/', 'Type: ANY']]], 10, True)

['Afar', 'Language', 'codes', 'ISO', '639-2', 'aar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Afar', 'Language', 'codes', 'ISO', '639-3', 'aar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Afar', 'Language', 'codes', 'ISO', '639-1', 'aa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dzongkha', 'Spoken', 'in', 'Bhutan', ',', 'Sikkim', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Bhutan', ',', 'Sikkim', '\\\\(', 'India', '\\\\)']]], 9, True)

['Dzongkha', 'Writing', 'system', 'Tibetan', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dzongkha', 'Official', 'language', 'in', 'Bhutan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Dzongkha', 'Language', 'codes', 'ISO', '639-2', 'dzo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dzongkha', 'Language', 'codes', 'ISO', '639-3', 'dzo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dzongkha', 'Language', 'codes', 'ISO', '639-1', 'dz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Fijian', 'Total', 'speakers', '450', ',', '000', 'native', 'speakers', ',', '200', ',', '000', 'second-language', 'users']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'native', 'speakers', ',', 'NUM', ',', 'NUM', 'second-

language', 'users']], ['NUM', '450'], ['NUM', '000'], ['NUM', '200'],  
['NUM', '000']], 14, True)

['Fijian', 'Spoken', 'in', 'Fiji']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fijian', 'Region', 'Spoken', 'as', 'first', 'language', 'on', 'Vanua',  
'Levu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Spoken', 'as',  
'first', 'language', 'on', 'Vanua', 'Levu']]], 9, True)

['Fijian', 'Region', 'the', 'eastern', 'half', 'of', 'Viti', 'Levu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'the',  
'eastern', 'half', 'of', 'Viti', 'Levu']]], 8, True)

['Fijian', 'Region', 'and', 'on', 'the', 'lesser', 'islands', 'of',  
'Kadavu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'on',  
'the', 'lesser', 'islands', 'of', 'Kadavu']]], 9, True)

['Fijian', 'Region', 'Nayau']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Fijian', 'Region', 'Lakeba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Fijian', 'Region', 'Oneata']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Fijian', 'Region', 'Moce']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Fijian', 'Region', 'Komo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Fijian', 'Region', 'Namuka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Fijian', 'Region', 'Kabara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Vulaga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Official', 'language', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Fijian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Fijian', 'Language', 'codes', 'ISO', '639-2', 'fij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Fijian', 'Language', 'codes', 'ISO', '639-3', 'fij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Fijian', 'Language', 'codes', 'ISO', '639-1', 'fj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Kashmiri', 'Spoken', 'in', 'Jammu', 'and', 'Kashmir', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Jammu', 'and', 'Kashmir', '\\\\(', 'India', '\\\\)']]], 9, True)

['Kashmiri', 'Spoken', 'in', 'Azad', 'Kashmir', '(', 'Pakistan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Azad', 'Kashmir', '\\\\(', 'Pakistan', '\\\\)']]], 8, True)

['Kashmiri', 'Writing', 'system', 'Perso-Arabic', 'script', '(', 'contemporary', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Perso-Arabic', 'script', '\\\\(', 'contemporary', '\\\\)', ',']]], 9, True)

['Kashmiri', 'Writing', 'system', 'Devanagari', 'script', '(', 'contemporary', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devanagari', 'script', '\\\\(', 'contemporary', '\\\\)', ',']]], 9, True)

['Kashmiri', 'Writing', 'system', 'Sharada', 'script', '(', 'ancient', '/',

```

'liturgical', ')]':
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Sharada', 'script', '\\(', 'ancient', '/', 'liturgical', '\\)']]], 10,
True)

['Kashmiri', 'Official', 'language', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Kashmiri', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Kashmiri', 'Language', 'codes', 'ISO', '639-2', 'kas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kashmiri', 'Language', 'codes', 'ISO', '639-3', 'kas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kashmiri', 'Language', 'codes', 'ISO', '639-1', 'ks']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Rundi', 'Spoken', 'in', 'Burundi', ',', 'Uganda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Rundi', 'Region', 'Central', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Rundi', 'Official', 'language', 'in', 'Burundi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Rundi', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Rundi', 'Language', 'codes', 'ISO', '639-2', 'run']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Rundi', 'Language', 'codes', 'ISO', '639-3', 'run']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rundi', 'Language', 'codes', 'ISO', '639-1', 'rn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Latino', 'sine', 'flexione', 'Total', 'speakers', 'virtually', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'sine', 'flexione', 'Total', 'speakers', 'virtually', 'extinct']]], 7, True)

['Latino', 'Date', 'founded', '1903']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Latino', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Latino', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Latino', 'Created', 'by', 'Giuseppe', 'Peano']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Latino', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Latino', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Ligurian', 'Spoken', 'in', 'Liguria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ligurian', 'Language', 'codes', 'ISO', '639-3', 'xlg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sango', 'Region', 'Central', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sango', 'Spoken', 'in', 'Central', 'African', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Republic']]], 6, True)

['Sango', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sango', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Sango', 'Pronunciation', 'j', '\xc3', '\xa1', '\xc5', '\x8b', '\xc9', '\xa1', '\xc3', '\xa1', 't', '\xc3', '\xad', 's', '\xc4', '\x81', '\xc5', '\x8b', '\xc9', '\xa1', '\xc5', '\x8d']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'j', '\xc3', '\xa1', '\xc5', '\x8b', '\xc9', '\xa1', '\xc3', '\xa1', 't', '\xc3', '\xad', 's', '\xc4', '\x81', '\xc5', '\x8b', '\xc9', '\xa1', '\xc5', '\x8d']]], 23, True)

['Sango', 'Regulated', 'by', 'Institute', 'of', 'Applied', 'Linguistics']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Institute', 'of', 'Applied', 'Linguistics']]], 7, True)

['Sango', 'Language', 'codes', 'ISO', '639-2', 'sag', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sag', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Sango', 'Language', 'codes', 'ISO', '639-2', 'sag', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sag', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Sango', 'Language', 'codes', 'ISO', '639-3', 'sag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sango', 'Language', 'codes', 'ISO', '639-1', 'sg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Swazi', 'Spoken', 'in', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Swazi', 'Spoken', 'in', 'South', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Swazi', 'Spoken', 'in', 'Lesotho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Swazi', 'Spoken', 'in', 'Mozambique']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Swazi', 'Total', 'speakers', '3', ',', '000', ',', '000', '(',
'Ethnologue', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']]], ['NUM', '3'],
['NUM', '000'], ['NUM', '000']], 11, True)

['Swazi', 'Official', 'language', 'in', 'Swaziland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Swazi', 'Official', 'language', 'in', 'South', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swazi', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Swazi', 'Language', 'codes', 'ISO', '639-2', 'ssw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Swazi', 'Language', 'codes', 'ISO', '639-3', 'ssw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Swazi', 'Language', 'codes', 'ISO', '639-1', 'ss']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Tongan', 'Spoken', 'in', 'Tonga']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

['Tongan', 'Spoken', 'in', 'American', 'Samoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tongan', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tongan', 'Spoken', 'in', 'Niue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Official', 'language', 'in', 'Tonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Tongan', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tongan', 'Language', 'codes', 'ISO', '639-2', 'ton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tongan', 'Language', 'codes', 'ISO', '639-3', 'ton']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tongan', 'Language', 'codes', 'ISO', '639-1', 'to']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tsonga', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsonga', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tsonga', 'Spoken', 'in', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsonga', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsonga', 'Region', 'Limpopo', ',', 'Mpumalanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tsonga', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Tsonga', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Tsonga', 'Language', 'codes', 'ISO', '639-2', 'tso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tsonga', 'Language', 'codes', 'ISO', '639-3', 'tso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsonga', 'Language', 'codes', 'ISO', '639-1', 'ts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

```
['Zhuang', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Zhuang', 'Total', 'speakers', '14', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '14']], 5, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-2', 'zha']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zha', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zha', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zhd', '\xe2', '\x80',
'\x93', 'Dai', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zhd', '\xe2', '\x80', '\x93', 'Dai', 'Type: ANY']], ['NUM',
'639-3']], 11, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zgb', '\xe2', '\x80',
'\x93', 'Guibei', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zgb', '\xe2', '\x80', '\x93', 'Guibei', 'Type: ANY']],
['NUM', '639-3']], 11, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zgn', '\xe2', '\x80',
'\x93', 'Guibian', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zgn', '\xe2', '\x80', '\x93', 'Guibian', 'Type: ANY']],
['NUM', '639-3']], 11, True)
```

```
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zln', '\xe2', '\x80',
'\x93', 'Lianshan', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zln', '\xe2', '\x80', '\x93', 'Lianshan', 'Type: ANY']],
['NUM', '639-3']], 11, True)
```

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zlj', '\xe2', '\x80', '\x93', 'Liujiang', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zlj', '\xe2', '\x80', '\x93', 'Liujiang', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zfq', '\xe2', '\x80', '\x93', 'Liuqian', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zfq', '\xe2', '\x80', '\x93', 'Liuqian', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zgm', '\xe2', '\x80', '\x93', 'Minz', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zgm', '\xe2', '\x80', '\x93', 'Minz', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zhn', '\xe2', '\x80', '\x93', 'Nong', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zhn', '\xe2', '\x80', '\x93', 'Nong', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zqe', '\xe2', '\x80', '\x93', 'Qiubei', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zqe', '\xe2', '\x80', '\x93', 'Qiubei', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyg', '\xe2', '\x80', '\x93', 'Yang', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zyg', '\xe2', '\x80', '\x93', 'Yang', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyb', '\xe2', '\x80', '\x93', 'Yongbei', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zyb', '\xe2', '\x80', '\x93', 'Yongbei', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyn', '\xe2', '\x80', '\x93', 'Yongnan', 'Zhuang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

['ISO', 'NUM', 'zyn', '\xe2', '\x80', '\x93', 'Yongnan', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyj', '\xe2', '\x80',  
\x93', 'Youjiang', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zyj', '\xe2', '\x80', '\x93', 'Youjiang', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zzj', '\xe2', '\x80',  
\x93', 'Zuojiang', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zzj', '\xe2', '\x80', '\x93', 'Zuojiang', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-1', 'za']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['N\xc7\x80u', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['N\xc7\x80u', 'Total', 'speakers', '10']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['N\xc7\x80u', 'Language', 'codes', 'ISO', '639-3', 'ngh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auslan', 'Signed', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Algonquin', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Algonquin', 'Language', 'codes', 'ISO', '639-3', 'alq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['French', 'Signed', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['French', 'Signed', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Togo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dacian', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dacian', 'Spoken', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dacian', 'Spoken', 'in', 'parts', 'of', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts', 'of', 'Ukraine']], 6, True)

['Dacian', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dacian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dacian', 'Language', 'codes', 'ISO', '639-3', 'xdc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Galatian', 'Spoken', 'in', 'Galatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Galatian', 'Language', 'codes', 'ISO', '639-3', 'xga', 'Map', 'showing', 'the', 'Roman', 'province', 'of', 'Galatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'xga', 'Map', 'showing', 'the', 'Roman', 'province', 'of', 'Galatia']], ['NUM', '639-3']], 13, True)

['Eyak', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Eyak', 'Language', 'extinction', '2008']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Eyak', 'Region', 'Cordova', ',', 'Alaska']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Eyak', 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Eyak', 'Pronunciation', '\xca', '\x94', 'i', '\xcb', '\x90', 'ja', '\xcb', '\x90', 'q']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca', '\x94', 'i', '\xcb', '\x90', 'ja', '\xcb', '\x90', 'q']], 11, True)

['Eyak', 'Language', 'codes', 'ISO', '639-3', 'eya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wenedyk', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Wenedyk', 'Date', 'founded', '2002']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'alternative', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'Wenedyk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]]], 6, True)
```

```
['Wenedyk', 'Created', 'by', 'Jan', 'van', 'Steenbergen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Jan',  
'van', 'Steenbergen']]]], 6, True)
```

```
['Wenedyk', 'Category', '(', 'sources', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Type: ANY', 'Type: ANY']]]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'sources', ')', 'a', 'posteriori',  
'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'a', 'posteriori', 'language']]]], 8, True)
```

```
['Wenedyk', 'Setting', 'and', 'usage', 'A', 'thought', 'experiment', 'in',  
'the', 'alternate', 'history']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'A', 'thought', 'experiment', 'in', 'the', 'alternate',  
'history']]]], 11, True)
```

```
['Wenedyk', 'Setting', 'and', 'usage', 'Ill', 'Bethisad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Ill', 'Bethisad']]]], 6, True)
```

```
['Wenedyk', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']]], 6, True)
```

```
['Wenedyk', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']]], 8, True)
```

```
['Sranan', 'Spoken', 'in', 'Suriname']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sranan', 'Language', 'codes', 'ISO', '639-2', 'srn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Sranan', 'Language', 'codes', 'ISO', '639-3', 'srn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sranan', 'Language', 'codes', 'Linguasphere', '52-ABB-aw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
["Martha's", 'Signed', 'in', 'formerly', 'in', 'the', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in',  
'formerly', 'in', 'the', 'United', 'States']]], 8, True)
```

```
["Martha's", 'Language', 'codes', 'ISO', '639-2', 'sgn-US-MA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Khoekhoegowab', 'Region', 'Orange', 'River', ',', 'Great', 'Namaland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Orange',  
'River', ',', 'Great', 'Namaland']]], 7, True)
```

```
['Khoekhoegowab', 'Official', 'language', 'in', 'National', 'language',  
'in', 'Namibia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'National', 'language', 'in', 'Namibia']]], 8, True)
```

```
['Khoekhoegowab', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Khoekhoegowab', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Khoekhoegowab', 'Language', 'codes', 'ISO', '639-3', ':', 'hgm', '\xe2',  
'\x80', '\x93', 'Hai', '\xc7', '\x81', 'om']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'hgm', '\xe2', '\x80', '\x93', 'Hai', '\xc7', '\x81',
```



'om']], ['NUM', '639-3']], 14, True)

['Khoekhoegowab', 'Language', 'codes', 'ISO', '639-3', 'naq', '\xe2',  
\x80', \x93', 'Nama']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'naq', '\xe2', '\x80', '\x93', 'Nama']], ['NUM', '639-3']],  
10, True)

['Khoekhoegowab', 'Language', 'codes', 'ISO', '639-3', 'The',  
'distribution', 'of', 'the', 'Nama', 'language', 'in', 'Namibia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'The', 'distribution', 'of', 'the', 'Nama', 'language', 'in',  
'Namibia']], ['NUM', '639-3']], 13, True)

['Avar', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Avar', 'Spoken', 'in', 'Azerbaijan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Avar', 'Spoken', 'in', 'Kazakhstan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Avar', 'Spoken', 'in', 'Georgia', 'and', 'Turkey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']], 6, True)

['Avar', 'Region', 'Republic', 'of', 'Dagestan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Avar', 'Language', 'codes', 'ISO', '639-2', 'ava']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Avar', 'Language', 'codes', 'ISO', '639-3', 'ava']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Avar', 'Language', 'codes', 'ISO', '639-1', 'av']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ruthenian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Ruthenian', 'Language', 'codes', 'ISO', '639-3', 'orv-olr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Ancient', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Old', 'Writing', 'system', 'Old', 'Persian', 'Cuneiform']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'peo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'peo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Europanto', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Europanto', 'Setting', 'and', 'usage', 'European', 'Union', 'administration']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'European', 'Union', 'administration']], 7, True)

['Europanto', 'Created', 'by', 'Diego', 'Marani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Europanto', 'Category', '(', 'purpose', ')', 'jest']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY']], 6, True)

['Europanto', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',

'purpose', '\\)', 'int', '.', 'auxiliary', 'languages']]], 9, True)

['Europanto', 'Category', '(', 'purpose', ')', 'Europanto']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Europanto', 'Date', 'founded', '1996']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Europanto', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Europanto', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Middle', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'dum']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'dum']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'southern', 'Germany', '(', 'south', 'of',  
'the', 'Benrath', 'line', ')', ',', 'parts', 'of', 'Austria', 'and',  
'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'southern', 'Germany', '\\(', 'south', 'of', 'the', 'Benrath', 'line', '\\  
)', ',', 'parts', 'of', 'Austria', 'and', 'Switzerland']]], 18, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'gmh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'gmh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Early', 'Spoken', 'in', 'southern', 'Germany', '(', 'south', 'of', 'the',

'Benrath', 'line', ')', ', ', 'parts', 'of', 'Austria', 'and',  
'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'southern', 'Germany', '\\(', 'south', 'of', 'the', 'Benrath', 'line', '\\  
)', ', ', 'parts', 'of', 'Austria', 'and', 'Switzerland']]], 18, True)

['Early', 'Language', 'codes', 'ISO', '639-2', 'deu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Early', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

["Dena'ina", 'Total', 'speakers', '75']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Dena'ina", 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

["Dena'ina", 'Writing', 'system', 'Latin', '(', "Dena'ina", 'variant',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

["Dena'ina", 'Language', 'codes', 'ISO', '639-3', 'tfn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ilocano', 'Spoken', 'in', 'Philippines', ', ']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ', ']]], 5, True)

['Ilocano', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ilocano', 'Region', 'Northern', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ilocano', 'Writing', 'system', 'Historically', 'written', 'in',  
'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Ilocano', 'Official', 'language', 'in', 'Regional', 'language', 'in',  
'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Ilocano', 'Language', 'codes', 'ISO', '639-2', 'ilo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ilocano', 'Language', 'codes', 'ISO', '639-3', 'ilo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Phoenician', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Formerly', 'spoken', 'in', 'Lebanon']]], 7, True)

['Phoenician', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'Malta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'Sicily']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Phoenician', 'Spoken', 'in', 'and', 'other', 'coastal', 'outposts',  
'and', 'islands', 'throughout', 'the', 'Mediterranean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',  
'other', 'coastal', 'outposts', 'and', 'islands', 'throughout', 'the',  
'Mediterranean']]], 12, True)

['Phoenician', 'Writing', 'system', 'Phoenician', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Phoenician', 'Language', 'codes', 'ISO', '639-2', 'phn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Phoenician', 'Language', 'codes', 'ISO', '639-3', 'phn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zauruan', 'Spoken', 'in', 'Nauru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Zauruan', 'Language', 'codes', 'ISO', '639-2', 'nau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zauruan', 'Language', 'codes', 'ISO', '639-3', 'nau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zauruan', 'Language', 'codes', 'ISO', '639-1', 'na']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Aequian', 'Spoken', 'in', 'Aequian', 'country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Aequian', 'Region', 'Latium', ',', 'east-central', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Latium', ',',  
'east-central', 'Italy']]], 6, True)

['Alemannic', 'Spoken', 'in', 'Switzerland', ':', 'entire', 'German-  
speaking', 'part']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',

'Switzerland', ':', 'entire', 'German-speaking', 'part']]], 8, True)

['Alemannic', 'Spoken', 'in', 'Germany', ':', 'most', 'of', 'Baden-W',  
'\xc3', '\xbc', 'rttemberg', 'and', 'Bavarian', 'Swabia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Germany', ':', 'most', 'of', 'Baden-W', '\xc3', '\xbc', 'rttemberg',  
'and', 'Bavarian', 'Swabia']]], 14, True)

['Alemannic', 'Spoken', 'in', 'Austria', ':', 'Vorarlberg', 'and', 'some',  
'parts', 'of', 'Tyrol']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Austria', ':', 'Vorarlberg', 'and', 'some', 'parts', 'of', 'Tyrol']]], 11,  
True)

['Alemannic', 'Spoken', 'in', 'Liechtenstein', ':', 'entire', 'country']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Liechtenstein', ':', 'entire', 'country']]], 7, True)

['Alemannic', 'Spoken', 'in', 'France', ':', 'Alsace']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'France',

['Alemannic', 'Spoken', 'in', 'Italy', ':', 'some', 'parts', 'of', 'Aosta',  
'Valley', 'and', 'northern', 'Piedmont']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy',  
'Piedmont']]], 13, True)

['Alemannic', 'Spoken', 'in', 'Venezuela', ':', 'Alem', '\xc3', '\xa1',  
'n', 'Coloniero']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Venezuela', ':', 'Alem', '\xc3', '\xa1', 'n', 'Coloniero']]], 10, True)

['Alemannic', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Alemannic', 'Pronunciation', 'al', '\xc9', '\x9b', '\xcb', '\x88', 'man',  
'(', '\xcb', '\x90', ')', '\xc9', '\xaa', '\xca', '\x83']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'al',  
'\xc9', '\x9b', '\xcb', '\x88', 'man', '\\(', '\xcb', '\x90', '\\)',  
'\xc9', '\xaa', '\xca', '\x83']]], 16, True)

['Alemannic', 'Language', 'codes', 'ISO', '639-2', 'gsw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Alemannic', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Alemannic', 'Language', 'codes', 'ISO', '639-3', ':', 'gct', '\xe2',
'\x80', '\x93', 'Alem', '\xc3', '\xa1', 'n', 'Coloniero']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'gct', '\xe2', '\x80', '\x93', 'Alem', '\xc3', '\xa1',
'n', 'Coloniero']], ['NUM', '639-3']], 15, True)
```

```
['Brithenig', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Brithenig', 'Date', 'founded', '1996']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'alternative', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'Brithenig']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY']]], 6, True)
```

```
['Brithenig', 'Created', 'by', 'Andrew', 'Smith']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Brithenig', 'Category', '(', 'sources', ')', 'constructed', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)'], 'Type: ANY', 'Type: ANY']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'sources', ')', 'a', 'posteriori',
'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
```



'sources', '\\)'), 'a', 'posteriori', 'language']]], 8, True)

['Brithenig', 'Category', '(', 'sources', ')', '(', 'Romance', 'language', 'based', 'on', 'Celtic', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)'), '\\(', 'Romance', 'language', 'based', 'on', 'Celtic', '\\\n)']]], 12, True)

['Brithenig', 'Setting', 'and', 'usage', 'A', 'thought', 'experiment', 'in', 'alternate', 'history']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'A', 'thought', 'experiment', 'in', 'alternate', 'history']]], 10, True)

['Brithenig', 'Setting', 'and', 'usage', 'Ill', 'Bethisad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Ill', 'Bethisad']]], 6, True)

['Brithenig', 'Language', 'codes', 'ISO', '639-3', 'bzt']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zimakani', 'Total', 'speakers', '1500']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zimakani', 'Language', 'codes', 'ISO', '639-3', 'zik']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Palatine', 'Spoken', 'in', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Palatine', 'Total', 'speakers', 'unknown']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Palatine', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Palatine', 'Language', 'codes', 'ISO', '639-3', 'pfl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Old', 'Spoken', 'in', 'southern', 'Germany', '(', 'south', 'of', 'the',  
'Benrath', 'line', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'southern', 'Germany', '\\(', 'south', 'of', 'the', 'Benrath', 'line', '\\  
)']]], 12, True)
```

```
['Old', 'Spoken', 'in', 'parts', 'of', 'Austria', 'and', 'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',  
'of', 'Austria', 'and', 'Switzerland']]], 8, True)
```

```
['Old', 'Spoken', 'in', 'Southern', 'Bohemia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'goh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'goh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'Roman', 'Republic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Old', 'Region', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Old', 'Official', 'language', 'in', 'Rome']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Old', 'Regulated', 'by', 'Schools', 'of', 'grammar', 'and', 'rhetoric']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Schools', 'of', 'grammar', 'and', 'rhetoric']]], 8, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'lat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Old', 'Language', 'codes', 'ISO', '639-3', 'Expansion', 'of', 'the', 'Roman', 'Republic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Expansion', 'of', 'the', 'Roman', 'Republic']], ['NUM', '639-3']], 10, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'and', 'was', 'by', 'no', 'means', 'ubiquitous', 'within', 'it']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'and', 'was', 'by', 'no', 'means', 'ubiquitous', 'within', 'it']], ['NUM', '639-3']], 13, True)

['Old', 'Language', 'codes', 'ISO', '639-1', 'la']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Xavante', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xavante', 'Region', 'Mato', 'Grosso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xavante', 'Language', 'codes', 'ISO', '639-3', 'xav']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mogholi', 'Total', 'speakers', '200']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mogholi', 'Spoken', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mogholi', 'Region', 'near', 'Herat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mogholi', 'Language', 'codes', 'ISO', '639-3', 'mhj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marwari', 'Total', 'speakers', '13', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'NUM', 'million']], ['NUM', '13']], 5, True)

['Marwari', 'Region', 'Rajasthan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Marwari', 'Region', 'Gujarat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Marwari', 'Region', 'Haryana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Marwari', 'Region', 'Sindh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Marwari', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marwari', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marwari', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marwari', 'Writing', 'system', 'Devanagari', 'script', 'and', 'Mahajani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devanagari', 'script', 'and', 'Mahajani']], 7, True)

['Marwari', 'Language', 'codes', 'ISO', '639-2', 'mwr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'mwr', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mwr', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', ':', 'dhd', '\xe2', '\x80', '\x93', 'Dhundari']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'dhd', '\xe2', '\x80', '\x93', 'Dhundari']], ['NUM', '639-3']], 11, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'wry', '\xe2', '\x80', '\x93', 'Merwari']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wry', '\xe2', '\x80', '\x93', 'Merwari']], ['NUM', '639-3']], 10, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'mtr', '\xe2', '\x80', '\x93', 'Mewari']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mtr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'smv', '\xe2', '\x80', '\x93', 'Shekhawati']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'smv', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Mozarabic', 'Spoken', 'in', 'Iberia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mozarabic', 'Language', 'codes', 'ISO', '639-3', 'mxi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hattic', 'Spoken', 'in', 'Anatolia', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Hattic', 'Spoken', 'in', 'Hattians']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hattic', 'Language', 'codes', 'ISO', '639-3', 'xht']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Somali', 'Spoken', 'in', 'Somalia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Somaliland', '(', 'not', 'recognized',
'internationally', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Somaliland', '\\(', 'not', 'recognized', 'internationally', '\\)']]], 9,
True)
```

```
['Somali', 'Spoken', 'in', 'Djibouti']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Ethiopia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Yemen']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Kenya']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Total', 'speakers', '13-25', 'million', 'native', 'speakers',
'and', 'maybe', '500', ',', '000', 'second', 'language', 'speakers']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'native', 'speakers', 'and', 'maybe', 'NUM', ',', 'NUM',
'second', 'language', 'speakers']], ['NUM', '13-25'], ['NUM', '500'],
['NUM', '000']], 15, True)
```

```
['Somali', 'Writing', 'system', 'Latin']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Writing', 'system', 'Arabic']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Somali', 'Writing', 'system', 'Osmanya', 'script']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Somali', 'Writing', 'system']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system']]],  
3, True)
```

```
['Somali', 'Official', 'language', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Somali', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Somali', 'Language', 'codes', 'ISO', '639-2', 'som']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Somali', 'Language', 'codes', 'ISO', '639-3', 'som']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Somali', 'Language', 'codes', 'ISO', '639-1', 'so']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Gutnish', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gutnish', 'Region', 'Gotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Gutnish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)
```

```
['Gutnish', 'Language', 'codes', 'ISO', '639-3', 'non-gut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['East', 'Spoken', 'in', 'Germany', ',', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['East', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['East', 'Language', 'codes', 'ISO', '639-2', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['East', 'Language', 'codes', 'ISO', '639-3', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Venda', 'Region', 'Limpopo', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Venda', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Venda', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Venda', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Venda', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Venda', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Venda', 'Language', 'codes', 'ISO', '639-2', 'ven']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Venda', 'Language', 'codes', 'ISO', '639-3', 'ven']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Venda', 'Language', 'codes', 'ISO', '639-1', 've']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bemba', 'Spoken', 'in', 'Zambia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bemba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Bemba', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bemba', 'Official', 'language', 'in', 'Zambia', '(', 'regional', 'language', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Zambia', '\\\\(', 'regional', 'language', '\\\\)']], 9, True)

['Bemba', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Bemba', 'Language', 'codes', 'ISO', '639-2', 'bem']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bemba', 'Language', 'codes', 'ISO', '639-3', 'bem']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hopi', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Hopi', 'Region', 'Northeastern', 'Arizona']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hopi', 'Language', 'codes', 'ISO', '639-3', 'hop']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adele', 'Spoken', 'in', 'Ghana', 'and', 'Togo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Adele', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Adele', 'Language', 'codes', 'ISO', '639-3', 'ade']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaonde', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kaonde', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Kaonde', 'Language', 'codes', 'ISO', '639-3', 'kqn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zizilivakan', 'Spoken', 'in', 'Cameroon', ',', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Zizilivakan', 'Language', 'codes', 'ISO', '639-3', 'ziz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lozi', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lozi', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lozi', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lozi', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lozi', 'Spoken', 'in', 'South', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lozi', 'Language', 'codes', 'ISO', '639-2', 'loz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lozi', 'Language', 'codes', 'ISO', '639-3', 'loz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ekajuk', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ekajuk', 'Language', 'codes', 'ISO', '639-3', 'eka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Region', 'Mainly', 'in', 'the', 'Punjab', 'area', 'of', 'Pakistan', '.', 'Also', 'spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mainly', 'in', 'the', 'Punjab', 'area', 'of', 'Pakistan', '.', 'Also', 'spoken', 'in', 'Afghanistan']]], 14, True)

['Western', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Writing', 'system', 'Shahmukhi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'lah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'pnb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lunda', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lunda', 'Spoken', 'in', 'Angola']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lunda', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Lunda', 'Region', 'Copperbelt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lunda', 'Language', 'codes', 'ISO', '639-2', 'lun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lunda', 'Language', 'codes', 'ISO', '639-3', 'lun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Masaba', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Masaba', 'Region', 'Eastern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Masaba', 'Region', 'south', 'of', 'the', 'Kupsabiny']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'south', 'of', 'the', 'Kupsabiny']]], 6, True)

['Masaba', 'Region', 'Bugisu', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Masaba', 'Language', 'codes', 'ISO', '639-3', 'myx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yorkshire', 'Spoken', 'in', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yorkshire', 'Region', 'Yorkshire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yorkshire', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Mauritian', 'Spoken', 'in', 'Mauritius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mauritian', 'Language', 'codes', 'ISO', '639-3', 'mfe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mauritian', 'Language', 'codes', 'Linguasphere', '51-AAC-cec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Tahitian', 'Spoken', 'in', 'French', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tahitian', 'Language', 'codes', 'ISO', '639-2', 'tah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tahitian', 'Language', 'codes', 'ISO', '639-3', 'tah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tahitian', 'Language', 'codes', 'ISO', '639-1', 'ty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Palauan', 'Spoken', 'in', 'Palau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palauan', 'Spoken', 'in', 'Guam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palauan', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Japanese',  
'katakana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

```
['Palauan', 'Official', 'language', 'in', 'Palau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Palauan', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Palauan', 'Language', 'codes', 'ISO', '639-2', 'pau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Palauan', 'Language', 'codes', 'ISO', '639-3', 'pau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Aosta', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Piedmont']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Foggia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Franche-Comt', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Savoie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Bresse']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Bugey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Franco-Proven\xc3\xa7al,', 'Region', 'Dombes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY'[]], 3, True)

['Franco-Proven\xc3\xa7al,', 'Region', 'Beaujolais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Franco-Proven\xc3\xa7al,', 'Region', 'Dauphin', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 5, True)

['Franco-Proven\xc3\xa7al,', 'Region', 'Lyonnais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Franco-Proven\xc3\xa7al,', 'Region', 'Forez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Franco-Proven\xc3\xa7al,', 'Region', 'Romandy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Franco-Proven\xc3\xa7al,', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Franco-Proven\xc3\xa7al,', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Franco-Proven\xc3\xa7al,', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Franco-Proven\xc3\xa7al,', 'Writing', 'system', 'Latin', 'alphabet', 'with', 'diacritical', 'marks']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY'[]]], 8, True)

['Franco-Proven\xc3\xa7al,', 'Pronunciation', 'patu', '\xcb', '\x88', 'e', ',', 'patu', '\xcb', '\x88', '\xc9', '\x91']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'patu', '\xcb', '\x88', 'e', ',', 'patu', '\xcb', '\x88', '\xc9', '\x91'[]]], 12, True)

['Franco-Proven\xc3\xa7al,', 'Official', 'language', 'in', 'protected',

'by', 'statute', 'in', 'Italy', 'and', 'Aosta', 'Valley', 'Autonomous', 'Region']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'protected', 'by', 'statute', 'in', 'Italy', 'and', 'Aosta', 'Valley', 'Autonomous', 'Region']]], 14, True)

['Franco-Proven\xc3\xa7al,', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Franco-Proven\xc3\xa7al,', 'Language', 'codes', 'ISO', '639-3', 'frp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Franco-Proven\xc3\xa7al,', 'Language', 'codes', 'Linguasphere', '51-AAA-j']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Franco-Proven\xc3\xa7al,', 'Language', 'codes', 'Linguasphere', ':', 'Dark', 'Blue', ':', 'Protected', '.', '\xe2', '\x80', '\x94', 'Medium', 'Blue', ':', 'General', 'regions']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', ':', 'Dark', 'Blue', ':', 'Protected', '.', '\xe2', '\x80', '\x94', 'Medium', 'Blue', ':', 'General', 'regions']]], 18, True)

['Franco-Proven\xc3\xa7al,', 'Language', 'codes', 'Linguasphere', 'Light', 'Blue', ':', 'Historical', 'transition', 'zone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Light', 'Blue', ':', 'Historical', 'transition', 'zone']]], 10, True)

['Sicilian', 'Spoken', 'in', 'Sicily']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicilian', 'Spoken', 'in', 'Apulia', '(', 'Salento', ',', 'Lecce', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Apulia', '\\(', 'Salento', ',', 'Lecce', '\\)']]], 9, True)

['Sicilian', 'Spoken', 'in', 'Campania', '(', 'Cilento', ',', 'Campania', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Campania', '\\(', 'Cilento', ',', 'Campania', '\\)']]], 9, True)



['Sicilian', 'Spoken', 'in', 'Canada', '(', 'Vancouver', ',', 'Qu', '\xc3', '\xa9', 'bec', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada', '\\(', 'Vancouver', ',', 'Qu', '\xc3', '\xa9', 'bec', '\\)']], 12, True)

['Sicilian', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sicilian', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sicilian', 'Spoken', 'in', 'Argentina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sicilian', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sicilian', 'Language', 'codes', 'ISO', '639-2', 'scn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sicilian', 'Language', 'codes', 'ISO', '639-3', 'scn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Comorian', 'Spoken', 'in', 'Comoros', 'and', 'Mayotte']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Comorian', 'Region', 'Throughout', 'Comoros', 'and', 'Mayotte', ';', 'also', 'in', 'Madagascar', 'and', 'R', '\xc3', '\xa9', 'union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Throughout', 'Comoros', 'and', 'Mayotte', ';', 'also', 'in', 'Madagascar', 'and', 'R', '\xc3', '\xa9', 'union']], 15, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', ':', 'zdz', '\xe2', '\x80', '\x93', 'Ngazidja', 'dialect']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```
'ISO', 'NUM', ':', 'zdj', '\xe2', '\x80', '\x93', 'Ngazidja', 'dialect']],  
['NUM', '639-3']], 12, True)
```

```
['Comorian', 'Language', 'codes', 'ISO', '639-3', 'wni', '\xe2', '\x80',  
\x93', 'Ndzwani', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'wni', '\xe2', '\x80', '\x93', 'Ndzwani', 'dialect']],  
['NUM', '639-3']], 11, True)
```

```
['Comorian', 'Language', 'codes', 'ISO', '639-3', 'swb', '\xe2', '\x80',  
\x93', 'Maore', 'Comorian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'swb', '\xe2', '\x80', '\x93', 'Maore', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Comorian', 'Language', 'codes', 'ISO', '639-3', 'wlc', '\xe2', '\x80',  
\x93', 'Mwali', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'wlc', '\xe2', '\x80', '\x93', 'Mwali', 'dialect']], ['NUM',  
'639-3']], 11, True)
```

```
['Maithili', 'Region', 'Bihar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Maithili', 'Region', 'Jharkhand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Maithili', 'Region', 'Terai', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Maithili', 'Spoken', 'in', 'India', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Maithili', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Maithili', 'Writing', 'system', 'Kaithi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Maithili', 'Writing', 'system', 'Mithilakshar']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maithili', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Maithili', 'Language', 'codes', 'ISO', '639-2', 'mai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Maithili', 'Language', 'codes', 'ISO', '639-3', 'mai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Maithili', 'Language', 'codes', 'ISO', '639-1', 'bh', '(', 'Bihari', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bh', '\\(', 'Bihari', '\\)']]], ['NUM', '639-1']], 9, True)

['Dakhni', 'Spoken', 'in', 'South', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dakhni', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Dakhni', 'Writing', 'system', 'Urdu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dakhni', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Dakhni', 'Language', 'codes', 'ISO', '639-3', 'hin-dak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Haitian', 'Spoken', 'in', 'Haiti', '(', 'official', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Haitian', 'Spoken', 'in', 'Bahamas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'Cuba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'Dominican', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Haitian', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Haitian', 'Official', 'language', 'in', 'Haiti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Haitian', 'Language', 'codes', 'ISO', '639-2', 'hat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Haitian', 'Language', 'codes', 'ISO', '639-3', 'hat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Haitian', 'Language', 'codes', 'Linguasphere', '51-AAC-cb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Haitian', 'Language', 'codes', 'ISO', '639-1', 'ht']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Old', 'Spoken', 'in', 'Eastern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'orv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bhojpuri', 'Dialects', 'Caribbean', 'Hindustani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bhojpuri', 'Dialects', 'Sarnami']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bhojpuri', 'Dialects', 'Mauritian', 'Bhojpuri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bhojpuri', 'Region', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bhojpuri', 'Region', 'Bihar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bhojpuri', 'Region', 'Uttar', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bhojpuri', 'Region', 'Jharkhand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bhojpuri', 'Spoken', 'in', 'moribund', 'in', 'Guyana', 'and', 'Trinidad',  
'and', 'Tobago']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'moribund', 'in', 'Guyana', 'and', 'Trinidad', 'and', 'Tobago']]], 10,  
True)

['Bhojpuri', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bhojpuri', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Bhojpuri', 'Spoken', 'in', 'Mauritius']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Spoken', 'in', 'Suriname']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Writing', 'system', 'Devanagari', ',', 'Kaithi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bhojpuri', 'Language', 'codes', 'ISO', '639-2', 'bho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Bhojpuri', 'Language', 'codes', 'ISO', '639-3', 'bho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bhojpuri', 'Language', 'codes', 'Linguasphere', '59-AAF-sa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Bhojpuri', 'Language', 'codes', 'ISO', '639-1', 'bh', '(', 'Bihari', 'languages', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bh', '\\\\(', 'Bihari', 'languages', '\\\\)']]], ['NUM', '639-1']], 10, True)

['Cajun', 'Spoken', 'in', 'Louisiana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Cajun', 'Language', 'codes', 'ISO', '639-3', 'frc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cajun', 'Language', 'codes', 'Linguasphere', '51-AAA-hp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Picard', 'Spoken', 'in', 'France']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Picard', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Picard', 'Official', 'language', 'in', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Picard', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Picard', 'Language', 'codes', 'ISO', '639-3', 'pcd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Picard', 'Language', 'codes', 'Linguasphere', '51-AAA-he']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Crimean', 'Spoken', 'in', 'formerly', 'Crimea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Crimean', 'Writing', 'system', 'originally']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Writing', 'system', 'probably', 'Runic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Crimean', 'Writing', 'system', 'later', 'Gothic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Crimean', 'Writing', 'system', 'Greek']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Writing', 'system', 'and', 'Latin', 'alphabets']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

```

['Crimean', 'Language', 'codes', 'ISO', '639-2', 'got']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Crimean', 'Language', 'codes', 'ISO', '639-3', 'got', '\xe2', '\x80',
'\x93', 'Gothic']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'got', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Selonian,', 'Spoken', 'in', 'Latvia', 'and', 'Lithuania']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']], 6, True)

['Selonian,', 'Region', 'Europe']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Selonian,', 'Language', 'codes', 'ISO', '639-3', 'sxl']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Zhou', 'Dynasty']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Old', 'Spoken', 'in', 'Warring', 'States', 'Period']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Warring', 'States', 'Period']], 6, True)

['Old', 'Region', 'Ancient', 'China']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Old', 'Writing', 'system', 'Oracle', 'Bone', 'Script']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Old', 'Writing', 'system', 'Bronze', 'script']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

['Old', 'Writing', 'system', 'Seal', 'Script']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

```



```

['Old', 'Language', 'codes', 'ISO', '639-3', 'och']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thracian', 'Language', 'codes', 'ISO', '639-3', 'txh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chewa,', 'Spoken', 'in', 'Zambia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Chewa,', 'Spoken', 'in', 'Malawi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Chewa,', 'Spoken', 'in', 'Mozambique']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Chewa,', 'Spoken', 'in', 'Zimbabwe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Chewa,', 'Official', 'language', 'in', 'Malawi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Chewa,', 'Official', 'language', 'in', 'Zambia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Chewa,', 'Regulated', 'by', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Chewa,', 'Language', 'codes', 'ISO', '639-2', 'nya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chewa,', 'Language', 'codes', 'ISO', '639-3', 'nya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chewa,', 'Language', 'codes', 'ISO', '639-1', 'ny']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Slavey', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Slavey', 'Region', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Slavey', 'Official', 'language', 'in', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Slavey', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Slavey', 'Language', 'codes', 'ISO', '639-2', 'den']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Slavey', 'Language', 'codes', 'ISO', '639-3', 'den', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'den', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Slavey', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Wu', 'Dialects', 'Oujiang', '(', 'Wenzhou', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Wu', 'Dialects', 'Taihu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Wu', 'Dialects', 'Taizhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Wu', 'Dialects', 'Wuzhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

ANY']]], 3, True)

['Wu', 'Dialects', 'Chuju']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wu', 'Dialects', 'Xuanzhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wu', 'Region', 'Shanghai', ';', 'most', 'of', 'Zhejiang', 'province',  
;', 'southern', 'Jiangsu', 'province', ';', 'Xuancheng', 'prefecture-  
level', 'city', 'of', 'Anhui', 'province', ';', 'Shangrao', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Shanghai',  
;', 'most', 'of', 'Zhejiang', 'province', ';', 'southern', 'Jiangsu',  
'province', ';', 'Xuancheng', 'prefecture-level', 'city', 'of', 'Anhui',  
'province', ';', 'Shangrao', 'County']]], 22, True)

['Wu', 'Region', 'Guangfeng', 'County', 'and', 'Yushan', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Guangfeng',  
'County', 'and', 'Yushan', 'County']]], 7, True)

['Wu', 'Region', 'Jiangxi', 'province', ';', 'Pucheng', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jiangxi',  
'province', ';', 'Pucheng', 'County']]], 7, True)

['Wu', 'Region', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wu', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Wu', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Wu', 'Language', 'codes', 'ISO', '639-3', 'wuu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wu', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tuvan', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tuvan', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tuvan', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tuvan', 'Region', 'Tuva']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tuvan', 'Official', 'language', 'in', 'Tuva', '(', 'federal', 'subject', 'of', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Tuva', '\\(', 'federal', 'subject', 'of', 'Russia', '\\)']]], 11, True)

['Tuvan', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tuvan', 'Language', 'codes', 'ISO', '639-2', 'tyv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tuvan', 'Language', 'codes', 'ISO', '639-3', 'tyv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Church', 'Spoken', 'in', 'Eastern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Church', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Church', 'Writing', 'system', 'Glagolitic', 'alphabet', '(', 'Glag', ')', ', ', 'Cyrillic', 'alphabet', '(', 'Cyr', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Glagolitic', 'alphabet', '\\(', 'Glag', '\\)', ', ', 'Cyrillic', 'alphabet', '\\(', 'Cyr', '\\)', ', ', 'Cyrillic', 'alphabet', '\\(', 'Cyr', '\\)']]], 11, True)

'alphabet', '\\(', 'Cyr', '\\)']]], 14, True)

['Church', 'Language', 'codes', 'ISO', '639-2', 'chu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Church', 'Language', 'codes', 'ISO', '639-3', 'chu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Church', 'Language', 'codes', 'Linguasphere', '53-AAA-a']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Church', 'Language', 'codes', 'ISO', '639-1', 'cu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Southern', 'Spoken', 'in', 'Norway', ',', 'Sweden']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Southern', 'Total', 'speakers', '600']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Southern', 'Official', 'language', 'in', 'Sn', '\\xc3', '\\xa5', 'sa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Sn', '\\xc3', '\\xa5', 'sa']], 8, True)

['Southern', 'Official', 'language', 'in', '(', 'minority', 'language',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', '\\(', 'minority', 'language', '\\)']], 8, True)

['Southern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'sma']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```

['Southern', 'Language', 'codes', 'ISO', '639-3', 'sma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auvergnat', 'Spoken', 'in', 'France']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Auvergnat', 'Region', 'Auvergne']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Auvergnat', 'Region', 'Puy-de-D', '\xc3', '\xb4', 'me']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Puy-de-D',
'\xc3', '\xb4', 'me']]], 6, True)

['Auvergnat', 'Region', 'Haute-Loire']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Auvergnat', 'Region', 'Allier']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Auvergnat', 'Region', 'Cantal']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Auvergnat', 'Region', 'Communities', 'in', 'Limousin']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Auvergnat', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Auvergnat', 'Language', 'codes', 'ISO', '639-3', 'oci-auv']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auvergnat', 'Language', 'codes', 'Linguasphere', '51-AAA-gi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

```

['Soqotri', 'Spoken', 'in', 'Yemen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Soqotri', 'Region', 'Socotra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Soqotri', 'Language', 'codes', 'ISO', '639-3', 'sqt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['L\xc3\xa1adan', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['L\xc3\xa1adan', 'Date', 'founded', '1982']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'artistic', 'and', 'philosophical', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'artistic', 'and', 'philosophical', 'language']], 9, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'fictional', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'L', '\xc3', '\xa1', 'adan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'L', '\xc3', '\xa1', 'adan']], 9, True)

['L\xc3\xa1adan', 'Created', 'by', 'Suzette', 'Haden', 'Elgin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Suzette', 'Haden', 'Elgin']], 6, True)

['L\xc3\xa1adan', 'Language', 'codes', 'ISO', '639-3', 'ldn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piedmontese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Piedmontese', 'Region', 'northwest', 'Italy', ',', 'Piedmont']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northwest',  
'Italy', ',', 'Piedmont']], 6, True)

['Piedmontese', 'Language', 'codes', 'ISO', '639-3', 'pms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piedmontese', 'Language', 'codes', 'Linguasphere', '51-AAA-of']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Suzhou', 'dialect', 'Total', 'speakers', 'approx', '.', '5-7',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '5-7']], 8, True)

['Suzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Suzhou', 'Language', 'codes', 'ISO', '639-3', 'wu-suh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['K\xc3\xb6lsch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['K\xc3\xb6lsch', 'Region', 'Cologne', 'and', 'environs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['K\xc3\xb6lsch', 'Language', 'codes', 'ISO', '639-3', 'ksh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Low', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Low', 'Spoken', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Low', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Low', 'Total', 'speakers', '\xe2', '\x89', '\x88', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\xe2', '\x89', '\x88', 'NUM', 'million']], ['NUM', '5']], 8, True)

['Low', 'Language', 'codes', 'ISO', '639-2', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Low', 'Language', 'codes', 'Linguasphere', '52-ACB']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Wyandot', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']], 7, True)

['Wyandot', 'Region', 'northeastern', 'Oklahoma', ',', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northeastern', 'Oklahoma', ',', 'Type: ANY']], 6, True)

['Wyandot', 'Writing', 'system', 'modified', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Wyandot', 'Language', 'codes', 'ISO', '639-3', 'wya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Elamite', 'Region', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Elamite', 'Language', 'codes', 'ISO', '639-2', 'elx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Elamite', 'Language', 'codes', 'ISO', '639-3', 'elx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chiac,', 'Shiac', ',', 'Shiak', ',', 'Chiak', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Shiac', ',', 'Shiak', ',', 'Chiak', 'Total', 'speakers', 'Unknown']], 9, True)

['Chiac,', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chiac,', 'Region', 'mainly', 'around', 'Moncton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Chiac,', 'Region', 'Shediac', 'and', 'Memramcook']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-2', 'fre', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fre', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-2', 'fra', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fra', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-3', 'fra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chiac,', 'Language', 'codes', 'Linguasphere', '51-AAA-am']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-1', 'fr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Isan', 'Total', 'speakers', '20', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'NUM', 'million']], ['NUM', '20']], 5, True)

['Isan', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Isan', 'Region', 'Isan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Isan', 'Language', 'codes', 'ISO', '639-3', 'tts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norfolk', 'Spoken', 'in', 'Norfolk', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Norfolk', 'Spoken', 'in', 'Pitcairn', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Norfolk', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Norfolk', 'Pronunciation', 'n', '\xc9', '\x94', '\xcb', '\x90', 'fuk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\x94', '\xcb', '\x90', 'fuk']], 8, True)

['Norfolk', 'Official', 'language', 'in', 'Norfolk', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Norfolk', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Norfolk', 'Language', 'codes', 'ISO', '639-3', 'pih']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bundeli', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Bundeli', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bundeli', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Bundeli', 'Official', 'language', 'in', 'India', '(', 'Madhya',  
'Pradesh', 'and', 'parts', 'of', 'Uttar', 'Pradesh', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'India', '\\(', 'Madhya', 'Pradesh', 'and', 'parts', 'of', 'Uttar',  
'Pradesh', '\\)']]], 14, True)
```

```
['Bundeli', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Bundeli', 'Language', 'codes', 'ISO', '639-3', 'bns']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bagheli', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bagheli', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bagheli', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Bagheli', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Bagheli', 'Language', 'codes', 'ISO', '639-3', 'bfy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ewe', 'Total', 'speakers', 'approximately', '6', 'million', ',', 'with',  
'2', ',', '000', ',', '000', 'second', 'language', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'approximately', 'NUM', 'million', ',', 'with', 'NUM', ',', 'NUM', ',',
```

```
['NUM', 'second', 'language', 'speakers']], ['NUM', '6'], ['NUM', '2'],
['NUM', '000'], ['NUM', '000']], 16, True)
```

```
['Ewe', 'Spoken', 'in', 'Ghana', ',', 'Togo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]]], 6, True)
```

```
['Ewe', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Ewe', 'Language', 'codes', 'ISO', '639-2', 'ewe']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Ewe', 'Language', 'codes', 'ISO', '639-3', 'ewe']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Ewe', 'Language', 'codes', 'ISO', '639-1', 'ee']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Khwarezmian', 'Spoken', 'in', 'Khwarezm']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['Khwarezmian', 'Writing', 'system', 'Aramaic', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Khwarezmian', 'Writing', 'system', 'Sogdian', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Khwarezmian', 'Writing', 'system', 'Pahlavi', 'script']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Khwarezmian', 'Language', 'codes', 'ISO', '639-3', 'xco']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Hawaii', 'Spoken', 'in', 'US']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```

['Hawaii', 'Language', 'codes', 'ISO', '639-3', 'hwc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hawaii', 'Language', 'codes', 'Linguasphere', '52-ABB-dc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Torres', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Torres', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)

['Torres', 'Language', 'codes', 'ISO', '639-2', 'cpe', '(', '?', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'cpe', '\\(', '?', '\\)']], ['NUM', '639-2']], 9, True)

['Torres', 'Language', 'codes', 'ISO', '639-3', 'tcs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Korandje', 'Total', 'speakers', '3000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Korandje', 'Spoken', 'in', 'Algeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Korandje', 'Region', 'Tabelbala', ',', 'wilaya', 'of', 'Bechar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
',', 'Type: ANY', 'of', 'Type: ANY']]], 7, True)

['Korandje', 'Language', 'codes', 'ISO', '639-3', 'kcy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zarma', 'Spoken', 'in', 'southwest', 'Niger']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Zarma', 'Language', 'codes', 'ISO', '639-3', 'dje']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zyrian', 'Region', 'Komi', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zyrian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zyrian', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Zyrian', 'Official', 'language', 'in', 'Komi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Zyrian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Zyrian', 'Language', 'codes', 'ISO', '639-2', 'kom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zyrian', 'Language', 'codes', 'ISO', '639-3', 'kpv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zyrian', 'Language', 'codes', 'ISO', '639-1', 'kv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Saraiki', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saraiki', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saraiki', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saraiki', 'Writing', 'system', 'Perso-Arabic', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Saraiki', 'Writing', 'system', 'Gurmukhi', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Saraiki', 'Writing', 'system', 'Devanagari', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Saraiki', 'Language', 'codes', 'ISO', '639-3', 'skr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Shanghainese', 'Total', 'speakers', '14', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']]], ['NUM', '14']], 5, True)

['Shanghainese', 'Pronunciation', 'z', '\xcc', '\xa5', '\xc9', '\x91',  
'\xcc', '\x83', '\xcc', '\x80', 'h', '\xc3', '\xa9', '\xc9', '\xa6',  
'\xc9', '\x9b', '\xcc', '\x80', '\xc9', '\xa6', '\xca', '\x8a', '\xcc',  
'\x80', ',', '\xc9', '\xa6', 'u', '\xc9', '\xb2', 'y']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'z',  
'\xcc', '\xa5', '\xc9', '\x91', '\xcc', '\x83', '\xcc', '\x80', 'h',  
'\xc3', '\xa9', '\xc9', '\xa6', '\xc9', '\x9b', '\xcc', '\x80', '\xc9',  
'\xa6', '\xca', '\x8a', '\xcc', '\x80', ',', '\xc9', '\xa6', 'u', '\xc9',  
'\xb2', 'y']]], 33, True)

['Shanghainese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)

['Shanghainese', 'Language', 'codes', 'ISO', '639-3', 'wuu-sha']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lakota', 'Region', 'but', 'also', 'northern', 'Nebraska']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'but', 'also',  
'northern', 'Nebraska']]], 6, True)

['Lakota', 'Region', 'southern', 'Minnesota']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lakota', 'Region', 'and', 'northern', 'Montana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lakota', 'Spoken', 'in', 'United', 'States', ',', 'with', 'some', 'speakers', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', ',', 'with', 'some', 'speakers', 'in', 'Canada']]], 11, True)

['Lakota', 'Pronunciation', 'la', '\xcb', '\x88', 'k', '\xcb', '\xa3', 'otijapi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'la', '\xcb', '\x88', 'k', '\xcb', '\xa3', 'otijapi']]], 9, True)

['Lakota', 'Language', 'codes', 'ISO', '639-3', 'lkt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rumsen', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Rumsen', 'Total', 'speakers', 'extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rumsen', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Rumsen', 'Language', 'codes', 'ISO', '639-2', 'nai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Rumsen', 'Language', 'codes', 'ISO', '639-3', 'css']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rumsen', 'Language', 'codes', 'ISO', '639-1', '-']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Komi-Permyak', 'Spoken', 'in', 'Russia']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Komi-Permyak', 'Region', 'Komi-Permyak', 'Okrug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Komi-Permyak', 'Language', 'codes', 'ISO', '639-2', 'kom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Komi-Permyak', 'Language', 'codes', 'ISO', '639-3', 'koi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Komi-Permyak', 'Language', 'codes', 'ISO', '639-1', 'kv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Kildin', 'Spoken', 'in', 'Kola', 'Peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kildin', 'Spoken', 'in', 'Murmansk', 'Oblast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kildin', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Kildin', 'Total', 'speakers', '753', 'total', ',', '608', 'in', 'Murmansk', 'Oblast', 'only']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'total', ',', 'NUM', 'in', 'Murmansk', 'Oblast', 'only']]], ['NUM', '753'], ['NUM', '608']], 11, True)
```

```
['Kildin', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Kildin', 'Language', 'codes', 'ISO', '639-3', 'sjd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Inari', 'Spoken', 'in', 'Finland']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Inari', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Inari', 'Language', 'codes', 'ISO', '639-2', 'smn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Erzya', 'Total', 'speakers', 'Ethnologue']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Estonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Kyrgyzstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Turkmenistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Erzya', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Region', 'Mordovia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Nizhny', 'Novgorod']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Region', 'Chuvashia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Ulyanovsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Samara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Penza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Saratov']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Orenburg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Tatarstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Bashkortostan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Language', 'codes', 'ISO', '639-2', 'myv']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Erzya', 'Language', 'codes', 'ISO', '639-3', 'myv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Evenki', 'Region', 'Inner', 'Mongolia', 'and', 'Heilongjiang', 'in', 'China', ';', 'Selenge', 'Province', 'in', 'Mongolia', ';', 'Krasnoyarsk', 'Krai', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Inner', 'Mongolia', 'and', 'Heilongjiang', 'in', 'China', ';', 'Selenge', 'Province', 'in', 'Mongolia', ';', 'Krasnoyarsk', 'Krai', 'in', 'Russia']]], 18, True)
```

```
['Evenki', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Evenki', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Evenki', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Evenki', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Evenki', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Evenki', 'Writing', 'system', 'Mongolian', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Evenki', 'Language', 'codes', 'ISO', '639-3', 'evn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shoshoni', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

['Shoshoni', 'Region', 'Wyoming']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Shoshoni', 'Region', 'Utah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Shoshoni', 'Region', 'Nevada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Shoshoni', 'Region', 'Idaho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Shoshoni', 'Language', 'codes', 'ISO', '639-3', 'shh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balti', 'Spoken', 'in', 'Baltistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balti', 'Spoken', 'in', 'Ladakh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balti', 'Language', 'codes', 'ISO', '639-3', 'bft']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalmyk', 'Region', 'Kalmykia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kalmyk', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalmyk', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalmyk', 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalmyk', 'Writing', 'system', 'Clear', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kalmyk', 'Official', 'language', 'in', 'Kalmykia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Kalmyk', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kalmyk', 'Language', 'codes', 'ISO', '639-2', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kalmyk', 'Language', 'codes', 'ISO', '639-3', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Beja', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Beja', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Beja', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Beja', 'Language', 'codes', 'ISO', '639-2', 'bej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Beja', 'Language', 'codes', 'ISO', '639-3', 'bej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quiripi', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Quiripi', 'Total', 'speakers', 'extinct']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Quiripi', 'Language', 'codes', 'ISO', '639-3', 'qyp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kabardian', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Spoken', 'in', 'Turkey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Spoken', 'in', 'Jordan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Spoken', 'in', 'Iraq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Spoken', 'in', 'Syria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Spoken', 'in', 'Germany']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Spoken', 'in', 'USA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kabardian', 'Region', 'Kabardino-Balkaria', ',', 'Karachay-Cherkessia',
'and', 'north', 'east', 'Turkey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kabardino-
Balkaria', ',', 'Karachay-Cherkessia', 'and', 'north', 'east', 'Turkey']]],
9, True)
```

```
['Kabardian', 'Language', 'codes', 'ISO', '639-2', 'kbd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```



['Kabardian', 'Language', 'codes', 'ISO', '639-3', 'kbd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Blackfoot', 'Total', 'speakers', '5', ',', '100', '/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '/']], ['NUM', '5'], ['NUM', '100']], 7, True)

['Blackfoot', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Blackfoot', 'Region', 'Blackfeet', 'Indian', 'Reservation', 'in',  
'Montana', 'and', 'Piikani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Blackfeet',  
'Indian', 'Reservation', 'in', 'Montana', 'and', 'Piikani']]], 9, True)

['Blackfoot', 'Region', 'Siksika']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Blackfoot', 'Region', 'and', 'Kainai', 'Reserves', 'in', 'southern',  
'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Kainai', 'Reserves', 'in', 'southern', 'Alberta']]], 8, True)

['Blackfoot', 'Language', 'codes', 'ISO', '639-2', 'bla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Blackfoot', 'Language', 'codes', 'ISO', '639-3', 'bla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kongo', 'Total', 'speakers', '7', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '7']], 5, True)

['Kongo', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kongo', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Kongo', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'the', 'Congo']]], 7, True)

['Kongo', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kongo', 'Official', 'language', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Kongo', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of',  
'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 9, True)

['Kongo', 'Official', 'language', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Kongo', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Kongo', 'Language', 'codes', 'ISO', '639-2', 'kon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', 'kon', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kon', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', ':', 'kng', '\xe2', '\x80',  
'\x93', 'Koongo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'kng', '\xe2', '\x80', '\x93', 'Koongo']], ['NUM',  
'639-3']], 11, True)

```
['Kongo', 'Language', 'codes', 'ISO', '639-3', 'ldi', '\xe2', '\x80',  
'\x93', 'Laari']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ldi', '\xe2', '\x80', '\x93', 'Laari']], ['NUM', '639-3']],  
10, True)
```

```
['Kongo', 'Language', 'codes', 'ISO', '639-1', 'kg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Lombard', 'Spoken', 'in', 'Italy', ',', 'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Lombard', 'Region', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lombard', 'Region', ':', 'Lombardy', ',', 'Piedmont']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', ':',  
'Lombardy', ',', 'Piedmont']], 6, True)
```

```
['Lombard', 'Region', 'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lombard', 'Region', ':', 'Ticino']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lombard', 'Region', 'Graub', '\xc3', '\xbc', 'nden']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Graub',  
'\xc3', '\xbc', 'nden']], 6, True)
```

```
['Lombard', 'Region', 'Val', 'Poschiavo', ',', 'Val', 'Mesolcina', 'and',  
'Val', 'Bregaglia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Val',  
'Poschiavo', ',', 'Val', 'Mesolcina', 'and', 'Val', 'Bregaglia']], 10,  
True)
```

```
['Lombard', 'Language', 'codes', 'ISO', '639-3', 'lmo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Moabite', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern',  
'Jordan']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern', 'Jordan']]], 8, True)

['Moabite', 'Language', 'codes', 'ISO', '639-3', 'obm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["O'odham", 'Spoken', 'in', 'United', 'States', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

["O'odham", 'Region', 'Primarily', 'south-central', 'Arizona', 'and', 'northern', 'Sonora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Primarily', 'south-central', 'Arizona', 'and', 'northern', 'Sonora']]], 8, True)

["O'odham", 'Official', 'language', 'in', 'One', 'of', 'the', 'national', 'languages', 'of', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'One', 'of', 'the', 'national', 'languages', 'of', 'Mexico']]], 11, True)

["O'odham", 'Regulated', 'by', 'Secretary', 'of', 'Public', 'Education', 'in', 'Mexico', ';', 'various', 'tribal', 'agencies', 'in', 'the', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Secretary', 'of', 'Public', 'Education', 'in', 'Mexico', ';', 'various', 'tribal', 'agencies', 'in', 'the', 'USA']]], 16, True)

["O'odham", 'Language', 'codes', 'ISO', '639-3', 'ood']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Phrygian', 'Language', 'codes', 'ISO', '639-3', 'xpg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adyghe', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Adyghe', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']]], 6, True)

['Adyghe', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Adyghe', 'Region', 'Russia', ':', 'Republic', 'of', 'Adygea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Russia', ':', 'Republic', 'of', 'Adygea']]], 7, True)

['Adyghe', 'Official', 'language', 'in', 'Republic', 'of', 'Adygea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'Adygea']]], 7, True)

['Adyghe', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Adyghe', 'Language', 'codes', 'ISO', '639-2', 'ady']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Adyghe', 'Language', 'codes', 'ISO', '639-3', 'ady']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['H\x3\xa4n', 'Total', 'speakers', '10']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['H\xc3\xa4n', 'Region', 'Yukon', ',', 'Alaska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['H\xc3\xa4n', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']]], 7, True)

['H\xc3\xa4n', 'Language', 'codes', 'ISO', '639-3', 'haa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ojibwe', 'Region', 'Canada', ':', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ojibwe', 'Region', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ojibwe', 'Region', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ojibwe', 'Region', 'Saskatchewan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ojibwe', 'Region', 'groups', 'in', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ojibwe', 'Region', 'Wisconsin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ojibwe', 'Region', 'Minnesota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ojibwe', 'Region', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ojibwe', 'Spoken', 'in', 'Canada', ',', ]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', ', ']]], 5, True)

['Ojibwe', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Ojibwe', 'Pronunciation', 'an', '\xc9', '\xaa', '\xca', '\x83', '\xc9',  
'\xaa', 'na', '\xcb', '\x90', 'pe', '\xcb', '\x90', 'mow', '\xc9', '\xaa',  
'n']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'an',  
'\xc9', '\xaa', '\xca', '\x83', '\xc9', '\xaa', 'na', '\xcb', '\x90', 'pe',  
'\xcb', '\x90', 'mow', '\xc9', '\xaa', 'n']]]], 19, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-2', 'oji']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'oji', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'oji', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM',  
'639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ciw', '\xe2', '\x80',  
'\x93', 'Chippewa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ciw', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'otw', '\xe2', '\x80',  
'\x93', 'Ottawa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'otw', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'alq', '\xe2', '\x80',  
'\x93', 'Algonquin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'alq', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-1', 'oj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Venetian', 'Region', 'The', 'Adriatic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Venetian', 'Language', 'codes', 'ISO', '639-3', 'vec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Venetian', 'Language', 'codes', 'Linguasphere', '51-AAA-n']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Gen', 'Spoken', 'in', 'Togo', ',', 'Benin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Gen', 'Language', 'codes', 'ISO', '639-3', 'gej']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Stadsfries', 'Spoken', 'in', 'Netherlands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Stadsfries', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Stadsfries', 'Pronunciation', '\xcb', '\x88', 'st', '\xc9', '\x91', 'tsfris']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'st', '\xc9', '\x91', 'tsfris']], 8, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-2', 'dut', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'dut', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-2', 'nld', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nld', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-3', 'nld']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-1', 'nl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['West', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['West', 'Region', 'Friesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['West', 'Region', 'Groningen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['West', 'Official', 'language', 'in', 'Province', 'of', 'Friesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Province', 'of', 'Friesland']]], 7, True)

['West', 'Regulated', 'by', 'Fryske', 'Akademy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['West', 'Language', 'codes', 'ISO', '639-2', 'fry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'fry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['West', 'Language', 'codes', 'Linguasphere', '52-ACA-b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['West', 'Language', 'codes', 'Linguasphere', 'Map', 'showing', 'the',  
'Frisian-speaking', 'areas', 'in', 'the', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Map', 'showing', 'the', 'Frisian-speaking', 'areas', 'in',  
'the', 'Netherlands']]], 12, True)

['West', 'Language', 'codes', 'ISO', '639-1', 'fy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Saterland', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saterland', 'Region', 'Lower', 'Saxony']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saterland', 'Language', 'codes', 'ISO', '639-2', 'frs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Saterland', 'Language', 'codes', 'ISO', '639-3', 'stq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['North', 'Region', 'Schleswig-Holstein']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['North', 'Official', 'language', 'in', 'Nordfriesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['North', 'Official', 'language', 'in', 'Heligoland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['North', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-2', 'frr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'frr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Classical', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-3', 'arb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Classical', 'Language', 'codes', 'ISO', '639-1', 'ar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Sireniki', 'Writing', 'system', 'Transcribed', 'with', 'Cyrillic', 'in',  
'old', 'monographs', '(', 'extended', 'with', 'diacritics', ')', ',',  
'but', 'new', 'publications', 'may', 'appear', 'also', 'romanized']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Transcribed', 'with', 'Cyrillic', 'in', 'old', 'monographs', '\\(',  
'extended', 'with', 'diacritics', '\\)', ',', 'but', 'new', 'publications',  
'may', 'appear', 'also', 'romanized']]], 22, True)
```

```
['Sireniki', 'Language', 'codes', 'ISO', '639-3', 'ysr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Limburgish', 'Total', 'speakers', '1', ',', '600', ',', '000', '(',  
'est', '.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'est', '.', '\\)']], ['NUM', '1'],  
['NUM', '600'], ['NUM', '000']], 12, True)
```

```
['Limburgish', 'Region', 'Limburg', '(', 'Netherlands', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Limburg', '\\(  
, 'Netherlands', '\\)']]], 6, True)
```

```
['Limburgish', 'Region', 'Limburg', '(', 'Belgium', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Limburg', '\\(  
, 'Belgium', '\\)']]], 6, True)
```

```
['Limburgish', 'Spoken', 'in', 'Netherlands', '(', 'Limburg', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)
```

```
['Limburgish', 'Spoken', 'in', 'Germany', '(', 'adjacent', 'parts', 'of',  
'the', 'Rhineland', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Germany', '\\(', 'adjacent', 'parts', 'of', 'the', 'Rhineland', '\\)']]],  
11, True)
```

```
['Limburgish', 'Writing', 'system', 'Latin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Limburgish', 'Pronunciation', '\\xcb', '\\x88', 'l', '\\xc9', '\\x9b', 'mb',  
'\\xc5', '\\x93', '\\xca', '\\x81', '\\xca', '\\x9d', 's']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb',  
'\\x88', 'l', '\\xc9', '\\x9b', 'mb', '\\xc5', '\\x93', '\\xca', '\\x81', '\\xca',  
'\\x9d', 's']]], 15, True)
```

```
['Limburgish', 'Official', 'language', 'in', 'the', 'Netherlands', '(',  
'as', 'a', 'regional', 'language', ')', ';', 'no', 'official', 'status',  
'in', 'Belgium', 'or', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'the', 'Netherlands', '\\(', 'as', 'a', 'regional', 'language', '\\  
\\)', ';', 'no', 'official', 'status', 'in', 'Belgium', 'or', 'Germany']]],  
20, True)
```

```
['Limburgish', 'Language', 'codes', 'ISO', '639-2', 'lim']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Limburgish', 'Language', 'codes', 'ISO', '639-3', 'lim']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Limburgish', 'Language', 'codes', 'ISO', '639-1', 'li']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Soga', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Soga', 'Region', 'Mainly', 'in', 'Busoga', 'region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mainly', 'in',  
'Busoga', 'region']]], 6, True)
```

```
['Soga', 'Language', 'codes', 'ISO', '639-3', 'xog']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pangasinense', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pangasinense', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']], 7, True)

['Pangasinense', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Pangasinense', 'Language', 'codes', 'ISO', '639-2', 'pag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pangasinense', 'Language', 'codes', 'ISO', '639-3', 'pag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ottoman', 'Spoken', 'in', 'Ottoman', 'Empire-Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ottoman', 'Writing', 'system', 'Ottoman', 'Turkish', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Ottoman', 'Language', 'codes', 'ISO', '639-2', 'ota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ottoman', 'Language', 'codes', 'ISO', '639-3', 'ota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Finnish', 'Signed', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bulgar', 'Language', 'codes', 'ISO', '639-3', 'xbo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hassaniyya', 'Spoken', 'in', 'south-western', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Spoken', 'in', 'Libya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hassaniyya', 'Spoken', 'in', 'northern', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Spoken', 'in', 'Mauritania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hassaniyya', 'Spoken', 'in', 'southern', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Spoken', 'in', 'north-western', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Spoken', 'in', 'northern', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Spoken', 'in', 'Western', 'Sahara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Hassaniyya', 'Language', 'codes', 'ISO', '639-3', 'mey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Limousin', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Limousin', 'Total', 'speakers', 'no', 'reliable', 'data']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'no', 'reliable', 'data']]], 6, True)
```

```
['Limousin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Limousin', 'Language', 'codes', 'ISO', '639-3', 'oci-lim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Limousin', 'Language', 'codes', 'Linguasphere', '51-AAA-gj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Meitei', 'Spoken', 'in', 'Northeast', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Meitei', 'Spoken', 'in', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Meitei', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Meitei', 'Writing', 'system', 'Eastern', 'Nagari', '(', 'currently',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Meitei', 'Writing', 'system', 'Meitei', 'Mayek', '(', 'formerly', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Meitei', 'Official', 'language', 'in', 'India', '(', 'Manipur', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'India', '\\(', 'Type: ANY', '\\)']]], 8, True)
```

```
['Meitei', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

['Meitei', 'Language', 'codes', 'ISO', '639-2', 'mni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Meitei', 'Language', 'codes', 'ISO', '639-3', 'mni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sylheti', 'Spoken', 'in', 'Including', 'the', 'Bangladeshi', 'diaspora',  
'communities', 'in', 'Britain', 'and', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Including', 'the', 'Bangladeshi', 'diaspora', 'communities', 'in',  
'Britain', 'and', 'the', 'United', 'States']]], 14, True)

['Sylheti', 'Spoken', 'in', 'Primarily', 'spoken', 'in', 'the', 'Sylhet',  
'region', 'of', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Primarily', 'spoken', 'in', 'the', 'Sylhet', 'region', 'of',  
'Bangladesh']]], 11, True)

['Sylheti', 'Spoken', 'in', 'and', 'the', 'Indian', 'districts', 'of', ':',  
'Barak', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',  
'the', 'Indian', 'districts', 'of', ':', 'Barak', 'Valley']]], 11, True)

['Sylheti', 'Spoken', 'in', 'Silchar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sylheti', 'Spoken', 'in', 'Hailakandi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sylheti', 'Spoken', 'in', 'Karimganj', 'and', 'Tripura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Sylheti', 'Writing', 'system', 'Sylheti', 'Nagari', ',', 'Bengali',  
'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Sylheti', 'Language', 'codes', 'ISO', '639-3', 'syl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Kemi', 'Spoken', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kemi', 'Language', 'codes', 'ISO', '639-3', 'sjk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papi\xc3\xa1', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Papi\xc3\xa1', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Papi\xc3\xa1', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Papi\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'mcm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papi\xc3\xa1', 'Language', 'codes', 'Linguasphere', '51-AAC-aha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Niuean', 'Spoken', 'in', 'Niue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Niuean', 'Spoken', 'in', 'Cook', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Niuean', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Niuean', 'Spoken', 'in', 'Tonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Niuean', 'Language', 'codes', 'ISO', '639-2', 'niu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Niuean', 'Language', 'codes', 'ISO', '639-3', 'niu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Mi'kmaq", 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']], 7, True)

["Mi'kmaq", 'Region', 'Nova', 'Scotia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Mi'kmaq", 'Region', 'New', 'Brunswick']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Mi'kmaq", 'Region', 'Newfoundland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

["Mi'kmaq", 'Region', 'northern', 'Maine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Mi'kmaq", 'Region', 'Boston', '(', 'Massachusetts', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Boston', '\\\\(', 'Massachusetts', '\\\)']], 6, True)

["Mi'kmaq", 'Language', 'codes', 'ISO', '639-2', 'mic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

["Mi'kmaq", 'Language', 'codes', 'ISO', '639-3', 'mic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mandaic', 'Total', 'speakers', '5500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ammonite', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern', 'Jordan']], 8, True)

['Ammonite', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Ammonite', 'Language', 'codes', 'ISO', '639-3', 'agg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Meroitic', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Meroitic', 'Region', 'Mero', '\xc3', '\xab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Meroitic', 'Writing', 'system', 'Meroitic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Meroitic', 'Language', 'codes', 'ISO', '639-3', 'xmr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khasi', 'Spoken', 'in', 'India', ',', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Khasi', 'Language', 'codes', 'ISO', '639-2', 'kha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Khasi', 'Language', 'codes', 'ISO', '639-3', 'kha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Santhali', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Santhali', 'Spoken', 'in', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Santhali', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Santhali', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Santhali', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Ol', 'Chiki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Santhali', 'Language', 'codes', 'ISO', '639-2', 'sat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Santhali', 'Language', 'codes', 'ISO', '639-3', 'sat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Galician\xe2\x80\x93Portuguese', 'Spoken', 'in', 'Kingdom', 'of',  
'Galicia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kingdom', 'of', 'Galicia']]], 6, True)

['Galician\xe2\x80\x93Portuguese', 'Region', 'North-west', 'Iberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Galician\xe2\x80\x93Portuguese', 'Language', 'codes', 'ISO', '639-3',  
'\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Galician\xe2\x80\x93Portuguese', 'Language', 'codes', 'ISO', '639-3',  
'079']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Carolinian', 'Region', 'Saipan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Carolinian', 'Region', 'Anatahan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Carolinian', 'Region', 'and', 'Agrihan', 'islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Carolinian', 'Region', 'Carolines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Carolinian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Carolinian', 'Language', 'codes', 'ISO', '639-3', 'cal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dene', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dene', 'Region', 'Northern', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Dene', 'Region', 'Saskatchewan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dene', 'Region', 'Manitoba', ';', 'southern', 'Northwest', 'Territories', 'and', 'Nunavut']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Manitoba', ';', 'southern', 'Northwest', 'Territories', 'and', 'Nunavut']]], 9, True)

['Dene', 'Language', 'codes', 'ISO', '639-2', 'chp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dene', 'Language', 'codes', 'ISO', '639-3', 'chp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Forro', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']]], 15, True)

['Forro', 'Language', 'codes', 'ISO', '639-2', 'cpp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Forro', 'Language', 'codes', 'ISO', '639-3', 'cri']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Forro', 'Language', 'codes', 'Linguasphere', '51-AAC-aca']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Kristi', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kristi', 'Total', 'speakers', '750']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kristi', 'Language', 'codes', 'ISO', '639-2', 'cpp']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Kristi', 'Language', 'codes', 'ISO', '639-3', 'vvp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kristi', 'Language', 'codes', 'Linguasphere', '51-AAC-agc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Angolar', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']], 15, True)
```

```
['Angolar', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Angolar', 'Language', 'codes', 'ISO', '639-3', 'aoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Angolar', 'Language', 'codes', 'Linguasphere', '51-AAC-ad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Principense', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']], 15, True)
```

```
['Principense', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Principense', 'Language', 'codes', 'ISO', '639-3', 'pre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Principense', 'Language', 'codes', 'Linguasphere', '51-AAC-acb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)
```

['Macanese', 'Spoken', 'in', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Macanese', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', '\\\\(', 'Type: ANY', '\\\\)']], 8, True)

['Macanese', 'Spoken', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Macanese', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Macanese', 'Language', 'codes', 'ISO', '639-3', 'mzs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Macanese', 'Language', 'codes', 'Linguasphere', '51-AAC-ai']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Gwich\u0099in', 'Region', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gwich\u0099in', 'Region', 'Yukon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gwich\u0099in', 'Region', 'Alaska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gwich\u0099in', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Gwich\u0099in', 'Official', 'language', 'in', 'Northwest', 'Territories', '(', 'Canada', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Northwest', 'Territories', '\\(', 'Canada', '\\)']]], 9, True)

['Gwich\u0099in', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Gwich\u0099in', 'Language', 'codes', 'ISO', '639-2', 'gwi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gwich\u0099in', 'Language', 'codes', 'ISO', '639-3', 'gwi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Penang', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Penang', 'Region', 'Penang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Penang', 'Region', 'Kedah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

ANY']]], 3, True)

['Penang', 'Region', 'Perak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Penang', 'Region', 'Medan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Penang', 'Region', 'Phuket']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Penang', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Penang', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Penang', 'Language', 'codes', 'ISO', '639-3', 'nan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Penang', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Amorite', 'Spoken', 'in', 'ancient', 'Mesopotamia', ',', 'by', 'the', 'Amorites']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'ancient', 'Mesopotamia', ',', 'by', 'the', 'Amorites']]], 9, True)

['Amorite', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dogrib', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dogrib', 'Region', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dogrib', 'Official', 'language', 'in', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Dogrib', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Dogrib', 'Language', 'codes', 'ISO', '639-2', 'dgr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dogrib', 'Language', 'codes', 'ISO', '639-3', 'dgr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dura', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dura', 'Spoken', 'in', 'East', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Dura', 'Language', 'codes', 'ISO', '639-3', 'drq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Daman', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Daman', 'Total', 'speakers', '4000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Daman', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Daman', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Daman', 'Language', 'codes', 'Linguasphere', '51-AAC-agb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Leponitic', 'Spoken', 'in', 'Cisalpine', 'Gaul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Leponitic', 'Language', 'codes', 'ISO', '639-3', 'xlp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Messapian', 'Spoken', 'in', 'Puglia', 'region', 'of', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Puglia', 'region', 'of', 'Italy']]], 7, True)

['Messapian', 'Language', 'codes', 'ISO', '639-3', 'cms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hurrian', 'Spoken', 'in', 'Mitanni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hurrian', 'Region', 'Mesopotamia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hurrian', 'Language', 'codes', 'ISO', '639-3', 'xhu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ongota', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ongota', 'Language', 'codes', 'ISO', '639-3', 'bxh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Menominee', 'Total', 'speakers', '130']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Menominee', 'Region', 'Northeastern', 'Wisconsin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Menominee', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Menominee', 'Pronunciation', 'om', '\xc3', '\xa6', '\xcb', '\x90',  
\xca', '\x94', 'nomenew']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'om',  
\xc3', '\xa6', '\xcb', '\x90', '\xca', '\x94', 'nomenew']]], 10, True)
```

```
['Menominee', 'Official', 'language', 'in', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Menominee', 'Language', 'codes', 'ISO', '639-2', 'alg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Menominee', 'Language', 'codes', 'ISO', '639-3', 'mez']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Urantian', 'Spoken', 'in', 'Urartu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Urantian', 'Language', 'codes', 'ISO', '639-3', 'xur']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Inuvialuktun', 'Spoken', 'in', 'Canada', '(', 'Northwest', 'Territories',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',  
'\\(', 'Northwest', 'Territories', '\\)']]], 8, True)
```

```
['Inuvialuktun', 'Region', 'North', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Inuvialuktun', 'Official', 'language', 'in', 'Northwest', 'Territories',  
'(', 'Canada', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Northwest', 'Territories', '\\(', 'Canada', '\\)']]], 9, True)
```

```
['Inuvialuktun', 'Language', 'codes', 'ISO', '639-2', 'iku']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inuvialuktun', 'Language', 'codes', 'ISO', '639-3', 'ikt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inuvialuktun', 'Language', 'codes', 'ISO', '639-1', 'iu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Inuinnaqtun', 'Spoken', 'in', 'Canada', '(', 'Nunavut', 'and',  
'Northwest', 'Territories', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',  
'\\(', 'Nunavut', 'and', 'Northwest', 'Territories', '\\)']]], 10, True)

['Inuinnaqtun', 'Official', 'language', 'in', 'Nunavut', 'and',  
'Northwest', 'Territories', '(', 'Canada', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Nunavut', 'and', 'Northwest', 'Territories', '\\(', 'Canada', '\\  
)']]], 11, True)

['Inuinnaqtun', 'Language', 'codes', 'ISO', '639-2', 'iku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inuinnaqtun', 'Language', 'codes', 'ISO', '639-3', 'ikt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inuinnaqtun', 'Language', 'codes', 'ISO', '639-1', 'iu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Palestinian', 'Spoken', 'in', 'Palestine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palestinian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palestinian', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palestinian', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Palestinian', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Palestinian', 'Language', 'codes', 'ISO', '639-3', 'ajp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Semigallian', 'Spoken', 'in', 'Latvia', 'and', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Semigallian', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Semigallian', 'Language', 'codes', 'ISO', '639-3', 'xzm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Extremaduran', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Extremaduran', 'Region', 'Autonomous', 'community', 'of', 'Extremadura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Extremaduran', 'Language', 'codes', 'ISO', '639-3', 'ext']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lezgian', 'Region', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Region', 'western', 'Caspian', 'Sea', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western', 'Caspian', 'Sea', 'coast']]], 6, True)

['Lezgian', 'Region', 'central', 'Caucasus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Spoken', 'in', 'Russia', 'and', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Lezgian', 'Spoken', 'in', 'also', 'spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'also', 'spoken', 'in', 'Georgia']]], 7, True)

['Lezgian', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Spoken', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lezgian', 'Spoken', 'in', 'and', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lezgian', 'Pronunciation', 'lez', '\xc9', '\xa1', 'i', 't', '\xca', '\x83', '\xca', '\xbc', 'al']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'lez', '\xc9', '\xa1', 'i', 't', '\xca', '\x83', '\xca', '\xbc', 'al']]], 12, True)

['Lezgian', 'Language', 'codes', 'ISO', '639-2', 'lez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lezgian', 'Language', 'codes', 'ISO', '639-3', 'lez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



```
['Lushootseed', 'Total', 'speakers', '200']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Lushootseed', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY', 'S']]], 7, True)
```

```
['Lushootseed', 'Language', 'codes', 'ISO', '639-3', 'lut']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Punic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Punic', 'Language', 'codes', 'ISO', '639-3', 'xpu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
["N'Ko", 'Spoken', 'in', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
["N'Ko", 'Spoken', 'in', 'Mali']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
["N'Ko", 'Region', 'West', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
["N'Ko", 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
["N'Ko", 'Language', 'codes', 'ISO', '639-2', 'nqo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
["N'Ko", 'Language', 'codes', 'ISO', '639-3', 'nqo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Fang', 'Spoken', 'in', 'Cameroon']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fang', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']]], 7, True)

['Fang', 'Spoken', 'in', 'Equatorial', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Fang', 'Spoken', 'in', 'Gabon', 'and', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Gabon', 'and', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']]], 17, True)

['Fang', 'Language', 'codes', 'ISO', '639-2', 'fan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Fang', 'Language', 'codes', 'ISO', '639-3', 'fan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Swiss', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swiss', 'Spoken', 'in', 'northeast', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Swiss', 'Spoken', 'in', 'northern', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Swiss', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Swiss', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Uripiv', 'Total', 'speakers', '9', ',', '000', 'for', 'dialect',

```

'continuum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'for', 'dialect', 'continuum']], ['NUM', '9'], ['NUM',
'000']], 9, True)

['Uripiv', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Uripiv', 'Region', 'Malakula']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Uripiv', 'Language', 'codes', 'ISO', '639-2', 'map']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Uripiv', 'Language', 'codes', 'ISO', '639-3', 'upv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laz', 'Spoken', 'in', 'Turkey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Laz', 'Spoken', 'in', 'Georgia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Laz', 'Spoken', 'in', 'Germany']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Laz', 'Total', 'speakers', '33', ',', '250', 'in', '1980']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '33'], ['NUM', '250']], 8,
True)

['Laz', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Georgian',
'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Laz', 'Language', 'codes', 'ISO', '639-3', 'lzz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Mingrelian', 'Region', 'Samegrelo', ',', 'Abkhazia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Mingrelian', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mingrelian', 'Writing', 'system', 'Georgian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mingrelian', 'Language', 'codes', 'ISO', '639-3', 'xmf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Moksha', 'Spoken', 'in', 'Russia', 'and', 'as', 'a', 'minority',  
'language', 'also', 'in', 'Armenia', ',', 'Australia', 'and', 'the',  
'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Russia',  
'and', 'as', 'a', 'minority', 'language', 'also', 'in', 'Armenia', ',',  
'Australia', 'and', 'the', 'United', 'States']]], 18, True)

['Moksha', 'Region', 'European', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Moksha', 'Official', 'language', 'in', 'Mordovia', '(', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Mordovia', '\\\\(', 'Russia', '\\\\)']]], 8, True)

['Moksha', 'Regulated', 'by', 'Mordovian', 'Research', 'Institute', 'of',  
'Language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Mordovian', 'Research', 'Institute', 'of', 'Type: ANY']]], 8, True)

['Moksha', 'Regulated', 'by', 'Literature']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Moksha', 'Regulated', 'by', 'History', 'and', 'Economics']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'History', 'and', 'Economics']]], 6, True)

['Moksha', 'Language', 'codes', 'ISO', '639-2', 'mdf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Moksha', 'Language', 'codes', 'ISO', '639-3', 'mdf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ecclesiastical', 'Writing', 'system', 'Spoken', 'and', 'written', 'Lingua', 'Franca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Ecclesiastical', 'Writing', 'system', 'Liturgical', 'and', 'administrative', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Liturgical', 'and', 'administrative', 'language']], 7, True)

['Ecclesiastical', 'Writing', 'system', 'Latin', 'alphabet', 'Literature']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Ecclesiastical', 'Writing', 'system', 'documents']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ecclesiastical', 'Writing', 'system', 'inscriptions']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Christian', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']], 2, True)

['Christian', 'Official', 'language', 'in', 'Holy', 'See']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Christian', 'Regulated', 'by', 'Style', 'not', 'regulated']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Style', 'not', 'regulated']], 6, True)

['Christian', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Christian', 'Language', 'codes', 'ISO', '639-3', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Christian', 'Language', 'codes', 'ISO', '639-3', 'The', 'spread', 'of', 'Christianity', 'in', 'Europe', 'to', '600', 'AD', '\xe2', '\x80', '\x94', 'the', 'dark', 'pockets', 'represent', 'initial', 'enclaves']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'The', 'spread', 'of', 'Christianity', 'in', 'Europe', 'to', 'NUM', 'AD', '\xe2', '\x80', '\x94', 'the', 'dark', 'pockets', 'represent', 'initial', 'enclaves']], ['NUM', '639-3'], ['NUM', '600']], 23, True)

['Christian', 'Language', 'codes', 'ISO', '639-1', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Svan', 'Total', 'speakers', 'less', 'than', '30', ',', '000', ';']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'less', 'than', 'NUM', ',', 'NUM', ';']], ['NUM', '30'], ['NUM', '000']], 9, True)

['Svan', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Svan', 'Pronunciation', 'lu', '\xca', '\x83', 'nu', 'nin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'lu', '\xca', '\x83', 'nu', 'nin']], 7, True)

['Svan', 'Language', 'codes', 'ISO', '639-3', 'sva']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Me\xc3\xa4nkieli', 'Spoken', 'in', 'Sweden', ',', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Me\xc3\xa4nkieli', 'Region', 'Torne', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Me\xc3\xa4nkieli', 'Official', 'language', 'in', 'Recognized', 'minority', 'language', 'of', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Recognized', 'minority', 'language', 'of', 'Sweden']], 9, True)

['Me\xc3\xa4nkieli', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:

ANY', 'of', 'Type: ANY']]], 6, True)

['Me\xc3\xa4nkieli', 'Language', 'codes', 'ISO', '639-3', 'fit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Han'er", 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Han'er", 'Region', 'Medieval', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Han'er", 'Writing', 'system', 'Seal', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

["Han'er", 'Writing', 'system', 'Clerical', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

["Han'er", 'Writing', 'system', 'Kaishu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Han'er", 'Writing', 'system', 'Semi-cursive', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

["Han'er", 'Writing', 'system', 'Grass', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

["Han'er", 'Writing', 'system', 'Phagspa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Han'er", 'Writing', 'system', 'Hangul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Han'er", 'Official', 'language', 'in', 'Yuan', 'Dynasty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

```
["Han'er", 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
["Han'er", 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Arawak', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arawak', 'Spoken', 'in', 'French', 'Guiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Arawak', 'Spoken', 'in', 'Guyana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arawak', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arawak', 'Region', 'Guianas', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arawak', 'Region', 'Arawak', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arawak', 'Language', 'codes', 'ISO', '639-3', 'arw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Khariboli', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Khariboli', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Khariboli', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
```



'Type: ANY', 'Type: ANY']]], 5, True)

['Khariboli', 'Language', 'codes', 'ISO', '639-2', 'hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Khariboli', 'Language', 'codes', 'ISO', '639-3', 'hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khariboli', 'Language', 'codes', 'ISO', '639-1', 'hi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Philippine', 'Region', 'Zamboanga', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Philippine', 'Region', 'Zamboanga', 'del', 'Norte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Philippine', 'Region', 'Zamboanga', 'Sibugay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Philippine', 'Region', 'Zamboanga', 'del', 'Sur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Philippine', 'Region', 'Basilan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Philippine', 'Region', 'Cavite', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Philippine', 'Region', 'Ternate']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Philippine', 'Region', 'Cavite']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Philippine', 'Region', 'Cotabato']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Philippine', 'Region', 'Davao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Philippine', 'Region', 'Jolo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Philippine', 'Region', 'Tawi-Tawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Philippine', 'Region', 'Semporna', 'in', 'Sabah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Philippine', 'Region', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Philippine', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Philippine', 'Writing', 'system', 'Latin', 'and', 'Spanish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Philippine', 'Official', 'language', 'in', 'Recognised', 'as', 'a', 'regional', 'language', 'in', 'the', 'Philippines', ';', 'recognized', 'as', 'a', 'minority', 'language', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Recognised', 'as', 'a', 'regional', 'language', 'in', 'the', 'Philippines', ';', 'recognized', 'as', 'a', 'minority', 'language', 'in', 'Malaysia']]], 20, True)

['Philippine', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Philippine', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Philippine', 'Language', 'codes', 'ISO', '639-3', 'cbk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Philippine', 'Language', 'codes', 'Linguasphere', '51-AAC-ba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Massachusetts', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Massachusetts', 'Region', 'Southeast', 'Massachusetts', 'Ethnicity']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Massachusetts', 'Language', 'codes', 'ISO', '639-3', 'wam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Greenlandic', 'Spoken', 'in', 'Greenland', ',', 'Kingdom', 'of',  
'Denmark']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Greenland', ',', 'Kingdom', 'of', 'Denmark']]], 8, True)

['Greenlandic', 'Region', 'North', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Greenlandic', 'Official', 'language', 'in', 'Greenland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Greenlandic', 'Language', 'codes', 'ISO', '639-2', 'kal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Greenlandic', 'Language', 'codes', 'ISO', '639-3', 'kal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Greenlandic', 'Language', 'codes', 'ISO', '639-1', 'kl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Seychellois', 'Spoken', 'in', 'Seychelles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Seychellois', 'Official', 'language', 'in', 'Seychelles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Seychellois', 'Regulated', 'by', 'Lenstiti', 'Kreol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Seychellois', 'Language', 'codes', 'ISO', '639-3', 'crs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Riograndenser', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Riograndenser', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Riograndenser', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Riograndenser', 'Language', 'codes', 'ISO', '639-3', 'hrx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abenaki', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Abenaki', 'Region', 'Odanak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Abenaki', 'Region', 'Centre-du-Qu', '\xc3', '\xa9', 'bec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Centre-du-Qu', '\xc3', '\xa9', 'bec']], 6, True)

['Abenaki', 'Region', 'Quebec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abenaki', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Venetic', 'Spoken', 'in', 'Northeastern', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Venetic', 'Language', 'codes', 'ISO', '639-3', 'xve']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jin', 'Total', 'speakers', '45', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '45']], 5, True)

['Jin', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jin', 'Region', 'most', 'of', 'Shanxi', 'province', ';', 'central', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Hebei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'most', 'of', 'Shanxi', 'province', ';', 'central', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Hebei']]], 14, True)

['Jin', 'Region', 'Henan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Henan']]], 3, True)

['Jin', 'Region', 'Shaanxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Jin', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Jin', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Jin', 'Language', 'codes', 'ISO', '639-3', 'cjy']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jin', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Bouyei', 'Spoken', 'in', 'China', '(', 'Guizhou', ',', 'Yunnan', ',', 'and', 'Sichuan', 'Provinces', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'China', '\\(', 'Guizhou', ',', 'Yunnan', ',', 'and', 'Sichuan', 'Provinces', '\\)']], 13, True)
```

```
['Bouyei', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Bouyei', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)
```

```
['Bouyei', 'Language', 'codes', 'ISO', '639-3', 'pcc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Inuktitut', 'Total', 'speakers', '35', ',', '000', '(', 'approx', '.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', ',', 'NUM', '\\(', 'approx', '.', '\\)']], ['NUM', '35'], ['NUM', '000']], 10, True)
```

```
['Inuktitut', 'Writing', 'system', 'Inuktitut', 'syllabics']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Inuktitut', 'Official', 'language', 'in', 'Nunavut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)
```

```
['Inuktitut', 'Official', 'language', 'in', 'Nunavik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)
```

```
['Inuktitut', 'Official', 'language', 'in', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

['Inuktitut', 'Official', 'language', 'in', 'Nunatsiavut', '(', 'Canada', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Nunatsiavut', '\\(', 'Canada', '\\)']], 8, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-2', 'iku']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-3', 'iku', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'iku', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-3', 'individual', 'code']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-1', 'iu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Hiligaynon', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hiligaynon', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']], 7, True)

['Hiligaynon', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Hiligaynon', 'Language', 'codes', 'ISO', '639-2', 'hil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hiligaynon', 'Language', 'codes', 'ISO', '639-3', 'hil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Sedang', 'Spoken', 'in', 'Vietnam', 'and', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)
```

```
['Sedang', 'Region', 'Kontum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sedang', 'Region', 'Quang', 'Nam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sedang', 'Region', 'Quang', 'Ngai', '(', 'Vietnam', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Quang',  
'Ngai', '\\(', 'Vietnam', '\\)']]], 7, True)
```

```
['Sedang', 'Language', 'codes', 'ISO', '639-3', 'sed']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kapampangan', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kapampangan', 'Region', 'Central', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kapampangan', 'Writing', 'system', 'Historically', 'written', 'in',  
'Matuang', 'Kudlitan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Kapampangan', 'Official', 'language', 'in', 'Regional', 'language', 'in',  
'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)
```

```
['Kapampangan', 'Language', 'codes', 'ISO', '639-2', 'pam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Kapampangan', 'Language', 'codes', 'ISO', '639-3', 'pam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```



['Milanese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Milanese', 'Total', 'speakers', 'n', '/', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'n', '/', 'a']], 6, True)

['Milanese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Norman', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Norman', 'Dialects', 'J', '\xc3', '\xa8', 'rriais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'J', '\xc3', '\xa8', 'rriais']], 6, True)

['Norman', 'Dialects', 'Guern', '\xc3', '\xa9', 'siais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Guern', '\xc3', '\xa9', 'siais']], 6, True)

['Norman', 'Dialects', 'Sercquiais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Norman', 'Region', 'Normandy', 'and', 'the', 'Channel', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Normandy', 'and', 'the', 'Channel', 'Islands']], 7, True)

['Norman', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Norman', 'Spoken', 'in', 'Normandy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Norman', 'Spoken', 'in', 'Jersey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Norman', 'Spoken', 'in', 'Guernsey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Norman', 'Spoken', 'in', 'Alderney', '\xe2', '\x80', '\xa0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Alderney', '\xe2', '\x80', '\xa0']]], 7, True)
```

```
['Norman', 'Spoken', 'in', 'Herm', '\xe2', '\x80', '\xa0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Herm',  
'\xe2', '\x80', '\xa0']]], 7, True)
```

```
['Norman', 'Spoken', 'in', 'Sark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Norman', 'Spoken', 'in', 'Canada', '\xe2', '\x80', '\xa0', '(', 'used',  
'to', 'a', 'certain', 'degree', 'in', 'eastern', 'Canada', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',  
'\xe2', '\x80', '\xa0', '\\(', 'used', 'to', 'a', 'certain', 'degree',  
'in', 'eastern', 'Canada', '\\)']]], 17, True)
```

```
['Norman', 'Spoken', 'in', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Norman', 'Spoken', 'in', 'Acadia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Norman', 'Spoken', 'in', 'Kingdom', 'of', 'Sicily', '\xe2', '\x80',  
'\xa0', '(', 'used', 'in', 'a', 'limited', 'degree', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kingdom', 'of', 'Sicily', '\xe2', '\x80', '\xa0', '\\(', 'used', 'in',  
'a', 'limited', 'degree', '\\)']]], 16, True)
```

```
['Norman', 'Spoken', 'in', 'Principality', 'of', 'Antioch', '\xe2', '\x80',  
'\xa0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Principality', 'of', 'Antioch', '\xe2', '\x80', '\xa0']]], 9, True)
```

```
['Norman', 'Spoken', 'in', '\xe2', '\x80', '\xa0', 'now', 'defunct', 'in',  
'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2',  
'\x80', '\xa0', 'now', 'defunct', 'in', 'area']]], 10, True)
```

```
['Norman', 'Writing', 'system', 'Latin', '(', 'French', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Norman', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Norman', 'Language', 'codes', 'ISO', '639-3', 'fra-nor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Norman', 'Language', 'codes', 'Linguasphere', 'Guernsey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Norman', 'Language', 'codes', 'Linguasphere', 'the', 'Cotentin', 'and',  
'the', 'Pays', 'de', 'Caux']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'the', 'Cotentin', 'and', 'the', 'Pays', 'de', 'Caux']]],  
11, True)
```

```
['Gallo', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gallo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Gallo', 'Language', 'codes', 'ISO', '639-3', 'fra-gal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gallo', 'Language', 'codes', 'Linguasphere', '51-AAA-hb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Middle', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Middle', 'Language', 'codes', 'ISO', '639-2', 'frm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Middle', 'Language', 'codes', 'ISO', '639-3', 'frm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Judeo-Arabic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Judeo-Arabic', 'Spoken', 'in', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2',
'\x80', '\x93']]], 6, True)
```

```
['Judeo-Arabic', 'Region', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Judeo-Arabic', 'Writing', 'system', 'Hebrew', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-2', 'jrb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'jrb', '\xe2',
'\x80', '\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'jrb', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'individual',
'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Karaim', 'Spoken', 'in', 'Crimea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Karaim', 'Spoken', 'in', 'Lithuania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

['Karaim', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Karaim', 'Total', 'speakers', 'less', 'than', '50', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'less', 'than', 'NUM', 'in', 'Lithuania']], ['NUM', '50']], 8, True)

['Karaim', 'Language', 'codes', 'ISO', '639-3', 'kdr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zazaki', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zazaki', 'Region', 'Main', 'in', 'Tunceli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Zazaki', 'Region', 'Bingol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Zazaki', 'Region', 'Erzincan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Zazaki', 'Region', 'Sivas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Zazaki', 'Region', 'Elazig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Zazaki', 'Region', 'Malatya', 'G', '\xc3', '\xbc', 'm', '\xc3', '\xbc', '\xc5', '\x9f', 'hane', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Malatya', 'G', '\xc3', '\xbc', 'm', '\xc3', '\xbc', '\xc5', '\x9f', 'hane', 'Province']], 13, True)

['Zazaki', 'Region', '\xc5', '\x9e', 'anl', '\xc4', '\xb1', 'urfa', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', '\xc5', '\x9e',

'anl', '\xc4', '\xb1', 'urfa', 'Province']]], 9, True)

['Zazaki', 'Region', 'and', 'Ad', '\xc4', '\xb1', 'yaman', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'Ad',  
'\xc4', '\xb1', 'yaman', 'Province']]], 8, True)

['Zazaki', 'Region', 'diasporic', 'in', 'Mutki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Zazaki', 'Region', 'Sar', '\xc4', '\xb1', 'z']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sar', '\xc4',  
'\xb1', 'z']]], 6, True)

['Zazaki', 'Region', 'Aksaray']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Zazaki', 'Region', 'and', 'Taraz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Zazaki', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Zazaki', 'Language', 'codes', 'ISO', '639-2', 'zza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zazaki', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mycenaean', 'Spoken', 'in', 'southern', 'Balkans', '/', 'Crete']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'southern', 'Balkans', '/', 'Crete']]], 7, True)

['Mycenaean', 'Writing', 'system', 'Linear', 'B']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Mycenaean', 'Language', 'codes', 'ISO', '639-3', 'gmy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mycenaean', 'Language', 'codes', 'ISO', '639-3', 'and', 'therefore', 'can', 'be', 'used', 'as', 'an', 'estimator', 'of', 'the', 'range']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'and', 'therefore', 'can', 'be', 'used', 'as', 'an', 'estimator', 'of', 'the', 'range']], ['NUM', '639-3']], 16, True)

['Guern\x3\xa9siais', 'Spoken', 'in', 'Guernsey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Guern\x3\xa9siais', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Guern\x3\xa9siais', 'Language', 'codes', 'Linguasphere', '51-AAA-hc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Ingrian', 'Total', 'speakers', '500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ingrian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ingrian', 'Region', 'Ingria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ingrian', 'Language', 'codes', 'ISO', '639-3', 'izh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Elfdalian/\xc3\x96vdalian', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Elfdalian/\xc3\x96vdalian', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Elfdalian/\xc3\x96vdalian', 'Regulated', 'by', 'R', '\xc3', '\xa5', '\xc3', '\xb0', 'dj', '\xc3', '\xa4', 'rum', ',', '\xc3', '\xa4', 'lvdalska', 'spr', '\xc3', '\xa5', 'kr', '\xc3', '\xa5', 'det']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'R',  
'\xc3', '\xa5', '\xc3', '\xb0', 'dj', '\xc3', '\xa4', 'rum', ',', '\xc3',  
'\xa4', 'lvdalska', 'spr', '\xc3', '\xa5', 'kr', '\xc3', '\xa5', 'det']]],  
23, True)

['Elfdalian/\xc3\x96vdalian', 'Regulated', 'by', '(', 'semi-official',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', '\\(',  
'semi-official', '\\)']]], 6, True)

['Elfdalian/\xc3\x96vdalian', 'Language', 'codes', 'ISO', '639-3', 'qer',  
'\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qer', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Bats', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bats', 'Region', 'Zemo-Alvani', 'in', 'Kakheti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Bats', 'Language', 'codes', 'ISO', '639-3', 'bbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Andalusian', 'Spoken', 'in', 'Al-Andalus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Andalusian', 'Region', 'Iberian', 'Peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Andalusian', 'Language', 'codes', 'ISO', '639-3', 'xaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Egyptian', 'Spoken', 'in', 'Egypt', 'and', 'a', 'few', 'other',  
'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Egypt',  
'and', 'a', 'few', 'other', 'countries']]], 9, True)

['Egyptian', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',



'Type: ANY', 'Type: ANY']]], 5, True)

['Egyptian', 'Pronunciation', '\xcb', '\x88', 'm', '\xc9', '\x91', 's',  
\xcb', '\xa4', '\xc9', '\xbe', 'i']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
\x88', 'm', '\xc9', '\x91', 's', '\xcb', '\xa4', '\xc9', '\xbe', 'i']]]],  
13, True)

['Egyptian', 'Language', 'codes', 'ISO', '639-3', 'arz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yemeni', 'Spoken', 'in', 'Yemen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]]], 4, True)

['Yemeni', 'Spoken', 'in', 'Somalia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]]], 4, True)

['Yemeni', 'Spoken', 'in', 'Somaliland', 'de', 'facto', 'state', 'not',  
currently', 'recognized', 'internationally']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
Somaliland', 'de', 'facto', 'state', 'not', 'currently', 'recognized',  
internationally']]]], 11, True)

['Yemeni', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
\xe2', '\x80', '\x93']]]], 6, True)

['Yemeni', 'Writing', 'system', 'Arabic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
Type: ANY', 'Type: ANY']]]], 5, True)

['Yemeni', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Yemeni', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Washo', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

```
['Washo', 'Region', 'Nevada', '(', 'Washoe', 'County', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nevada', '\\
(', 'Washoe', 'County', '\\\')']], 7, True)
```

```
['Washo', 'Language', 'codes', 'ISO', '639-2', 'was']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Washo', 'Language', 'codes', 'ISO', '639-3', 'was']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Levantine', 'Spoken', 'in', 'Syria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Levantine', 'Spoken', 'in', 'Lebanon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Levantine', 'Spoken', 'in', 'Israel']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Levantine', 'Spoken', 'in', 'Palestine']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Levantine', 'Spoken', 'in', 'Jordan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Levantine', 'Writing', 'system', 'Arabic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Levantine', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Levantine', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['(Eastern)', 'Spoken', 'in', 'Botswana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['(Eastern)', 'Language', 'codes', 'ISO', '639-3', 'huc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x81Xegwi', 'Language', 'extinction', '1988']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['\xc7\x81Xegwi', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['\xc7\x81Xegwi', 'Language', 'codes', 'ISO', '639-3', 'xeg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x80Xam', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['\xc7\x80Xam', 'Language', 'codes', 'ISO', '639-3', 'xam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Seroa', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Seroa', 'Language', 'codes', 'ISO', '639-3', 'kqu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ju\xc7\x80\xca\xbchoan', 'Spoken', 'in', 'Namibia', ',', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Ju\xc7\x80\xca\xbchoan', 'Region', 'near', 'border', 'with', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'near',  
'border', 'with', 'Angola']]], 6, True)

['Ju\xc7\x80\xca\xbchoan', 'Language', 'codes', 'ISO', '639-3', 'ktz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
["K'iche'", 'Region', 'Central', 'highlands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
["K'iche'", 'Spoken', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
["K'iche'", 'Pronunciation', '/', 'k', '\xca', '\xbc', 'it', '\xca',
'\x83', 'e', '\xca', '\x94', '/']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',
'k', '\xca', '\xbc', 'it', '\xca', '\x83', 'e', '\xca', '\x94', '/']]], 13,
True)
```

```
["K'iche'", 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
["K'iche'", 'Regulated', 'by', 'Academia', 'de', 'Lenguas', 'Mayas', 'de',
'Guatemala', '(', 'ALMG', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Academia', 'de', 'Lenguas', 'Mayas', 'de', 'Guatemala', '\\(', 'ALMG', '\\
)']]], 12, True)
```

```
["K'iche'", 'Recognised', 'minority', 'language', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'Type: ANY']]], 6, True)
```

```
["K'iche'", 'Language', 'codes', 'ISO', '639-3', 'quc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jiwarli', 'Spoken', 'in', 'Western', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Jiwarli', 'Language', 'codes', 'ISO', '639-3', 'djl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Warrungu', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Warrungu', 'Total', 'speakers', '0']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Warrungu', 'Language', 'codes', 'ISO', '639-3', 'wrg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Spoken', 'in', 'Namibia', ',',  
'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Region', 'Okavango', 'and',  
'Ovamboland', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Okavango',  
'and', 'Ovamboland', 'Territory']], 6, True)

['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3',  
'oun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbc0', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbc0', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbc0', 'Language', 'codes', 'ISO',  
'639-3', 'aue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Naro', 'Spoken', 'in', 'Botswana', ',', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Naro', 'Language', 'codes', 'ISO', '639-3', 'nhr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['G\xc7\x80wi', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['G\\xc7\\x80wi', 'Total', 'speakers', '2500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['G\\xc7\\x80wi', 'Language', 'codes', 'ISO', '639-3', 'gwj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kxoe', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kxoe', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kxoe', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kxoe', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Kxoe', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kxoe', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kxoe', 'Language', 'codes', 'ISO', '639-3', ':', 'xuu', '\\xe2', '\\x80',  
'\\x93', 'Khwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'xuu', '\\xe2', '\\x80', '\\x93', 'Khwe']], ['NUM',  
'639-3']], 11, True)

['Kxoe', 'Language', 'codes', 'ISO', '639-3', 'hnh', '\\xe2', '\\x80',  
'\\x93', '\\xc7', '\\x81', 'Ani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hnh', '\\xe2', '\\x80', '\\x93', '\\xc7', '\\x81', 'Ani']],  
['NUM', '639-3']], 12, True)

['G\\xc7\\x81ana', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['G\x7\x81ana', 'Total', 'speakers', '2000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['G\x7\x81ana', 'Language', 'codes', 'ISO', '639-3', 'gnk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsoa', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tsoa', 'Total', 'speakers', '9300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tsoa', 'Dialects', 'Hiechware']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Tsoa', 'Dialects', 'Cua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Tsoa', 'Dialects', 'Cire', 'Cire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tsoa', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsoa', 'Language', 'codes', 'ISO', '639-3', ':', 'hio', '\xe2', '\x80',  
'\x93', 'Tsoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'hio', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Tsoa', 'Language', 'codes', 'ISO', '639-3', 'tyu', '\xe2', '\x80',  
'\x93', 'Kua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tyu', '\xe2', '\x80', '\x93', 'Kua']], ['NUM', '639-3']],  
10, True)

```

["A'tong", 'Spoken', 'in', 'India', ',', 'Nepal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

["A'tong", 'Region', 'India', ',', 'region', 'of', 'Assam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
',', 'Type: ANY', 'of', 'Type: ANY']]], 7, True)

["A'tong", 'Official', 'language', 'in', 'No', 'official', 'status']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'No', 'official', 'status']]], 7, True)

["A'tong", 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

["A'tong", 'Language', 'codes', 'ISO', '639-3', 'aot']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['A-Pucikwar', 'Total', 'speakers', 'extinct', '(', 'disputed', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'extinct', '\\(', 'disputed', '\\)']]], 7, True)

['A-Pucikwar', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['A-Pucikwar', 'Region', 'Andaman', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['A-Pucikwar', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['A-Pucikwar', 'Language', 'codes', 'ISO', '639-3', ':', 'apq', '\\xe2',
'\\x80', '\\x93', 'A-Pucikwar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'apq', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM',
'639-3']], 11, True)

['A-Pucikwar', 'Language', 'codes', 'ISO', '639-3', 'gac', '\\xe2', '\\x80',
'\\x93', 'Great', 'Andamanese', 'creole']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'gac', '\\xe2', '\\x80', '\\x93', 'Great', 'Andamanese',

```



'creole']], ['NUM', '639-3']], 12, True)

['Aari', 'Total', 'speakers', '129', ',', '350', 'monolinguals']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'monolinguals']], ['NUM', '129'], ['NUM', '350']], 7,  
True)

['Aari', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Aari', 'Region', 'north', 'central', 'Omo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Aari', 'Language', 'codes', 'ISO', '639-3', 'aiw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abanyom', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Abanyom', 'Region', 'Cross', 'River', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Abanyom', 'Language', 'codes', 'ISO', '639-2', 'bnt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Abanyom', 'Language', 'codes', 'ISO', '639-3', 'abm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abau', 'Region', 'Sandaun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Abau', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Abau', 'Language', 'codes', 'ISO', '639-3', 'aau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hixkary\x3\xa1na', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hixkary\x3\xa1na', 'Language', 'codes', 'ISO', '639-3', 'hix']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shwa', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shwa', 'Total', 'speakers', '6000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shwa', 'Language', 'codes', 'ISO', '639-3', 'shg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xiri', 'Total', 'speakers', '87', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '?']], ['NUM', '87']], 5, True)

['Xiri', 'Language', 'codes', 'ISO', '639-3', 'xii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tama', 'Spoken', 'in', 'Chad', 'and', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Tama', 'Region', 'western', 'Sudan', 'and', 'eastern', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western', 'Sudan', 'and', 'eastern', 'Chad']], 7, True)

['Tama', 'Language', 'codes', 'ISO', '639-3', 'tma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tama', 'Language', 'codes', 'Linguasphere', '05-PEA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Korana', 'Total', 'speakers', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Korana', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Korana', 'Language', 'codes', 'ISO', '639-3', 'kqz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kwadi', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kwadi', 'Language', 'codes', 'ISO', '639-3', 'kwz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sandawe', 'Region', 'Rift', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sandawe', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sandawe', 'Pronunciation', 's', '\xc3', '\xa0', 'nd', '\xc3', '\xa0', 'w', '\xc4', '\x9b', '\xcb', '\x90', 'k', '\xc3', '\xad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's', '\xc3', '\xa0', 'nd', '\xc3', '\xa0', 'w', '\xc4', '\x9b', '\xcb', '\x90', 'k', '\xc3', '\xad']]], 16, True)

['Sandawe', 'Language', 'codes', 'ISO', '639-2', 'sad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sandawe', 'Language', 'codes', 'ISO', '639-3', 'sad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hadza', 'Spoken', 'in', 'Lake', 'Eyasi', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lake', 'Eyasi', ',', 'Tanzania']]], 7, True)

['Hadza', 'Language', 'codes', 'ISO', '639-3', 'hts']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Proto-Mandarin', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Mandarin', 'Region', 'Medieval', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Mandarin', 'Writing', 'system', 'Seal', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Proto-Mandarin', 'Writing', 'system', 'Clerical', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Proto-Mandarin', 'Writing', 'system', 'Kaishu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Mandarin', 'Writing', 'system', 'Semi-cursive', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Proto-Mandarin', 'Writing', 'system', 'Grass', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Proto-Mandarin', 'Writing', 'system', 'Phagspa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Mandarin', 'Writing', 'system', 'Hangul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Mandarin', 'Official', 'language', 'in', 'Song', 'Dynasty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Proto-Mandarin', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Proto-Mandarin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Sudanese', 'Spoken', 'in', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sudanese', 'Region', 'northern']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northern']], 3, True)

['Sudanese', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Sudanese', 'Language', 'codes', 'ISO', '639-3', 'apd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ab\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'aba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Awjilah', 'Spoken', 'in', 'Libya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Awjilah', 'Region', 'Cyrenaica']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Awjilah', 'Language', 'codes', 'ISO', '639-3', 'auj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sokna', 'Total', 'speakers', 'Extinct', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Extinct', '?']], 5, True)

['Sokna', 'Spoken', 'in', 'Libya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Sokna', 'Region', 'Fezzan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Sokna', 'Language', 'codes', 'ISO', '639-3', 'swn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guanche', 'Spoken', 'in', 'Formerly', ',', 'Canary', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Formerly', ',', 'Canary', 'Islands']]], 7, True)

['Guanche', 'Region', 'Limited', 'to', 'the', 'islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Limited',
'to', 'the', 'islands']]], 6, True)

['Guanche', 'Language', 'codes', 'ISO', '639-3', 'gnc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gilbertese', 'Spoken', 'in', 'Kiribati']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gilbertese', 'Spoken', 'in', 'Fiji']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gilbertese', 'Spoken', 'in', 'Marshall', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Gilbertese', 'Spoken', 'in', 'Nauru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gilbertese', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Gilbertese', 'Spoken', 'in', 'Tuvalu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gilbertese', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Gilbertese', 'Official', 'language', 'in', 'Kiribati']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Gilbertese', 'Language', 'codes', 'ISO', '639-2', 'gil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Gilbertese', 'Language', 'codes', 'ISO', '639-3', 'gil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Montenegrin', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Montenegrin', 'Region', 'Southern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Montenegrin', 'Spoken', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Montenegrin', 'Pronunciation', 't', '\xcd', '\xa1', 'sr', '\xcc', '\xa9',  
'\xcb', '\x88', 'n', '\xc9', '\x94', '\xcc', '\x8c', '\xc9', '\xa1',  
'\xc9', '\x94', 'rski', '\xcb', '\x90']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't',  
'\xcd', '\xa1', 'sr', '\xcc', '\xa9', '\xcb', '\x88', 'n', '\xc9', '\x94',  
'\xcc', '\x8c', '\xc9', '\xa1', '\xc9', '\x94', 'rski', '\xcb', '\x90']]],  
22, True)

['Montenegrin', 'Official', 'language', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Montenegrin', 'Recognised', 'minority', 'language', 'in', 'Serbia', '(',  
'Municipality', 'of', 'Mali', 'I', '\xc4', '\x91', 'o', '\xc5', '\xa1',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Serbia', '\\(', 'Municipality', 'of', 'Mali', 'I',  
'\xc4', '\x91', 'o', '\xc5', '\xa1', '\\)']]], 17, True)

['Montenegrin', 'Language', 'codes', 'ISO', '639-3', 'srp', '\xe2', '\x80',

'\x93', 'Ethnologue', 'identifies', 'it', 'as', 'an', 'alternate', 'name', 'of', 'Serbian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'srp', '\xe2', '\x80', '\x93', 'Ethnologue', 'identifies', 'it', 'as', 'an', 'alternate', 'name', 'of', 'Type: ANY']], ['NUM', '639-3']], 18, True)

['Judeo-Berber', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Berber', 'Total', 'speakers', '2000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Berber', 'Writing', 'system', 'Hebrew', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Judeo-Berber', 'Language', 'codes', 'ISO', '639-3', 'jbe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shilha', 'Total', 'speakers', 'Uncertain', ',', '3', 'to', '9', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Uncertain', ',', 'NUM', 'to', 'NUM', 'million']], ['NUM', '3'], ['NUM', '9']], 9, True)

['Shilha', 'Spoken', 'in', 'Morocco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shilha', 'Region', 'High', 'Atlas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shilha', 'Region', 'Anti-Atlas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Shilha', 'Region', 'Souss']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Shilha', 'Language', 'codes', 'ISO', '639-3', 'shi']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ahom', 'Total', 'speakers', '(', 'religious', 'use', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\\\(', 'religious', 'use', '\\\\)']], 7, True)

['Ahom', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ahom', 'Region', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ahom', 'Writing', 'system', 'Ahom', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Ahom', 'Language', 'codes', 'ISO', '639-3', 'aho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hui', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hui', 'Region', 'southern', 'Anhui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hui', 'Region', 'neighbouring', 'portions', 'of', 'Zhejiang', 'and', 'Jiangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'neighbouring', 'portions', 'of', 'Zhejiang', 'and', 'Jiangxi']], 8, True)

['Hui', 'Region', 'about', '12', 'counties', 'in', 'total']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'about', 'NUM', 'counties', 'in', 'total']], ['NUM', '12']], 7, True)

['Hui', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Hui', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Hui', 'Language', 'codes', 'ISO', '639-3', 'czh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hui', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Walser', 'Spoken', 'in', 'the', 'Alps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Walser', 'Language', 'codes', 'ISO', '639-2', 'gsw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Walser', 'Language', 'codes', 'ISO', '639-3', 'wae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marshallese', 'Spoken', 'in', 'Marshall', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Marshallese', 'Spoken', 'in', 'Nauru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marshallese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Marshallese', 'Language', 'codes', 'ISO', '639-2', 'mah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Marshallese', 'Language', 'codes', 'ISO', '639-3', 'mah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marshallese', 'Language', 'codes', 'ISO', '639-1', 'mh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

```
['Tlingit', 'Region', 'Alaska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tlingit', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tlingit', 'Region', 'Yukon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tlingit', 'Region', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tlingit', 'Spoken', 'in', 'USA', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Tlingit', 'Writing', 'system', 'Latin', '(', 'Tlingit', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Tlingit', 'Pronunciation', '/', '\xc9', '\xac', '\xc9', '\xaa', '\xc5',  
'\x8b', 'k', '\xc9', '\xaa', '\xcc', '\x81', 't', '/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'\xc9', '\xac', '\xc9', '\xaa', '\xc5', '\x8b', 'k', '\xc9', '\xaa',  
'\xcc', '\x81', 't', '/']]], 16, True)
```

```
['Tlingit', 'Language', 'codes', 'ISO', '639-2', 'tli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Tlingit', 'Language', 'codes', 'ISO', '639-3', 'tli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['M\x3\x3cheno', 'Spoken', 'in', 'Italy', ',', 'Fersina', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy',  
, ',', 'Fersina', 'Valley']]], 7, True)
```

```
['M\x3\x3cheno', 'Language', 'codes', 'ISO', '639-3', 'mhn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Ingush', 'Spoken', 'in', 'Russia', ',', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Ingush', 'Region', 'Ingushetia', ',', 'Chechnya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ingush', 'Official', 'language', 'in', 'Ingushetia', '(', 'federal', 'subject', 'of', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Ingushetia', '\\\\(', 'federal', 'subject', 'of', 'Russia', '\\\\)']]], 11, True)

['Ingush', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Ingush', 'Language', 'codes', 'ISO', '639-2', 'inh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ingush', 'Language', 'codes', 'ISO', '639-3', 'inh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Altay', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Altay', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Altay', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Altay', 'Official', 'language', 'in', 'Altai', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Altay', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

```

['Altay', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Samaritan', 'Total', 'speakers', 'fewer', 'than', '1', ',', '000', ';',
'liturgical', 'only']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'fewer', 'than', 'NUM', ',', 'NUM', ';', 'liturgical', 'only']], ['NUM',
'1'], ['NUM', '000']], 11, True)

['Samaritan', 'Writing', 'system', 'Samaritan', 'abjad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Samaritan', 'Language', 'codes', 'ISO', '639-3', 'heb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Samaritan', 'Language', 'codes', 'Linguasphere', '12-AAB']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Fur', 'Spoken', 'in', 'Sudan', ',', 'Chad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Fur', 'Region', 'Darfur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Fur', 'Language', 'codes', 'ISO', '639-3', 'fvr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fur', 'Language', 'codes', 'ISO', '639-3', 'Geographic', 'distribution',
'of', 'Fur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Geographic', 'distribution', 'of', 'Type: ANY']], ['NUM',
'639-3']], 9, True)

['Fon', 'Spoken', 'in', 'Benin', ',', 'Togo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Fon', 'Language', 'codes', 'ISO', '639-2', 'fon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Fon', 'Language', 'codes', 'ISO', '639-3', 'fon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chuvash', 'Region', 'Chuvashia', 'and', 'adjacent', 'areas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Chuvashia',  
'and', 'adjacent', 'areas']]], 6, True)

['Chuvash', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chuvash', 'Pronunciation', 't', '\xc9', '\x95', '\xc9', '\x99', '\xca',  
'\x8b', 'a', '\xca', '\x82', '\xcb', '\x88', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't',  
'\xc9', '\x95', '\xc9', '\x99', '\xca', '\x8b', 'a', '\xca', '\x82',  
'\xcb', '\x88', 'la']]], 15, True)

['Chuvash', 'Official', 'language', 'in', 'Chuvashia', '(', 'federal',  
'subject', 'of', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Chuvashia', '\\(', 'federal', 'subject', 'of', 'Russia', '\\)']]],  
11, True)

['Chuvash', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Chuvash', 'Language', 'codes', 'ISO', '639-2', 'chv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chuvash', 'Language', 'codes', 'ISO', '639-3', 'chv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chuvash', 'Language', 'codes', 'ISO', '639-1', 'cv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Jakaltek', 'Spoken', 'in', 'Guatemala', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Jakaltek', 'Region', 'Huehuetenango', ',', 'Chiapas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Jakaltek', 'Language', 'codes', 'ISO', '639-3', 'jac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Vilamovian', 'Total', 'speakers', '70']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Vilamovian', 'Region', 'Wilamowice']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Vilamovian', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Vilamovian', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Vilamovian', 'Official', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Vilamovian', 'Regulated', 'by', 'no', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Vilamovian', 'Language', 'codes', 'ISO', '639-3', 'wym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Judeo-Italian', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Judeo-Italian', 'Pronunciation', 'd', '\xca', '\x92', 'ud', '\xc9',  
'\x9b', 'o', 'italiano', ',', 'i', '\xcb', '\x80', 't', '\xc3', '\xa6',  
'lki', '\xce', '\xb8']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'd',  
'\xca', '\x92', 'ud', '\xc9', '\x9b', 'o', 'italiano', ',', 'i', '\xcb',  
'\x80', 't', '\xc3', '\xa6', 'lki', '\xce', '\xb8']]], 20, True)

['Judeo-Italian', 'Spoken', 'in', 'Ferrara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Florence']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Mantua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Modena']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Piedmont']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Reggio', 'Emilia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Judeo-Italian', 'Spoken', 'in', 'Rome']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Venice']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Italian', 'Spoken', 'in', 'Livorno', ';', 'Corfu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Livorno', ';', 'Corfu']], 6, True)

['Judeo-Italian', 'Language', 'codes', 'ISO', '639-3', 'itk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chan\xc3\xa9', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chan\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'caj']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Biblical', 'Spoken', 'in', 'Kingdom', 'of', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kingdom', 'of', 'Israel']]], 6, True)

['Biblical', 'Spoken', 'in', 'Global', '(', 'as', 'a', 'liturgical', 'language', 'for', 'Judaism', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Global', '\\\\(', 'as', 'a', 'liturgical', 'language', 'for', 'Judaism', '\\\\)']]], 12, True)

['Biblical', 'Writing', 'system', 'Paleo-Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Biblical', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Biblical', 'Writing', 'system', 'Samaritan', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Biblical', 'Language', 'codes', 'ISO', '639-3', 'hbo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc3\x84ynu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['\xc3\x84ynu', 'Region', 'Xinjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['\xc3\x84ynu', 'Language', 'codes', 'ISO', '639-3', 'aib']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc3\x84ynu', 'Language', 'codes', 'ISO', '639-3', 'Map', 'showing', 'locations', 'of', '\xc3', '\x84', 'ynu', '(', 'red', ')', 'within', 'Xinjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Map', 'showing', 'locations', 'of', '\xc3', '\x84', 'ynu',

'\\(','red','\\)','within','Xinjiang']], ['NUM', '639-3']], 17, True)

['Tualatin\\xe2\\x80\\x93Yamhill', 'Spoken', 'in', 'Northwest', 'Oregon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tualatin\\xe2\\x80\\x93Yamhill', 'Total', 'speakers', '0', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '?']], ['NUM', '0']], 5, True)

['Tualatin\\xe2\\x80\\x93Yamhill', 'Dialects', 'Tualatin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tualatin\\xe2\\x80\\x93Yamhill', 'Dialects', 'Yamhill']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tualatin\\xe2\\x80\\x93Yamhill', 'Language', 'codes', 'ISO', '639-3', 'nrt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Santiam', 'Spoken', 'in', 'Northwest', 'Oregon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Mossi', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Mossi', 'Spoken', 'in', 'Benin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mossi', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mossi', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mossi', 'Spoken', 'in', 'Togo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Mossi', 'Language', 'codes', 'ISO', '639-2', 'mos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mossi', 'Language', 'codes', 'ISO', '639-3', 'mos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Idiom', 'Neutral', 'Total', 'speakers', 'virtually', 'extinct']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Neutral', 'Total',
'speakers', 'virtually', 'extinct']]], 6, True)

['Idiom', 'Date', 'founded', '1902']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Idiom', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Idiom', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary',
'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'int', '.', 'auxiliary', 'languages']]], 9, True)

['Idiom', 'Category', '(', 'purpose', ')', 'Volapuk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Idiom', 'Created', 'by', 'Waldemar', 'Rosenberger']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Idiom', 'Category', '(', 'sources', ')', 'a', 'posteriori', 'language',
',', 'developed', 'from', 'a', 'heavily', 'revised', 'form', 'of', 'Volap',
'\\xc3', '\\xbc', 'k']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)', 'a', 'posteriori', 'language', ',', 'developed', 'from',
'a', 'heavily', 'revised', 'form', 'of', 'Volap', '\\xc3', '\\xbc', 'k']]],
20, True)

['Idiom', 'Setting', 'and', 'usage', 'international', 'auxiliary',
'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',
'usage', 'international', 'auxiliary', 'language']]], 7, True)

```

```

['Idiom', 'Language', 'codes', 'ISO', '639-2', 'art']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Idiom', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Lacandon', 'Region', 'Chiapas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Lacandon', 'Spoken', 'in', 'Mexico']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Lacandon', 'Pronunciation', 'xat', '\xca', '\x83', 't', '\xca', '\xbc',
'a', '\xcb', '\x90', 'n']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'xat',
'\xca', '\x83', 't', '\xca', '\xbc', 'a', '\xcb', '\x90', 'n']], 12, True)

['Lacandon', 'Language', 'codes', 'ISO', '639-3', 'lac']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Neapolitan', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Neapolitan', 'Language', 'codes', 'ISO', '639-2', 'nap']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Neapolitan', 'Language', 'codes', 'ISO', '639-3', 'nap']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Neapolitan', 'Language', 'codes', 'Linguasphere', '51-AAA-rb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

['Neapolitan', 'Language', 'codes', 'Linguasphere', 'Dialects']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

```

```
['Noric', 'Spoken', 'in', 'Austria', ',', 'Slovenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Noric', 'Language', 'codes', 'ISO', '639-3', 'nrc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Shawia', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Shawia', 'Language', 'codes', 'ISO', '639-3', 'shy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sened', 'Spoken', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sened', 'Region', 'Sened', 'and', 'Tmagourt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sened', 'Language', 'codes', 'ISO', '639-3', 'sds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Osage', 'Total', 'speakers', '5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Osage', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Osage', 'Region', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Osage', 'Language', 'codes', 'ISO', '639-2', 'osa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Osage', 'Language', 'codes', 'ISO', '639-3', 'osa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

['ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Winnebago', 'Total', 'speakers', '230']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Winnebago', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Winnebago', 'Region', 'Wisconsin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Winnebago', 'Language', 'codes', 'ISO', '639-3', 'win']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mandan', 'Total', 'speakers', '1']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandan', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Mandan', 'Language', 'codes', 'ISO', '639-2', 'sio']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mandan', 'Language', 'codes', 'ISO', '639-3', 'mhq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Crow', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Crow', 'Region', 'Montana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Crow', 'Language', 'codes', 'ISO', '639-3', 'cro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Konkani', 'Region', 'Konkan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Konkani', 'Region', 'includes', 'the', 'state', 'of', 'Goa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'includes', 'the', 'state', 'of', 'Goa']]], 7, True)

['Konkani', 'Region', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Konkani', 'Region', 'Maharashtra', 'and', 'some', 'parts', 'of', 'Kerala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maharashtra', 'and', 'some', 'parts', 'of', 'Kerala']]], 8, True)

['Konkani', 'Region', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Konkani', 'Region', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Konkani', 'Region', 'Persian', 'Gulf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Region', 'Lisbon', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Konkani', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Writing', 'system', 'Devanagari', '(', 'official', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Konkani', 'Writing', 'system', 'Roman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Writing', 'system', 'Kannada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Writing', 'system', 'Malayalam', 'and', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Konkani', 'Pronunciation', 'k', '\xc3', '\xb5', 'k', '\xc9', '\xb5', '\xc9', '\xb3', 'i', '(', 'standard', ')', ',', 'k', '\xc3', '\xb5', 'k', '\xc9', '\xb3', 'i', '(', 'popular', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k', '\xc3', '\xb5', 'k', '\xc9', '\xb5', '\xc9', '\xb3', 'i', '\\(', 'standard', '\\)', ',', 'k', '\xc3', '\xb5', 'k', '\xc9', '\xb3', 'i', '\\(', 'popular', '\\)']]], 25, True)

['Konkani', 'Official', 'language', 'in', 'Goa', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', ',', 'Type: ANY']]], 7, True)

['Konkani', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Goa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Goa']]], 10, True)

['Konkani', 'Language', 'codes', 'ISO', '639-2', 'kok']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Konkani', 'Language', 'codes', 'ISO', '639-3', 'kok', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kok', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)

['Konkani', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Konkani', 'Language', 'codes', 'ISO', '639-3', ':', 'gom', '\xe2', '\x80', '\x93', 'Goan', 'Konkani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'gom', '\xe2', '\x80', '\x93', 'Goan', 'Type: ANY']]], ['NUM', '639-3']], 12, True)

['Konkani', 'Language', 'codes', 'ISO', '639-3', 'knn', '\xe2', '\x80',



'\x93', 'Mararashtra', 'Konkani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'knn', '\xe2', '\x80', '\x93', 'Mararashtra', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Magadhi', 'Total', 'speakers', '13', ',', '000', ',', '000', '(', 'As',  
'per', 'SIL', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'As', 'per', 'SIL', '\\)']], ['NUM',  
'13'], ['NUM', '000'], ['NUM', '000']], 13, True)

['Magadhi', 'Region', 'Bihar', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Magadhi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Magadhi', 'Writing', 'system', 'Devanagari', ',', 'Kaithi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Magadhi', 'Official', 'language', 'in', 'Bihar', 'state', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Bihar', 'state', 'in', 'India']], 8, True)

['Magadhi', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Magadhi', 'Language', 'codes', 'ISO', '639-2', 'mag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Magadhi', 'Language', 'codes', 'ISO', '639-3', 'mag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chhattisgarhi', 'Total', 'speakers', '11', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '11']], 5, True)

['Chhattisgarhi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```

['Chhattisgarhi', 'Region', 'Chhattisgarh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Chhattisgarhi', 'Region', 'Madhya', 'Pradesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chhattisgarhi', 'Region', 'Bihar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Chhattisgarhi', 'Region', 'Orissa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Chhattisgarhi', 'Language', 'codes', 'ISO', '639-2', 'hin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chhattisgarhi', 'Language', 'codes', 'ISO', '639-3', 'hne']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chhattisgarhi', 'Language', 'codes', 'Linguasphere', '59-AAF-ta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Nuosu', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nuosu', 'Region', 'southern', 'Sichuan', ',', 'northern', 'Yunnan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'southern', 'Sichuan', ',', 'northern', 'Yunnan']], 7, True)

['Nuosu', 'Language', 'codes', 'ISO', '639-3', 'iii']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nuosu', 'Language', 'codes', 'ISO', '639-1', 'ii']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tujia', 'Spoken', 'in', 'China', 'Ethnicity']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tujia', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tujia', 'Language', 'codes', 'ISO', '639-3', ':', 'tjs', '\xe2', '\x80', '\x93', 'Southern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'tjs', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Tujia', 'Language', 'codes', 'ISO', '639-3', 'tji', '\xe2', '\x80', '\x93', 'Northern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tji', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Ladakhi', 'Total', 'speakers', '600', ',', '000', ',', 'all', 'dialects']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', ',', 'all', 'dialects']], ['NUM', '600'], ['NUM', '000']], 9, True)

['Ladakhi', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Ladakhi', 'Region', 'Kashmir', '(', 'Ladakh', 'and', 'Baltistan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kashmir', '\(', 'Ladakh', 'and', 'Baltistan', '\)']]], 8, True)

['Ladakhi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ladakhi', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ladakhi', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ladakhi', 'Writing', 'system', 'Tibetan', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ladakhi', 'Language', 'codes', 'ISO', '639-3', 'lbj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kabyle', 'Region', 'Kabylie', '(', 'Provinces', 'of', 'Algiers', ',', ',', 'Tizi', 'Ouzou', ',', ',', 'Bejaia', ',', ',', 'Bouira', ',', ',', 'Boumerdes', ',', ',', 'S', '\xc3', '\xa9', 'tif', ',', ',', 'BBA', ',', ',', 'and', 'parts', 'of', 'Jijel', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kabylie', '\\(', 'Provinces', 'of', 'Algiers', ',', ',', 'Tizi', 'Ouzou', ',', ',', 'Bejaia', ',', ',', 'Bouira', ',', ',', 'Boumerdes', ',', ',', 'S', '\xc3', '\xa9', 'tif', ',', ',', 'BBA', ',', ',', 'and', 'parts', 'of', 'Jijel', '\\)']]], 29, True)

['Kabyle', 'Spoken', 'in', 'Algeria', ';', 'immigrant', 'communities', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Algeria', ';', 'immigrant', 'communities', 'in', 'France']]], 9, True)

['Kabyle', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kabyle', 'Spoken', 'in', 'Canada', 'and', 'elsewhere']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Kabyle', 'Writing', 'system', 'Berber', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Kabyle', 'Language', 'codes', 'ISO', '639-2', 'kab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Crimean', 'Region', 'Black', 'Sea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Spoken', 'in', 'Uzbekistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Spoken', 'in', 'Romania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Spoken', 'in', 'Bulgaria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Crimean', 'Pronunciation', 'q', '\xc9', '\xaf', 'r', '\xc9', '\xaf', 'mt', '\xc9', '\x91', 't', '\xc9', '\x91', 'r', '\xcb', '\x88', 'd', '\xcd', '\xa1', '\xca', '\x92', '\xc9', '\x91']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'q', '\xc9', '\xaf', 'r', '\xc9', '\xaf', 'mt', '\xc9', '\x91', 't', '\xc9', '\x91', 'r', '\xcb', '\x88', 'd', '\xcd', '\xa1', '\xca', '\x92', '\xc9', '\x91']]], 24, True)

['Crimean', 'Official', 'language', 'in', 'Ukraine', '(', 'Crimea', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Ukraine', '\\(', 'Crimea', '\\)']]], 8, True)

['Crimean', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Crimean', 'Recognised', 'minority', 'language', 'in', 'Romania', '(', 'Dobroudja', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Romania', '\\(', 'Dobroudja', '\\)']]], 9, True)

['Crimean', 'Language', 'codes', 'ISO', '639-3', 'crh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ripuarian', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ripuarian', 'Spoken', 'in', 'Belgium']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Ripuarian', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ripuarian', 'Region', 'North', 'Rhine-Westphalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ripuarian', 'Region', 'Rhineland-Palatinate']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ripuarian', 'Region', 'Li', '\xc3', '\xa8', 'ge']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Li', '\xc3', '\xa8', 'ge']]], 6, True)

['Ripuarian', 'Region', 'Limburg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ripuarian', 'Language', 'codes', 'ISO', '639-3', 'ksh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gorani', 'Spoken', 'in', 'Iraq', 'and', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Gorani', 'Region', 'Primarily', 'Hawraman', 'and', 'Garmian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Primarily', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Gorani', 'Language', 'codes', 'ISO', '639-3', 'hac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gorani', 'Language', 'codes', 'Linguasphere', '58-AAA-b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Raetic', 'Spoken', 'in', 'Ancient', 'Raetia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Raetic', 'Region', 'Eastern', 'Alps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Raetic', 'Language', 'codes', 'ISO', '639-3', 'xrr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Palaic', 'Spoken', 'in', 'Anatolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Palaic', 'Language', 'codes', 'ISO', '639-3', 'plq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aja', 'Language', 'codes', 'ISO', '639-3', 'aja']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bantik', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bantik', 'Region', 'North', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bantik', 'Language', 'codes', 'ISO', '639-3', 'bnq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Manado', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Manado', 'Region', 'North', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Manado', 'Language', 'codes', 'ISO', '639-3', 'xmm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Qashqai', 'Region', 'Fars']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Qashqai', 'Spoken', 'in', 'Iran', ',', 'Iraq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Qashqai', 'Writing', 'system', 'Perso-Arabic', 'script', ',', 'Latin',  
'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Qashqai', 'Language', 'codes', 'ISO', '639-3', 'qxq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nogai', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nogai', 'Region', 'Caucasus']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Nogai', 'Language', 'codes', 'ISO', '639-2', 'nog']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Nogai', 'Language', 'codes', 'ISO', '639-3', 'nog']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Late', 'Spoken', 'in', 'Ostrogothic', 'Kingdom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Late', 'Spoken', 'in', 'Gallic', 'Empire']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Late', 'Spoken', 'in', 'Palmyrene', 'Empire']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Late', 'Region', 'mare', 'nostrum', '(', 'Mediterranean', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'mare',  
'nostrum', '\\(', 'Mediterranean', '\\)']]], 7, True)



['Late', 'Writing', 'system', 'Latin', 'alphabet', 'Literature']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Late', 'Writing', 'system', 'documents']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Late', 'Regulated', 'by', 'Schools', 'of', 'grammar', 'and', 'rhetoric']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Schools', 'of', 'grammar', 'and', 'rhetoric']]], 8, True)

['Late', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Late', 'Language', 'codes', 'ISO', '639-3', 'lat', 'The', 'Late-Latin',  
'speaking', 'world', ',', '271', 'AD']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lat', 'The', 'Late-Latin', 'speaking', 'world', ',', 'NUM',  
'AD']]], ['NUM', '639-3'], ['NUM', '271']], 13, True)

['Late', 'Language', 'codes', 'ISO', '639-1', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Lori', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lori', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lori', 'Spoken', 'in', 'Kuwait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lori', 'Spoken', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lori', 'Spoken', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Lori', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lori', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lori', 'Language', 'codes', 'ISO', '639-3', 'bqi', '\xe2', '\x80',  
'\x93', 'Bakhtiari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bqi', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Lori', 'Language', 'codes', 'ISO', '639-3', 'zum', '\xe2', '\x80',  
'\x93', 'Kumzari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zum', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Old', 'Spoken', 'in', 'northwest', 'Germany', ',', 'northeast',  
'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'northwest', 'Germany', ',', 'northeast', 'Netherlands']], 8, True)
```

```
['Old', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'osx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shenwa', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Shenwa', 'Region', 'Mt', '.', 'Chenoua', ',', 'wilaya', 'of', 'Tipasa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mt', '.',  
'Chenoua', ',', 'wilaya', 'of', 'Tipasa']], 9, True)
```

```
['Shenwa', 'Language', 'codes', 'ISO', '639-2', 'ber']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```

['Shenwa', 'Language', 'codes', 'ISO', '639-3', 'cnu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warlpiri', 'Region', 'North', 'Central', 'Desert']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Warlpiri', 'Signed', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Warlpiri', 'Total', 'signers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',
'\xe2', '\x80', '\x93']]], 6, True)

['Pitkern', 'Region', 'Pacific']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Pitkern', 'Language', 'codes', 'ISO', '639-3', 'pih']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pitkern', 'Language', 'codes', 'Linguasphere', '52-ABB-dd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Bikya', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bikya', 'Region', 'North', 'West', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Bikya', 'Region', 'Menchum', 'Division']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bikya', 'Region', 'Furu-Awa', 'Subdivision']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bikya', 'Region', 'Furubana', 'village']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Bikya', 'Language', 'codes', 'ISO', '639-3', 'byb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wakhi', 'Writing', 'system', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wakhi', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wakhi', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wakhi', 'Language', 'codes', 'ISO', '639-3', 'wbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wakhi', 'Language', 'codes', 'Linguasphere', '58-ABD-c']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Siberian', 'Region', 'Bering', 'Strait', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Siberian', 'Writing', 'system', 'Latin', ',', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Siberian', 'Language', 'codes', 'ISO', '639-3', 'ess']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Spanish', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Spanish', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

```

['Spanish', 'Language', 'codes', 'ISO', '639-2', 'spa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Spanish', 'Language', 'codes', 'ISO', '639-3', 'spa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Spanish', 'Language', 'codes', 'ISO', '639-1', 'es']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Soddo', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Soddo', 'Region', 'Gurage', 'Zone', ',', 'Southern']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Zone', ',', 'Type: ANY']]], 6, True)

['Soddo', 'Language', 'codes', 'ISO', '639-3', 'gru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iberian', 'Spoken', 'in', 'Modern', 'Spain', 'and', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Iberian', 'Language', 'codes', 'ISO', '639-3', 'xib']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'pal', '(', 'see', 'text',
'left', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'pal', '\\\\(', 'see', 'text', 'left', '\\\\)']], ['NUM',
'639-2']], 11, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'pal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

```
['Middle', 'Language', 'codes', 'Linguasphere', '58-AAC-ca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Bashkardi', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bashkardi', 'Language', 'codes', 'ISO', '639-3', 'bsg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Medieval', 'Region', 'Most', 'of', 'western', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Most', 'of',  
'western', 'Europe']]], 6, True)
```

```
['Medieval', 'Official', 'language', 'in', 'Most', 'states']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Medieval', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Medieval', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Medieval', 'Language', 'codes', 'ISO', '639-3', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Medieval', 'Language', 'codes', 'ISO', '639-3', 'Europe', ',', '1000',  
'AD']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Europe', ',', 'NUM', 'AD']], ['NUM', '639-3'], ['NUM',  
'1000']], 9, True)
```

```
['Medieval', 'Language', 'codes', 'ISO', '639-1', 'la']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Bench', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Bench', 'Language', 'codes', 'ISO', '639-3', 'bcq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bench', 'Language', 'codes', 'Linguasphere', '16-BBA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Kwasio', 'Language', 'codes', 'ISO', '639-3', 'nmg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Waray-Waray', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Waray-Waray', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Waray-Waray', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Waray-Waray', 'Regulated', 'by', 'Historically', 'regulated', 'by', 'the', 'Sanghiran', 'san', 'Binisaya', 'ha', 'Samar', 'ug', 'Leyte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Historically', 'regulated', 'by', 'the', 'Sanghiran', 'san', 'Binisaya', 'ha', 'Samar', 'ug', 'Leyte']]], 14, True)

['Waray-Waray', 'Language', 'codes', 'ISO', '639-2', 'war']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Waray-Waray', 'Language', 'codes', 'ISO', '639-3', 'war']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kinaray-a', 'Region', 'Iloilo', 'and', 'Antique', 'provinces', ',', 'western', 'Panay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Iloilo', 'and', 'Antique', 'provinces', ',', 'western', 'Panay']]], 9, True)

['Kinaray-a', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kinaray-a', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Kinaray-a', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Kinaray-a', 'Language', 'codes', 'ISO', '639-3', 'krj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taus\xc5\xabg', 'Region', 'Jolo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Taus\xc5\xabg', 'Region', 'Sulu', ';', 'Palawan', 'Island', ';', 'Basilan', 'Island', ';', 'Tawi-Tawi', ';', 'and', 'Zamboanga', 'City', 'and', 'environs', '.', 'Also', 'spoken', 'in', 'Indonesia', '(', 'Kalimantan', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sulu', ';', 'Palawan', 'Island', ';', 'Basilan', 'Island', ';', 'Tawi-Tawi', ';', 'and', 'Zamboanga', 'City', 'and', 'environs', '.', 'Also', 'spoken', 'in', 'Indonesia', '\\\\(', 'Kalimantan', '\\\\)']], 25, True)

['Taus\xc5\xabg', 'Region', 'and', 'Malaysia', '(', 'Sabah', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'Malaysia', '\\\\(', 'Sabah', '\\\\)']], 7, True)

['Taus\xc5\xabg', 'Spoken', 'in', 'Mainly', 'in', 'the', 'Philippines', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mainly', 'in', 'the', 'Philippines', ',']], 8, True)

['Taus\xc5\xabg', 'Spoken', 'in', 'also', 'in', 'Malaysia', '(', 'Sabah', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'also', 'in', 'Malaysia', '\\\\(', 'Sabah', '\\\\)']], 9, True)

['Taus\xc5\xabg', 'Spoken', 'in', 'and', 'Indonesia', '(', 'Kalimantan', ')']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'Indonesia', '\\\\(', 'Kalimantan', '\\\\)']], 8, True)

['Taus\\xc5\\xabg', 'Writing', 'system', 'Latin', ';']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Taus\\xc5\\xabg', 'Writing', 'system', 'Jawi', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Taus\\xc5\\xabg', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Taus\\xc5\\xabg', 'Language', 'codes', 'ISO', '639-3', 'tsg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maranao', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maranao', 'Region', 'Central', 'Mindanao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maranao', 'Writing', 'system', 'Historically', 'written', 'in', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']], 7, True)

['Maranao', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Maranao', 'Language', 'codes', 'ISO', '639-3', 'mrw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maguindanao', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maguindanao', 'Region', 'Southwest', 'Mindanao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Maguindanao', 'Writing', 'system', 'Historically', 'written', 'in',  
'Jawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Maguindanao', 'Official', 'language', 'in', 'Regional', 'language', 'in',  
'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Maguindanao', 'Language', 'codes', 'ISO', '639-3', 'mdh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laal', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Laal', 'Region', 'Gori']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Laal', 'Region', 'Damtar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Laal', 'Region', 'Mailao', 'villages', 'in', 'Moyen-Chari',  
'prefecture']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mailao',  
'villages', 'in', 'Moyen-Chari', 'prefecture']]], 7, True)

['Laal', 'Language', 'codes', 'ISO', '639-3', 'gdm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gulf', 'Spoken', 'in', 'Kuwait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gulf', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Gulf', 'Spoken', 'in', 'Bahrain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gulf', 'Spoken', 'in', 'Qatar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gulf', 'Spoken', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gulf', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Gulf', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Gulf', 'Language', 'codes', 'ISO', '639-3', 'afb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mpre', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mpre', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mpre', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jalaa', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jalaa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jalaa', 'Region', 'Loojaa', 'settlement', 'in', 'Balanga', 'Local',

'Government', 'Area', ',', 'Bauchi', 'State']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Region', 'Loojaa',  
 'settlement', 'in', 'Balanga', 'Local', 'Government', 'Area', ',',  
 'Bauchi', 'State' ] ] ], 12, True)

['Jalaa', 'Language', 'codes', 'ISO', '639-3', 'cet']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes',  
 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Shabo', 'Spoken', 'in', 'Ethiopia']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY' ] ] ], 4, True)

['Shabo', 'Region', 'western', 'SNNPR']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY' ] ] ], 4, True)

['Shabo', 'Language', 'codes', 'ISO', '639-3', 'sbf']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes',  
 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Shabo', 'Language', 'codes', 'Linguasphere', '05-PEA-aa']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes',  
 'Linguasphere', 'Type: ANY' ] ] ], 5, True)

['Nepal', 'Dialects', 'Dolakhae']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type:  
 ANY' ] ] ], 3, True)

['Nepal', 'Region', 'South', 'Asia']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY' ] ] ], 4, True)

['Nepal', 'Spoken', 'in', 'Nepal']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY' ] ] ], 4, True)

['Nepal', 'Spoken', 'in', 'India']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY' ] ] ], 4, True)

['Nepal', 'Spoken', 'in', 'Bhutan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY' ] ] ], 4, True)

['Nepal', 'Writing', 'system', 'Devanagari']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Nepal', 'Writing', 'system', 'Kutakshari', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Writing', 'system', 'Ranjana', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Writing', 'system', 'Prachalit', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Writing', 'system', 'Brahmi', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Writing', 'system', 'Gupta', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Writing', 'system', 'Bhujimol', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Writing', 'system', 'Golmol', 'script']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Official', 'language', 'in', 'Nepal']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)
```

```
['Nepal', 'Language', 'codes', 'ISO', '639-2', 'new']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Nepal', 'Language', 'codes', 'ISO', '639-3', 'new']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yerkish', 'Total', 'speakers', '3', '(', 'ape', ')']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\(', 'ape', '\\)']], ['NUM', '3']], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'Animal', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'Yerkish']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY']]], 6, True)
```

```
['Yerkish', 'Created', 'by', 'Ernst', 'von', 'Glasersfeld']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Ernst',  
'von', 'Glasersfeld']]], 6, True)
```

```
['Yerkish', 'Category', '(', 'sources', ')', 'Use', 'a', 'keyboard', 'to',  
'punch', 'keys', 'with', 'so', 'called', 'lexigrams', ',', 'symbols',  
'corresponding', 'to', 'objects', 'or', 'ideas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)'], 'Use', 'a', 'keyboard', 'to', 'punch', 'keys', 'with',  
'so', 'called', 'lexigrams', ',', 'symbols', 'corresponding', 'to',  
'objects', 'or', 'ideas']]], 22, True)
```

```
['Yerkish', 'Setting', 'and', 'usage', 'Animal', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Type: ANY', 'language']]], 6, True)
```

```
['Yerkish', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Scythian', 'Spoken', 'in', 'Scythia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Scythian', 'Language', 'codes', 'ISO', '639-3', 'xsc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Scythian', 'Language', 'codes', 'Linguasphere', '58', '(', 'Iranic',  
'in', 'general', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'NUM', '\\\\(', 'Iranic', 'in', 'general', '\\\\)']], ['NUM',  
'58']], 10, True)
```

```
['Hunnic', 'Spoken', 'in', 'from', 'Eurasian', 'steppe', 'into', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'from',  
'Eurasian', 'steppe', 'into', 'Europe']]], 8, True)
```

```
['Hunnic', 'Language', 'codes', 'ISO', '639-3', 'xhc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cuman', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cuman', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cuman', 'Region', 'Cumania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cuman', 'Language', 'codes', 'ISO', '639-3', 'qwm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nuer', 'Total', 'speakers', '740', ',', '000', 'in', 'Sudan', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Sudan', ',', '']], ['NUM', '740'], ['NUM', '000']],  
9, True)
```

```
['Nuer', 'Total', 'speakers', '153', ',', '407', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '153'], ['NUM', '407']], 8,  
True)
```

```
['Nuer', 'Spoken', 'in', 'Sudan', ',', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Nuer', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nuer', 'Language', 'codes', 'ISO', '639-3', 'nus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Beothuk', 'Language', 'extinction', '1829']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Beothuk', 'Spoken', 'in', 'Newfoundland']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Beothuk', 'Language', 'codes', 'ISO', '639-3', 'bue']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shambala', 'Spoken', 'in', 'Tanzania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Shambala', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Shambala', 'Language', 'codes', 'ISO', '639-3', 'ksb']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bagirmi', 'Spoken', 'in', 'Chad', ',', 'Nigeria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Bagirmi', 'Language', 'codes', 'ISO', '639-3', 'bmi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Faliscan', 'Spoken', 'in', 'ancient', 'Italy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Faliscan', 'Region', 'Lazio']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Faliscan', 'Language', 'codes', 'ISO', '639-3', 'xfa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yola', 'Spoken', 'in', 'Ireland']:
```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yola', 'Region', 'Wexford']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yola', 'Language', 'codes', 'ISO', '639-3', 'yol']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yola', 'Language', 'codes', 'Linguasphere', '52-ABA-bd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['New', 'Language', 'codes', 'ISO', '639-2', 'sgn-NZ']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Yagnobi', 'Spoken', 'in', 'Tajikistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yagnobi', 'Region', 'originally', 'from', 'Yagnob', 'Valley']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'originally', 'from', 'Yagnob', 'Valley']]], 6, True)

['Yagnobi', 'Language', 'codes', 'ISO', '639-3', 'yai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yagnobi', 'Language', 'codes', 'Linguasphere', '58-ABC-a']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Livvi-Karelian', 'Region', 'Republic', 'of', 'Karelia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Livvi-Karelian', 'Region', 'Inhabiting', 'the', 'area', 'between', 'Ladoga', 'and', 'Onega', 'lakes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Inhabiting', 'the', 'area', 'between', 'Ladoga', 'and', 'Onega', 'lakes']]], 10, True)

['Livvi-Karelian', 'Region', 'northward', 'of', 'Svir', 'River']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'north', 'of',

'Type: ANY', 'River']]], 6, True)

['Livvi-Karelian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Livvi-Karelian', 'Spoken', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Livvi-Karelian', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']]], 9, True)

['Livvi-Karelian', 'Official', 'language', 'in', ':', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Livvi-Karelian', 'Official', 'language', 'in', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'Karelia']]], 7, True)

['Livvi-Karelian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Livvi-Karelian', 'Language', 'codes', 'ISO', '639-3', 'olo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shan', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shan', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shan', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shan', 'Region', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY']]], 3, True)

['Shan', 'Language', 'codes', 'ISO', '639-2', 'shn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Shan', 'Language', 'codes', 'ISO', '639-3', 'shn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kujarg\xc3\xa9', 'Spoken', 'in', 'Chad', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Kujarg\xc3\xa9', 'Region', 'Jebel', 'Mirra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kujarg\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'vkj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ibanag', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Ibanag', 'Region', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Ibanag', 'Language', 'codes', 'ISO', '639-3', 'ibg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amami', 'Spoken', 'in', 'Kagoshima', 'Prefecture', 'and', 'Okinawa',  
'Prefecture', ',', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kagoshima', 'Prefecture', 'and', 'Okinawa', 'Prefecture', ',', 'Japan']],  
10, True)

['Amami', 'Region', 'Amami', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Amami', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Amami', 'Language', 'codes', 'ISO', '639-3', 'ryn', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ryn', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Amami', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Miyako', 'Spoken', 'in', 'Okinawa', ',', 'Japan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Miyako', 'Region', 'Miyako', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Miyako', 'Language', 'codes', 'ISO', '639-3', 'mvi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yaeyama', 'Spoken', 'in', 'Japan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yaeyama', 'Region', 'Yaeyama', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yaeyama', 'Language', 'codes', 'ISO', '639-3', 'rys']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yonaguni', 'Total', 'speakers', '800']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yonaguni', 'Spoken', 'in', 'Japan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yonaguni', 'Region', 'Yonaguni']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY'[]], 3, True)

['Yonaguni', 'Language', 'codes', 'ISO', '639-3', 'yoi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yucatec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yucatec', 'Spoken', 'in', 'Belize']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yucatec', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yucatec', 'Region', '098']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Yucatec', 'Region', 'Quintana', 'Roo', '163']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Yucatec', 'Region', '477']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Yucatec', 'Region', 'Campeche', '75']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yucatec', 'Region', '847']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Yucatec', 'Official', 'language', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Yucatec', 'Regulated', 'by', 'INALI']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Yucatec', 'Language', 'codes', 'ISO', '639-3', 'yua']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Griko', 'Spoken', 'in', 'Italy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Griko', 'Region', 'Southern']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Griko', 'Region', 'east', 'of', 'Reggio', 'Calabria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'east', 'of',
'Reggio', 'Calabria']]], 6, True)
```

```
['Griko', 'Region', 'Salento', 'and', 'Aspromonte']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Griko', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)
```

```
['Griko', 'Language', 'codes', 'ISO', '639-3', 'ell-gri']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kham,', 'Region', 'Rapti', 'Zone', ',', 'Rolpa', 'and', 'Rukum',
'Districts']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rapti',
'Zone', ',', 'Rolpa', 'and', 'Rukum', 'Districts']]], 9, True)
```

```
['Kham,', 'Region', 'Dhaulagiri', 'Zone', ',', 'Baglung', 'District']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Dhaulagiri',
'Zone', ',', 'Baglung', 'District']]], 7, True)
```

```
['Kham,', 'Spoken', 'in', 'Nepal']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kham,', 'Writing', 'system', 'Devanagari']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

['Kham,', 'Official', 'language', 'in', 'No', 'official', 'status']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'No', 'official', 'status']]], 7, True)

['Kham,', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Kham,', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kham,', 'Language', 'codes', 'ISO', '639-3', 'kgj', '\xe2', '\x80',  
'\x93', 'Gamale', 'Kham']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kgj', '\xe2', '\x80', '\x93', 'Gamale', 'Kham']]], ['NUM',  
'639-3']], 11, True)

['Kham,', 'Language', 'codes', 'ISO', '639-3', 'kip', '\xe2', '\x80',  
'\x93', 'Sheshi', 'Kham']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kip', '\xe2', '\x80', '\x93', 'Sheshi', 'Kham']]], ['NUM',  
'639-3']], 11, True)

['Yugtun', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Yugtun', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Yugtun', 'Region', 'Central', 'Alaska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yugtun', 'Writing', 'system', 'Yugtun', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Yugtun', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Oropom', 'Spoken', 'in', 'Kenya', ',', 'Uganda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Oropom', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['T\x3\xafcuna', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['T\x3\xafcuna', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['T\x3\xafcuna', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['T\x3\xafcuna', 'Language', 'codes', 'ISO', '639-3', 'tca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Amdang', 'Spoken', 'in', 'Chad', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Amdang', 'Region', 'Wadi', 'Fira']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Amdang', 'Language', 'codes', 'ISO', '639-3', 'amj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Inupiaq', 'Spoken', 'in', 'United', 'States', ',', 'formerly', 'Russia', ';', 'Northwest', 'Territories', 'of', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', ',', 'formerly', 'Russia', ';', 'Northwest', 'Territories', 'of', 'Canada']]], 13, True)

['Inupiaq', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Inupiaq', 'Language', 'codes', 'ISO', '639-2', 'ipk']:



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Inupiaq', 'Language', 'codes', 'ISO', '639-3', 'ipk', '\xe2', '\x80', '\x93', 'Macrolanguage']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ipk', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)
```

```
['Inupiaq', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Inupiaq', 'Language', 'codes', 'ISO', '639-3', 'esk', '\xe2', '\x80', '\x93', 'Northwest', 'Alaska', 'Inupiatun']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'esk', '\xe2', '\x80', '\x93', 'Northwest', 'Alaska', 'Inupiatun']], ['NUM', '639-3']], 12, True)
```

```
['Inupiaq', 'Language', 'codes', 'ISO', '639-1', 'ik']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Manyika', 'Spoken', 'in', 'Zimbabwe']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Manyika', 'Spoken', 'in', 'Mozambique']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Manyika', 'Region', 'Southern', 'Africa']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Manyika', 'Language', 'codes', 'ISO', '639-3', 'mxc']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Komi', 'Spoken', 'in', 'Russia']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Komi', 'Writing', 'system', 'Cyrillic', 'alphabet']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)
```

['Komi', 'Official', 'language', 'in', 'Komi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Komi', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Komi', 'Language', 'codes', 'ISO', '639-2', 'kom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Komi', 'Language', 'codes', 'ISO', '639-3', 'kom', '\xe2', '\x80',  
\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kom', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Komi', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Komi', 'Language', 'codes', 'ISO', '639-3', ':', 'koi', '\xe2', '\x80',  
\x93', 'Komi-Permyak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'koi', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Komi', 'Language', 'codes', 'ISO', '639-3', 'kpv', '\xe2', '\x80',  
\x93', 'Komi-Zyrian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kpv', '\xe2', '\x80', '\x93', 'Komi-Zyrian']], ['NUM',  
'639-3']], 10, True)

['Komi', 'Language', 'codes', 'ISO', '639-1', 'kv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Yevanic', 'Spoken', 'in', 'Originally', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Yevanic', 'Spoken', 'in', 'more', 'recently', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'more',  
'recently', 'Israel']]], 6, True)

['Yevanic', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yevanic', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yevanic', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Yevanic', 'Language', 'codes', 'ISO', '639-3', 'yej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yevanic', 'Language', 'codes', 'Linguasphere', '56-AAA-am']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Huarijio', 'Spoken', 'in', 'Mexico', ':', 'Chihuahua', ',', 'Sonora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', ':', 'Chihuahua', ',', 'Sonora']], 8, True)

['Huarijio', 'Total', 'speakers', '5000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Huarijio', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Huarijio', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']], 16, True)

['Huarijio', 'Language', 'codes', 'ISO', '639-3', 'var']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shtokavian', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Shtokavian', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']]]], 6, True)

['Shtokavian', 'Spoken', 'in', 'Serbia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Spoken', 'in', 'Kosovo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Spoken', 'in', 'Montenegro']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Spoken', 'in', 'Romania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Spoken', 'in', 'Hungary']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Pronunciation']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']]]], 2, True)

['Shtokavian', 'Standard', 'forms', 'Serbian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Standard', 'forms', 'Croatian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Standard', 'forms', 'Bosnian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Standard', 'forms', 'Montenegrin']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)

['Shtokavian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:

```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Shtokavian', 'Language', 'codes', 'ISO', '639-3', 'hrv-sht']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kajkavian', 'Pronunciation']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']]], 2, True)
```

```
['Kajkavian', 'Spoken', 'in', 'Croatia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Kajkavian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Kajkavian', 'Language', 'codes', 'ISO', '639-3', 'hrv-kay']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chakavian', 'Pronunciation']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']]], 2, True)
```

```
['Chakavian', 'Spoken', 'in', 'Croatia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Chakavian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Chakavian', 'Language', 'codes', 'ISO', '639-3', 'hrv-cha']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chakavian', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', '\xc4', '\x8c', 'akavian']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Distribution', 'of', '\xc4', '\x8c', 'akavian']]], ['NUM',
```

'639-3']], 10, True)

['Jamtlandic', 'Spoken', 'in', 'J', '\xc3', '\xa4', 'mtland', ',', 'H', '\xc3', '\xa4', 'rjedalen']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'J', '\xc3', '\xa4', 'mtland', ',', 'H', '\xc3', '\xa4', 'rjedalen']]], 12, True)

['Jamtlandic', 'Official', 'language', 'in', 'Republic', 'of', 'Jamtland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'Jamtland']]], 7, True)

['Jamtlandic', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Jamtlandic', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Jamtlandic', 'Language', 'codes', 'ISO', '639-3', 'jmk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bajan', 'Spoken', 'in', 'Barbados']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bajan', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Bajan', 'Language', 'codes', 'ISO', '639-3', 'bjs']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bajan', 'Language', 'codes', 'Linguasphere', '52-ABB-ar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Early', 'Spoken', 'in', 'England']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Early', 'Spoken', 'in', 'southern', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Early', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Early', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Tianjin', 'dialect', 'Total', 'speakers', 'Approx', '.', '6', 'million',  
'est']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'Approx', '.', 'NUM', 'million', 'est']], ['NUM', '6']], 9,  
True)
```

```
['Tianjin', 'Region', 'Northeastern', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tianjin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Old', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'ojp', '\xe2', '\x80', '\x93',  
'but', 'see', 'caveat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ojp', '\xe2', '\x80', '\x93', 'but', 'see', 'caveat']],  
['NUM', '639-3']], 12, True)
```

```
['Goguryeo', 'Spoken', 'in', 'Goguryeo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Goguryeo', 'Dialects', 'Buyeo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Goguryeo', 'Dialects', '?', 'Baekje']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Goguryeo', 'Region', 'Manchuria', ',', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Goguryeo', 'Language', 'codes', 'ISO', '639-3', 'zkg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Guugu', 'Spoken', 'in', 'Hopevale', ',', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Guugu', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Guugu', 'Language', 'codes', 'ISO', '639-3', 'kky']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kala', 'Total', 'speakers', '3000-4000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kala', 'Language', 'codes', 'ISO', '639-3', 'mwp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Antakarinya', 'Language', 'codes', 'ISO', '639-3', 'ant']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yankunytjatjara', 'Spoken', 'in', 'South', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Yankunytjatjara', 'Language', 'codes', 'ISO', '639-3', 'kdd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ghotuo', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



['Ghotuo', 'Region', 'Edo', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ghotuo', 'Language', 'codes', 'ISO', '639-3', 'aaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Roma', 'Spoken', 'in', 'Jerusu', 'village']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Roma', 'Spoken', 'in', 'Roma', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Roma', 'Spoken', 'in', 'Maluku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Roma', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Roma', 'Language', 'codes', 'ISO', '639-3', 'rmm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tobian', 'Total', 'speakers', '100']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tobian', 'Spoken', 'in', 'Palau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tobian', 'Region', 'Hatohebei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tobian', 'Region', 'Koror']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tobian', 'Region', 'Sonsorol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

ANY'']], 3, True)

['Tobian', 'Language', 'codes', 'ISO', '639-2', 'map']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tobian', 'Language', 'codes', 'ISO', '639-3', 'tox']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Herero', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Herero', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Herero', 'Region', 'Kunene', ',', 'Omaheke', 'and', 'Otjozondjupa', 'in',  
'Namibia', ';', 'Ghanzi', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kunene', ',',  
'Omaheke', 'and', 'Otjozondjupa', 'in', 'Namibia', ';', 'Ghanzi', 'in',  
'Botswana']], 13, True)

['Herero', 'Language', 'codes', 'ISO', '639-2', 'her']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Herero', 'Language', 'codes', 'ISO', '639-3', 'her']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Herero', 'Language', 'codes', 'ISO', '639-1', 'hz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hz']], ['NUM', '639-1']], 6, True)

['Choctaw', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Choctaw', 'Language', 'codes', 'ISO', '639-2', 'cho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Choctaw', 'Language', 'codes', 'ISO', '639-3', 'cho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kutchi', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kutchi', 'Spoken', 'in', 'Trinidad', 'And', 'Tobago', 'Also', 'UK']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Trinidad', 'And', 'Tobago', 'Also', 'UK']], 8, True)

['Kutchi', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kutchi', 'Spoken', 'in', 'Kenya', '(', 'Some', 'Parts', 'Of', 'Guyana', ')', 'and', 'others']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kenya', '\\\\(', 'Some', 'Parts', 'Of', 'Guyana', '\\\\)', 'and', 'others']], 12, True)

['Kutchi', 'Writing', 'system', 'Arabic', 'script', ',', 'Gujarati', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Kutchi', 'Official', 'language', 'in', 'No', 'official', 'status']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'No', 'official', 'status']], 7, True)

['Kutchi', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Kutchi', 'Language', 'codes', 'ISO', '639-3', 'kfr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kumzari', 'Spoken', 'in', 'Oman']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kumzari', 'Region', 'Kumzar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kumzari', 'Writing', 'system', 'Persian', 'alphabet']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kumzari', 'Language', 'codes', 'ISO', '639-3', 'zum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sogdian', 'Spoken', 'in', 'Sogdiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sogdian', 'Language', 'codes', 'ISO', '639-2', 'sog']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Sogdian', 'Language', 'codes', 'ISO', '639-3', 'sog']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kok-borok', 'Spoken', 'in', 'India', 'and', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)
```

```
['Kok-borok', 'Region', 'Tripura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Kok-borok', 'Region', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Kok-borok', 'Region', 'Mizoram']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Kok-borok', 'Region', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Kok-borok', 'Region', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Kok-borok', 'Official', 'language', 'in', 'India', '(', 'Tripura', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\\\(', 'Type: ANY', '\\\\)']]], 8, True)
```

```
['Kok-borok', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Kok-borok', 'Language', 'codes', 'ISO', '639-3', 'trp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Gikuyu', 'Region', 'Central', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gikuyu', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gikuyu', 'Pronunciation', '\xc9', '\xa3', '\xc4', '\x93', 'k', '\xc5',  
'\x8d', 'j', '\xc3', '\xb3']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',  
'\xa3', '\xc4', '\x93', 'k', '\xc5', '\x8d', 'j', '\xc3', '\xb3']]], 12,  
True)
```

```
['Gikuyu', 'Language', 'codes', 'ISO', '639-2', 'kik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Gikuyu', 'Language', 'codes', 'ISO', '639-3', 'kik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Gikuyu', 'Language', 'codes', 'ISO', '639-1', 'ki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Saurashtra', 'Spoken', 'in', 'Tamil', 'Nadu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Saurashtra', 'Spoken', 'in', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Saurashtra', 'Spoken', 'in', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Saurashtra', 'Spoken', 'in', 'Andhra', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Saurashtra', 'Spoken', 'in', 'Kerala', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Saurashtra', 'Writing', 'system', 'Saurashtra', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Saurashtra', 'Writing', 'system', 'Tamil', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Saurashtra', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Saurashtra', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Saurashtra', 'Language', 'codes', 'ISO', '639-3', 'saz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Renaissance', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Renaissance', 'Official', 'language', 'in', 'Most', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Renaissance', 'Regulated', 'by', 'The', 'community', 'of', 'scholars', 'at', 'the', 'earliest', 'universities']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'The', 'community', 'of', 'scholars', 'at', 'the', 'earliest', 'universities']]], 11, True)

['Renaissance', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Renaissance', 'Language', 'codes', 'ISO', '639-3', 'lat', 'Europe', 'in', '1648', 'AD']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lat', 'Europe', 'in', 'NUM', 'AD']], ['NUM', '639-3'], ['NUM', '1648']], 10, True)

['Renaissance', 'Language', 'codes', 'ISO', '639-1', 'la']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Chol', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chol', 'Region', 'North', 'Central', 'Chiapas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Chol', 'Language', 'codes', 'ISO', '639-3', 'ctu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalanga', 'Spoken', 'in', 'Zimbabwe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalanga', 'Spoken', 'in', 'Botswana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalanga', 'Region', 'Southern', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalanga', 'Language', 'codes', 'ISO', '639-3', 'kck']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Penang', 'Region', 'originally']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Penang', 'Region', 'but', 'not', 'limited', 'to']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'but', 'not', 'limited', 'to']], 6, True)

```

['Penang', 'Region', 'Penang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Penang', 'Signed', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Selangor', 'Region', 'Kuala', 'Lumpur', 'and', 'Selangor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kuala', 'Lumpur', 'and', 'Type: ANY']], 6, True)

['Selangor', 'Signed', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malaysian', 'Signed', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malaysian', 'Language', 'codes', 'ISO', '639-2', 'sgn-MY']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Beijing', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Beijing', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Taishun', 'Manjiang', '\xe6', '\xb3', '\xb0', '\xe9', '\xa0', '\x86', '\xe8', '\xa0', '\xbb', '\xe8', '\xac', '\x9b', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Manjiang', '\xe6', '\xb3', '\xb0', '\xe9', '\xa0', '\x86', '\xe8', '\xa0', '\xbb', '\xe8', '\xac', '\x9b', 'Total', 'speakers']], 16, True)

['Taishun', 'Spoken', 'in', 'Southern', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Taishun', 'Region', 'Taishun']:

```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Taishun', 'Region', 'Cangnan', 'county']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taishun', 'Region', 'Wenzhou', 'prefecture']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taishun', 'Region', 'Zhejiang', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taishun', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Taishun', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Taishun', 'Language', 'codes', 'ISO', '639-3', 'cdo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Taishun', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Dongxiang', 'Region', 'Gansu', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dongxiang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dongxiang', 'Writing', 'system', 'Arabic', 'alphabet', ',', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Dongxiang', 'Language', 'codes', 'ISO', '639-3', 'sce']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cantonese', 'Spoken', 'in', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'United', 'Kingdom', ':', 'London']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'Kingdom', ':', 'London']]], 7, True)

['Cantonese', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Canada', ':', 'Vancouver']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',

['Cantonese', 'Spoken', 'in', 'Toronto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Edmonton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Calgary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Malaysia', ':', 'Kuala', 'Lumpur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Malaysia', ':', 'Kuala', 'Lumpur']]], 7, True)

['Cantonese', 'Spoken', 'in', 'Perak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Sandakan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Ipoh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Kota', 'Kinabalu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'San', 'Francisco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Los', 'Angeles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Oakland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Seattle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cantonese', 'Spoken', 'in', 'Can', 'Tho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Bac', 'Lieu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Da', 'Nang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Kien', 'Giang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Spoken', 'in', 'Quang', 'Ninh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
\xe2', '\x80', '\x93']]], 6, True)

['Cantonese', 'Writing', 'system', 'Written', 'Cantonese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Cantonese', 'Official', 'language', 'in', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Cantonese', 'Official', 'language', 'in', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Cantonese', 'Regulated', 'by', 'Civil', 'Service', 'Bureau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Civil', 'Service', 'Bureau']]], 6, True)

['Cantonese', 'Regulated', 'by', 'Government', 'of', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Government', 'of', 'Hong', 'Kong']]], 7, True)

['Cantonese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Cantonese', 'Language', 'codes', 'ISO', '639-3', 'yue-can']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Neo-Mandaic', 'Spoken', 'in', 'Iran', ',', 'Iraq', '(', 'extinct', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iran',

',' , 'Iraq', '\\(', 'extinct', '\\)']]], 9, True)

['Neo-Mandaic', 'Writing', 'system', 'Mandaic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Neo-Mandaic', 'Pronunciation', 'm', '\\xc3', '\\xa6', 'nd', '\\xc9', '\\x94',  
'\\xcb', '\\x90', '\\xcb', '\\x88', 'ji', ',', 'r', '\\xc3', '\\xa6', 't',  
'\\xcb', '\\xa4', 'n', '\\xc9', '\\x94']]:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm',  
'\\xc3', '\\xa6', 'nd', '\\xc9', '\\x94', '\\xcb', '\\x90', '\\xcb', '\\x88', 'ji',  
' ,', 'r', '\\xc3', '\\xa6', 't', '\\xcb', '\\xa4', 'n', '\\xc9', '\\x94']]], 23,  
True)

['Neo-Mandaic', 'Language', 'codes', 'ISO', '639-3', 'mid']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mbula', 'Region', 'Morobe', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mbula', 'Language', 'codes', 'ISO', '639-3', 'mna']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Teochew', 'Spoken', 'in', 'China']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Vietnam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'India', 'Malaysia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Teochew', 'Spoken', 'in', 'Singapore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Teochew', 'Spoken', 'in', 'United', 'States', 'of', 'America']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Teochew', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Teochew', 'Region', 'in', 'China', ':', 'eastern', 'Guangdong', 'province', 'including', 'Chaozhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'in', 'China', ':', 'eastern', 'Guangdong', 'province', 'including', 'Chaozhou']]], 10, True)

['Teochew', 'Region', 'Shantou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Teochew', 'Region', 'Jieyang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Teochew', 'Region', 'Chaoyang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

```
['Teochew', 'Region', 'Puning']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Teochew', 'Region', "Chao'an"]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Teochew', 'Region', 'Raoping']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Teochew', 'Region', 'Huilai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Teochew', 'Region', 'and', 'Shanwei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Teochew', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Teochew', 'Language', 'codes', 'ISO', '639-3', 'nan-cha', '(', 'Chao-  
Shan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nan-cha', '\\(', 'Chao-Shan', '\\)']], ['NUM', '639-3']], 9,  
True)
```

```
['Teochew', 'Language', 'codes', 'ISO', '639-3', 'nan-chs', '(',  
'Chaozhou', 'proper', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nan-chs', '\\(', 'Chaozhou', 'proper', '\\)']], ['NUM',  
'639-3']], 10, True)
```

```
['Masalit', 'Spoken', 'in', 'Sudan', ',', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Masalit', 'Language', 'codes', 'ISO', '639-3', 'mls']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Suba', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Suba', 'Region', 'Lake', 'Victoria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Suba', 'Language', 'codes', 'ISO', '639-3', 'sxb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saho', 'Total', 'speakers', '191', ',', '000', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '191'], ['NUM', '000']], 8, True)

['Saho', 'Total', 'speakers', '32', ',', '832', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '32'], ['NUM', '832']], 8, True)

['Saho', 'Spoken', 'in', 'Eritrea', ',', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Saho', 'Region', 'southern', 'Eritrea', ',', 'Tigray', 'Province', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'southern', 'Eritrea', ',', 'Tigray', 'Province', 'in', 'Ethiopia']], 9, True)

['Saho', 'Language', 'codes', 'ISO', '639-3', 'ssy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Malaysian', 'Spoken', 'in', 'Brunei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malaysian', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malaysian', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



['Malaysian', 'Official', 'language', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Malaysian', 'Official', 'language', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Malaysian', 'Language', 'codes', 'ISO', '639-3', 'zsm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Grand', 'Duchy', 'of', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Grand', 'Duchy', 'of', 'Lithuania']]], 7, True)

['Old', 'Spoken', 'in', 'Polish-Lithuanian', 'Commonwealth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Spoken', 'in', 'part', 'of', 'the', 'Grand', 'Duchy', 'of', 'Moscow', '(', 'probably', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'part', 'of', 'the', 'Grand', 'Duchy', 'of', 'Moscow', '\\(', 'probably', '\\)']]], 13, True)

['Old', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Senhaja', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Senhaja', 'Spoken', 'in', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Senhaja', 'Region', 'Rif']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Senhaja', 'Language', 'codes', 'ISO', '639-3', 'sjs']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ghomara', 'Spoken', 'in', 'Morocco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ghomara', 'Language', 'codes', 'ISO', '639-3', 'gho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Riff', 'Total', 'speakers', '4', 'million', '(', 'incl', '.', 'abroad', 'Morocco', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', '\\\\(', 'incl', '.', 'abroad', 'Morocco', '\\\\)']], ['NUM', '4']], 11, True)

['Riff', 'Region', 'Rif', ',', 'Ceuta', 'and', 'Melilla']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rif', ',', 'Ceuta', 'and', 'Melilla']], 7, True)

['Riff', 'Spoken', 'in', 'Morocco', 'Spain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Riff', 'Writing', 'system', 'Arabic', ',', 'Tifinagh', 'and', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Riff', 'Language', 'codes', 'ISO', '639-3', 'rif']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Riff', 'Language', 'codes', 'ISO', '639-3', 'Area', 'of', 'Rif']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY', 'of', 'Type: ANY']], ['NUM', '639-3']], 8, True)

['Akawaio', 'Spoken', 'in', 'Guyana', ',', 'Venezuela', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Guyana', ',', 'Venezuela', 'Ethnicity']], 7, True)

['Akawaio', 'Spoken', 'in', 'Ingarik', '\\xc3', '\\xb3']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ingarik', '\\xc3', '\\xb3']], 6, True)

['Akawaio', 'Total', 'speakers', '5000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Akawaio', 'Language', 'codes', 'ISO', '639-3', 'ake']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dari', '(', 'Zoroastrian', ')', 'Total', 'speakers', '8', ',', ',', '000-15', ',', ',', '000', 'native', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', '\\\\(', 'Zoroastrian', '\\\\)', 'Total', 'speakers', 'NUM', ',', ',', 'NUM', ',', ',', 'NUM', 'native', 'speakers']], ['NUM', '8'], ['NUM', '000-15'], ['NUM', '000']], 13, True)

['Dari', 'Spoken', 'in', 'Yazd', 'and', 'Kerman', ',', ',', '(', 'Iran', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Yazd', 'and', 'Kerman', ',', ',', '\\\\(', 'Iran', '\\\\)']]], 10, True)

['Dari', 'Region', 'Central', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Language', 'codes', 'ISO', '639-3', 'gbz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rodi', 'Spoken', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rodi', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Rodi', 'Language', 'codes', 'ISO', '639-3', 'rmg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Udmurt', 'Spoken', 'in', 'Russia', ',', ',', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', ',', 'Type: ANY']]], 6, True)

['Udmurt', 'Region', 'Udmurtia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Udmurt', 'Official', 'language', 'in', 'Udmurtia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Udmurt', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Udmurt', 'Language', 'codes', 'ISO', '639-2', 'udm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Udmurt', 'Language', 'codes', 'ISO', '639-3', 'udm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Shimaore', 'Spoken', 'in', 'Mayotte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shimaore', 'Language', 'codes', 'ISO', '639-3', 'swb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Egypt', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Old', 'Region', 'Along', 'the', 'banks', 'of', 'the', 'Nile', 'in', 'what', 'is', 'now', 'southern', 'Egypt', 'and', 'northern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Along', 'the', 'banks', 'of', 'the', 'Nile', 'in', 'what', 'is', 'now', 'southern', 'Egypt', 'and', 'northern', 'Sudan']]], 17, True)

['Old', 'Writing', 'system', 'Coptic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'nub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'onw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'found', 'at', 'Pakhoras']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'found', 'at', 'Pakhoras']], ['NUM', '639-3']], 8, True)

['Harauti', 'Spoken', 'in', 'India', '(', 'Hadoti', 'region', 'of', 'Rajasthan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', '\\(', 'Hadoti', 'region', 'of', 'Rajasthan', '\\)']], 10, True)

['Harauti', 'Language', 'codes', 'ISO', '639-2', 'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Harauti', 'Language', 'codes', 'ISO', '639-3', 'hoj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ludic', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ludic', 'Region', 'Near', 'the', 'northeastern', 'shore', 'of', 'Lake', 'Onega', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Near', 'the', 'northeastern', 'shore', 'of', 'Lake', 'Onega', ',']], 10, True)

['Ludic', 'Region', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Ludic', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']], 9, True)

['Ludic', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Karelia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Karelia']], 8, True)

['Ludic', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Ludic', 'Language', 'codes', 'ISO', '639-3', 'lud']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Tagoi', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba',  
'Hills', ',', 'Sudan']]], 7, True)
```

```
['Tagoi', 'Language', 'codes', 'ISO', '639-3', 'tag']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taiap', 'Language', 'codes', 'ISO', '639-3', 'gpn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Delaware', 'Spoken', 'in', 'United', 'States', ',', 'in', 'modern',  
'times', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', ',', 'in', 'modern', 'times', 'Canada']]], 10, True)
```

```
['Delaware', 'Language', 'codes', 'ISO', '639-2', 'del']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Delaware', 'Language', 'codes', 'ISO', '639-3', 'del', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'del', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)
```

```
['Delaware', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Delaware', 'Language', 'codes', 'ISO', '639-3', ':', 'umu', '\xe2',  
'\x80', '\x93', 'Munsee']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'umu', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Delaware', 'Language', 'codes', 'ISO', '639-3', 'unm', '\xe2', '\x80',  
'\x93', 'Unami']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'unm', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Northern', 'Spoken', 'in', 'Norway']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Northern', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Northern', 'Spoken', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Northern', 'Total', 'speakers', '15', ',', '000-25', ',', '000', '(',  
'estimated', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', '\\\\(', 'estimated', '\\\\)']], ['NUM', '15'],  
['NUM', '000-25'], ['NUM', '000']], 11, True)

['Northern', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Northern', 'Language', 'codes', 'ISO', '639-2', 'sme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'sme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-1', 'se']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Louisiana', 'Spoken', 'in', 'Louisiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Louisiana', 'Spoken', 'in', 'particularly', 'St', '.', 'Martin',  
'Parish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'particularly', 'St', '.', 'Martin', 'Parish']], 8, True)

['Louisiana', 'Spoken', 'in', 'Natchitoches', 'Parish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Louisiana', 'Spoken', 'in', 'St', '.', 'Landry', 'Parish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'St',

['.', 'Landry', 'Parish']]], 7, True)

['Louisiana', 'Spoken', 'in', 'Jefferson', 'Parish', 'and', 'Lafayette', 'Parish']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Jefferson', 'Parish', 'and', 'Lafayette', 'Parish']]], 8, True)

['Louisiana', 'Language', 'codes', 'ISO', '639-3', 'lou']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Louisiana', 'Language', 'codes', 'Linguasphere', '51-AAC-ca']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Tokelauan', 'Spoken', 'in', 'Tokelau']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tokelauan', 'Official', 'language', 'in', 'Tokelau']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Tokelauan', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tokelauan', 'Language', 'codes', 'ISO', '639-3', 'tkl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ivatan/', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ivatan/', 'Region', 'Batanes', 'Islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ivatan/', 'Language', 'codes', 'ISO', '639-3', 'ivv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yahgan/Y\xc3\xa1mana', 'Total', 'speakers', '1']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



['Yahgan/Y\xc3\xa1mana', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yahgan/Y\xc3\xa1mana', 'Region', 'Tierra', 'del', 'Fuego']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Yahgan/Y\xc3\xa1mana', 'Language', 'codes', 'ISO', '639-3', 'yag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Matumbi', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Matumbi', 'Region', 'Kilwa', 'district']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Matumbi', 'Language', 'codes', 'ISO', '639-3', 'mgw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yuchi', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Yuchi', 'Region', 'East', 'central', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Yuchi', 'Language', 'codes', 'ISO', '639-3', 'yuc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Totonac', 'Spoken', 'in', 'Mexico', ':', 'Puebla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', ':', 'Puebla']], 6, True)

['Totonac', 'Spoken', 'in', 'Veracruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Totonac', 'Spoken', 'in', 'Hidalgo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Totonac', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Totonac', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']]], 16, True)

['Totonac', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Totonac', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

["Weyt'o", 'Spoken', 'in', 'Lake', 'Tana', ',', 'Ethiopia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lake', 'Tana', ',', 'Ethiopia', 'Ethnicity']]], 8, True)

["Weyt'o", 'Spoken', 'in', 'Weyto', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

["Weyt'o", 'Language', 'codes', 'ISO', '639-3', 'woy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bung', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bung', 'Region', 'Adamawa', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bung', 'Language', 'codes', 'ISO', '639-3', 'bqd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bete', 'Spoken', 'in', 'Takum', 'Local', 'Government', 'Authority']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Takum', 'Local', 'Government', 'Authority']]], 7, True)

['Bete', 'Spoken', 'in', 'Taraba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bete', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bete', 'Language', 'codes', 'ISO', '639-3', 'byf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tigre', 'Spoken', 'in', 'Eritrea', ',', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Tigre', 'Language', 'codes', 'ISO', '639-2', 'tig']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tigre', 'Language', 'codes', 'ISO', '639-3', 'tig']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adamorobe', 'Signed', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sabaic', 'Total', 'speakers', 'Extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sabaic', 'Spoken', 'in', 'Yemen']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sabaic', 'Spoken', 'in', 'Oman']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sabaic', 'Spoken', 'in', 'Saudi', 'Arabia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sabaic', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sabaic', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sabaic', 'Language', 'codes', 'ISO', '639-3', 'xsa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tzeltal', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tzeltal', 'Region', 'Chiapas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tzeltal', 'Language', 'codes', 'ISO', '639-3', 'tzh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tzotzil', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tzotzil', 'Region', 'Chiapas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tzotzil', 'Language', 'codes', 'ISO', '639-3', 'tzo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Irish', 'Total', 'signers', '40', ',', '000', 'daily']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'NUM', ',', 'NUM', 'daily']], ['NUM', '40'], ['NUM', '000']], 7, True)

['Nubi', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nubi', 'Spoken', 'in', 'Kenya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nubi', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nubi', 'Language', 'codes', 'ISO', '639-3', 'kcn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kwavi', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kwavi', 'Language', 'codes', 'ISO', '639-3', 'mas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lufu', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lufu', 'Region', 'Taraba', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lufu', 'Language', 'codes', 'ISO', '639-3', 'ldq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lengue', 'Spoken', 'in', 'Equatorial', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lengue', 'Language', 'codes', 'ISO', '639-3', 'bvb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'formerly', 'the', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'formerly', 'the', 'Netherlands']]], 6, True)

['Old', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Northern', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Old', 'Spoken', 'in', 'Western', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'frk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['F\&#x3\xa1', 'Spoken', 'in', 'Equatorial', 'Guinea', '(', 'mainly', 'on',  
'Annob', '\&#x3', '\&#xb3', 'n', 'island', ';', 'some', 'speakers', 'on',  
'Bioko', 'island', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Equatorial', 'Guinea', '\\&#x3', 'mainly', 'on', 'Annob', '\&#x3', '\&#xb3',  
'n', 'island', ';', 'some', 'speakers', 'on', 'Bioko', 'island', '\\&#x3']]],  
20, True)

['F\&#x3\xa1', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['F\&#x3\xa1', 'Language', 'codes', 'ISO', '639-3', 'fab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['F\&#x3\xa1', 'Language', 'codes', 'Linguasphere', '51-AAC-ae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Pothohari', 'Spoken', 'in', 'Azad', 'Kashmir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Pothohari', 'Spoken', 'in', 'Pothohar', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Pothohari', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Pothohari', 'Spoken', 'in', 'United', 'Kingdom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Pothohari', 'Language', 'codes', 'ISO', '639-3', 'phr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Southern', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Southern', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Tsat', 'Spoken', 'in', 'Hainan', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tsat', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsat', 'Language', 'codes', 'ISO', '639-3', 'huq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Samogitian', 'Spoken', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Samogitian', 'Region', 'Samogitia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Samogitian', 'Language', 'codes', 'ISO', '639-3', 'sgs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wangkumara', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wangkumara', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wangkumara', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Wangkumara', 'Language', 'codes', 'ISO', '639-3', 'nbx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Antillean', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Antillean', 'Language', 'codes', 'ISO', '639-3', ':', 'gcf', '\xe2', '\x80', '\x93', 'Guadeloupean', 'Creole', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'gcf', '\xe2', '\x80', '\x93', 'Guadeloupean', 'Creole', 'Type: ANY']], ['NUM', '639-3']], 13, True)

['Tumbuka', 'Total', 'speakers', '2', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '2']], 5, True)

['Tumbuka', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tumbuka', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tumbuka', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tumbuka', 'Region', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tumbuka', 'Language', 'codes', 'ISO', '639-2', 'tum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)



```

['Tumbuka', 'Language', 'codes', 'ISO', '639-3', 'tum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mandingo', 'Spoken', 'in', 'Mali']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mandingo', 'Spoken', 'in', 'Senegal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mandingo', 'Spoken', 'in', 'The', 'Gambia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Mandingo', 'Spoken', 'in', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mandingo', 'Spoken', 'in', 'Burkina', 'Faso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Mandingo', 'Spoken', 'in', 'Sierra', 'Leone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Mandingo', 'Spoken', 'in', 'Liberia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mandingo', 'Spoken', 'in', 'Guinea-Bissau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mandingo', 'Spoken', 'in', 'Chad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mandingo', 'Language', 'codes', 'ISO', '639-3', 'mnk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['T\x3\xbbr\x3\xb4yo', 'Region', 'Mardin', 'Province', 'of',

```

'southeastern', 'Turkey', ';', 'Al', 'Hasakah', 'and', 'Qamishli', 'in', 'northeastern', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mardin', 'Province', 'of', 'southeastern', 'Turkey', ';', 'Al', 'Hasakah', 'and', 'Qamishli', 'in', 'northeastern', 'Syria']]], 15, True)

['T\x3\xbb\x3\x4yo', 'Spoken', 'in', 'Turkey', ',', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['T\x3\xbb\x3\x4yo', 'Spoken', 'in', 'Assyrian', 'diaspora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['T\x3\xbb\x3\x4yo', 'Pronunciation', 't', '\xcb', '\xa4', 'ur', '\xcb', '\x88', '\xc9', '\x94', 'j', '\xc9', '\x94']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't', '\xcb', '\xa4', 'ur', '\xcb', '\x88', '\xc9', '\x94', 'j', '\xc9', '\x94']]], 13, True)

['T\x3\xbb\x3\x4yo', 'Pronunciation', 's', '\xcb', '\xa4', 'u', '\xcb', '\x88', 'rajt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's', '\xcb', '\xa4', 'u', '\xcb', '\x88', 'rajt']]], 9, True)

['T\x3\xbb\x3\x4yo', 'Pronunciation', 'sur', '\xcb', '\x88', 'j', '\xc9', '\x94', 'j', '\xc9', '\x94']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sur', '\xcb', '\x88', 'j', '\xc9', '\x94', 'j', '\xc9', '\x94']]], 11, True)

['T\x3\xbb\x3\x4yo', 'Language', 'codes', 'ISO', '639-3', 'tru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Region', 'Kermanshah']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Southern', 'Region', 'Ilam', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Region', 'Kirind']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

```
['Southern', 'Region', 'and', 'Qorwaq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-2', 'kur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'sdh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Language', 'codes', 'Linguasphere', '58-AAA-c']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Hmong', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Hmong', 'Spoken', 'in', 'Vietnam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Hmong', 'Spoken', 'in', 'Laos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Hmong', 'Spoken', 'in', 'Thailand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Hmong', 'Spoken', 'in', 'and', 'USA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Hmong', 'Total', 'speakers', 'over', '4', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'over', 'NUM', 'million']]], ['NUM', '4']], 6, True)
```

```
['Hmong', 'Writing', 'system', 'Pahawh', 'Hmong', 'script', ',', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Hmong', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hmong', 'Language', 'codes', 'ISO', '639-3', 'cqd', '\xe2', '\x80', '\x93', 'Chuanqiandian', 'Cluster', 'Miao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cqd', '\xe2', '\x80', '\x93', 'Chuanqiandian', 'Cluster', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hrm', '\xe2', '\x80', '\x93', 'Horned', 'Miao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hrm', '\xe2', '\x80', '\x93', 'Horned', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Hmong', 'Language', 'codes', 'ISO', '639-3', 'sfm', '\xe2', '\x80', '\x93', 'Small', 'Flowery', 'Miao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sfm', '\xe2', '\x80', '\x93', 'Small', 'Flowery', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmd', '\xe2', '\x80', '\x93', 'Large', 'Flowery', 'Miao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hmd', '\xe2', '\x80', '\x93', 'Large', 'Flowery', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hml', '\xe2', '\x80', '\x93', 'Luopohe', 'Hmong']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hml', '\xe2', '\x80', '\x93', 'Luopohe', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Enets', 'Total', 'speakers', '70']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Enets', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Enets', 'Region', 'Krasnoyarsk', 'Krai', ',', 'along', 'the', 'lower', 'Yenisei', 'River']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Krasnoyarsk', 'Krai', ',', 'along', 'the', 'lower', 'Yenisei', 'River']], 10, True)

['Enets', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mlahs\xc3\xb4', 'Spoken', 'in', 'Syria', ',', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Mlahs\xc3\xb4', 'Region', 'Qamishli', 'in', 'northeastern', 'Syria', ',', 'two', 'villages', 'in', 'Diyarbak', '\xc4', '\xb1', 'r', 'Province', 'of', 'southeastern', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Qamishli', 'in', 'northeastern', 'Syria', ',', 'two', 'villages', 'in', 'Diyarbak', '\xc4', '\xb1', 'r', 'Province', 'of', 'southeastern', 'Turkey']], 18, True)

['Mlahs\xc3\xb4', 'Language', 'codes', 'ISO', '639-3', 'lhs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yauma', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yauma', 'Language', 'codes', 'ISO', '639-3', 'yax']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fer', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Fer', 'Spoken', 'in', 'Central', 'African', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Republic']], 6, True)

['Fer', 'Region', 'Biraó']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Fer', 'Language', 'codes', 'ISO', '639-3', 'kah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yeni', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yeni', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yeni', 'Language', 'codes', 'ISO', '639-3', 'yei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Creek', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Creek', 'Region', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Creek', 'Region', 'Alabama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Creek', 'Region', 'Georgia', 'and', 'Florida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Creek', 'Language', 'codes', 'ISO', '639-2', 'mus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Creek', 'Language', 'codes', 'ISO', '639-3', 'mus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gayle', 'Spoken', 'in', 'Pretoria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gayle', 'Spoken', 'in', 'Cape', 'Town']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Gayle', 'Spoken', 'in', 'Durban']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gayle', 'Spoken', 'in', 'Bloemfontein']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Gayle', 'Spoken', 'in', 'and', 'Port', 'Elizabeth']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'Port', 'Elizabeth']]], 6, True)
```

```
['Gayle', 'Total', 'speakers', '20', ',', '000', 'second-', 'or', 'third-language', 'speakers']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'second-', 'or', 'third-language', 'speakers']], ['NUM', '20'], ['NUM', '000']], 10, True)
```

```
['Gayle', 'Language', 'codes', 'ISO', '639-3', 'gic']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lanc-Patu\x3\xa1', 'Spoken', 'in', 'Brazil', '(', 'in', 'the', 'state', 'of', 'Amap', '\x3\xa1', ')']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Brazil', '\\(', 'in', 'the', 'state', 'of', 'Amap', '\x3\xa1', '\\)']]], 13, True)
```

```
['Lanc-Patu\x3\xa1', 'Language', 'codes', 'ISO', '639-3', 'kmv']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lanc-Patu\x3\xa1', 'Language', 'codes', 'Linguasphere', '51-AAC-cdd']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Yaqui', 'Region', 'Sonora', ',', 'Arizona']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Yaqui', 'Pronunciation', 'jo', '\xca', '\x94', 'em', 'noki']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'jo', '\xca', '\x94', 'em', 'noki']]], 7, True)
```

```
['Yaqui', 'Language', 'codes', 'ISO', '639-3', 'yaq']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lomavren', 'Spoken', 'in', 'Armenia']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

['Lomavren', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lomavren', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lomavren', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lomavren', 'Total', 'speakers', '50']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lomavren', 'Language', 'codes', 'ISO', '639-3', 'rmi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Badaga', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Badaga', 'Writing', 'system', 'Tamil', 'script', 'or', 'English', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Badaga', 'Language', 'codes', 'ISO', '639-3', 'bfq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Angami', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Angami', 'Region', 'Nagaland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Angami', 'Language', 'codes', 'ISO', '639-3', 'njm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



```
['Bruu', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bruu', 'Total', 'speakers', '20', ',', '000', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'speakers']], ['NUM', '20'], ['NUM', '000']], 7, True)
```

```
['Bruu', 'Writing', 'system', 'Lao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bruu', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bruu', 'Language', 'codes', 'ISO', '639-3', 'brv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aimaq', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aimaq', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aimaq', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aimaq', 'Region', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aimaq', 'Language', 'codes', 'ISO', '639-3', 'aiq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Alyutor', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Alyutor', 'Region', 'Kamchatka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Alyutor', 'Language', 'codes', 'ISO', '639-3', 'alr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Askunu', 'Total', 'speakers', '1', ',', '200', '(', 'Ethnologue', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']], ['NUM', '1'], ['NUM',  
'200']], 9, True)

['Askunu', 'Spoken', 'in', 'Afghanistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Askunu', 'Region', 'Pech', 'Valley']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Askunu', 'Language', 'codes', 'ISO', '639-3', 'ask']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Askunu', 'Language', 'codes', 'Linguasphere', '58-ACA-a']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Awadhi', 'Dialects', 'Fiji', 'Hindi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Awadhi', 'Region', 'Kapilbastu', 'District', ';', 'Bheri', 'Zone']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kapilbastu',  
'District', ';', 'Bheri', 'Zone']]], 7, True)

['Awadhi', 'Region', 'Banke', 'District', 'and', 'Bardiya', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Banke',  
'District', 'and', 'Bardiya', 'District']]], 7, True)

['Awadhi', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Awadhi', 'Spoken', 'in', 'Nepal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```

['Awadhi', 'Writing', 'system', 'Devanagari', ',', 'Kaithi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Awadhi', 'Official', 'language', 'in', 'No', 'official', 'status']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'No', 'official', 'status']]], 7, True)

['Awadhi', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Awadhi', 'Language', 'codes', 'ISO', '639-2', 'awa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Awadhi', 'Language', 'codes', 'ISO', '639-3', 'awa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Asi', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Asi', 'Region', 'Western', 'Visayas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Asi', 'Language', 'codes', 'ISO', '639-3', 'bno']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Surigaonon', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Surigaonon', 'Region', 'Surigao', 'del', 'Norte', ',', 'most', 'part',
'of', 'Surigao', 'del', 'Sur', 'and', 'some', 'portion', 'of', 'Agusan',
'del', 'Norte']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Surigao',
'del', 'Norte', ',', 'most', 'part', 'of', 'Surigao', 'del', 'Sur', 'and',
'some', 'portion', 'of', 'Agusan', 'del', 'Norte']]], 19, True)

['Surigaonon', 'Language', 'codes', 'ISO', '639-3', 'sgd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romblomanon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Romblomanon', 'Region', 'Eastern', 'Visayas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Romblomanon', 'Language', 'codes', 'ISO', '639-3', 'rol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nafaanra', 'Language', 'codes', 'ISO', '639-3', 'nfr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nafaanra', 'Language', 'codes', 'ISO', '639-3', 'Nafaanra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nafaanra', 'Language', 'codes', 'ISO', '639-3', 'some', 'neighbouring', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'some', 'neighbouring', 'languages']], ['NUM', '639-3']], 8, True)

['Nafaanra', 'Language', 'codes', 'ISO', '639-3', 'and', 'other', 'Senufo', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'and', 'other', 'Senufo', 'languages']], ['NUM', '639-3']], 9, True)

['Masbate\x3\xb1o', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Masbate\x3\xb1o', 'Region', 'Eastern', 'Visayas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Masbate\x3\xb1o', 'Language', 'codes', 'ISO', '639-3', 'msb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Ratagnon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ratagnon', 'Language', 'codes', 'ISO', '639-3', 'btn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Pontic', 'Region', 'Southeastern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pontic', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Pontic', 'Writing', 'system', 'Greek', 'alphabet', ',', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Pontic', 'Language', 'codes', 'ISO', '639-3', 'pnt']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cappadocian', 'Total', 'speakers', 'very', 'few', ',', 'previously', 'thought', 'to', 'be', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'very', 'few', ',', 'previously', 'thought', 'to', 'be', 'extinct']]], 11, True)

['Cappadocian', 'Language', 'codes', 'ISO', '639-3', 'cpg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warsaw', 'dialect', 'Spoken', 'in', 'Poland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Spoken', 'in', 'Poland']]], 5, True)

['Warsaw', 'Region', 'Warsaw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Warsaw', 'Writing', 'system', 'Latin', '(', 'Polish', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Warsaw', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Zuni', 'Region', 'Western', 'New', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Zuni', 'Language', 'codes', 'ISO', '639-2', 'zun']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zuni', 'Language', 'codes', 'ISO', '639-3', 'zun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zuni', 'Language', 'codes', 'ISO', '639-3', 'Pre-European', 'contact', 'distribution', 'of', 'Zuni']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Pre-European', 'contact', 'distribution', 'of', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Carib', 'Spoken', 'in', 'Kali'na']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Carib', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Carib', 'Spoken', 'in', 'French', 'Guiana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Carib', 'Spoken', 'in', 'Guyana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Carib', 'Spoken', 'in', 'Suriname']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Carib', 'Spoken', 'in', 'and', 'Venezuela', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'Venezuela', 'Ethnicity']], 6, True)

['Carib', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Carib', 'Language', 'codes', 'ISO', '639-2', 'car']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Carib', 'Language', 'codes', 'ISO', '639-3', 'car']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ligbi', 'Spoken', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ligbi', 'Language', 'codes', 'ISO', '639-3', 'lig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alabama', 'Total', 'speakers', '100']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Alabama', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Alabama', 'Region', 'Texas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Alabama', 'Language', 'codes', 'ISO', '639-3', 'akz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laki', 'Spoken', 'in', 'Iran', ',', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Laki', 'Region', 'Central', 'Zagros']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Laki', 'Language', 'codes', 'ISO', '639-3', 'lki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laki', 'Language', 'codes', 'Linguasphere', '58-AAC-aac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Koasati', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koasati', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)



['Koasati', 'Region', 'Elton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Koasati', 'Region', 'Louisiana', 'and', 'Livingston']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Koasati', 'Region', 'Texas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Koasati', 'Language', 'codes', 'ISO', '639-3', 'cku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cuyonon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cuyonon', 'Region', 'Palawan', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cuyonon', 'Region', 'Cuyo', 'Islands', 'between', 'Palawan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Cuyo', 'Islands', 'between', 'Palawan']], 6, True)

['Cuyonon', 'Region', 'and', 'Panay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cuyonon', 'Language', 'codes', 'ISO', '639-3', 'cyo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Caluyanon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Caluyanon', 'Region', 'Caluya', 'Islands', ',', 'Antique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Island', ',', 'Type: ANY']], 6, True)

['Caluyanon', 'Language', 'codes', 'ISO', '639-3', 'clu']:

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Paiwan', 'Spoken', 'in', 'Taiwan']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Paiwan', 'Total', 'speakers', '66', ',', '000', '(', 'est', '.', ')']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\(', 'est', '.', '\\)']], ['NUM', '66'], ['NUM', '000']], 10, True)
```

```
['Paiwan', 'Language', 'codes', 'ISO', '639-3', 'pwn']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yugh', 'Spoken', 'in', 'Russia']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Yugh', 'Region', 'Yenisei', 'River']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Yugh', 'Pronunciation', '\\xc9', '\\x9f', 'uk']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xc9', '\\x9f', 'uk']], 5, True)
```

```
['Yugh', 'Language', 'codes', 'ISO', '639-3', 'yuu']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chickasaw', 'Spoken', 'in', 'United', 'States']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Chickasaw', 'Region', 'South', 'central', 'Oklahoma']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Chickasaw', 'Language', 'codes', 'ISO', '639-3', 'cic']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cypriot', 'Spoken', 'in', 'Cyprus']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Cypriot', 'Language', 'codes', 'ISO', '639-3', 'acy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karakalpak', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karakalpak', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karakalpak', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karakalpak', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karakalpak', 'Region', 'Karakalpakstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Karakalpak', 'Official', 'language', 'in', 'Uzbekistan', '(',  
'Karakalpakstan', 'is', 'the', 'official', 'area', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Uzbekistan', '\\\\(', 'Karakalpakstan', 'is', 'the', 'official',  
'area', '\\\\)']]], 12, True)

['Karakalpak', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Karakalpak', 'Language', 'codes', 'ISO', '639-2', 'kaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Karakalpak', 'Language', 'codes', 'ISO', '639-3', 'kaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gan', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Gan', 'Region', 'central', 'and', 'northern', 'Jiangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central',  
'and', 'northern', 'Jiangxi']]], 6, True)

['Gan', 'Region', 'eastern', 'Hunan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gan', 'Region', 'parts', 'of', 'Fujian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Gan', 'Region', 'Anhui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Gan', 'Region', 'Hubei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Gan', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Gan', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Gan', 'Language', 'codes', 'ISO', '639-3', 'gan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gan', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Tunisian', 'Spoken', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tunisian', 'Writing', 'system', 'Arabic', 'alphabet', '-', 'Latin',  
'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Tunisian', 'Official', 'language', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Tunisian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tunisian', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tunisian', 'Language', 'codes', 'ISO', '639-3', 'aeb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eteocretan', 'Spoken', 'in', 'Crete']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eteocretan', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eteocretan', 'Language', 'codes', 'ISO', '639-3', ':', 'omn', '\xe2',  
'\x80', '\x93', 'Minoan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'omn', '\xe2', '\x80', '\x93', 'Minoan']], ['NUM',  
'639-3']], 11, True)

['Eteocretan', 'Language', 'codes', 'ISO', '639-3', 'ecr', '\xe2', '\x80',  
'\x93', 'Eteocretan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ecr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Kumyk', 'Region', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kumyk', 'Region', 'Chechnya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kumyk', 'Region', 'North', 'Ossetia']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Kumyk', 'Region', 'Turkey']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Kumyk', 'Spoken', 'in', 'Russia']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Kumyk', 'Writing', 'system', 'Cyrillic', 'alphabet']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kumyk', 'Language', 'codes', 'ISO', '639-2', 'kum']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Kumyk', 'Language', 'codes', 'ISO', '639-3', 'kum']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tewa', 'Total', 'speakers', '1300']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tewa', 'Spoken', 'in', 'United', 'States']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Tewa', 'Region', 'New', 'Mexico', 'Ethnicity']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Tewa', 'Region', 'Tewa', 'people']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tewa', 'Language', 'codes', 'ISO', '639-3', 'tew']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Halkomelem', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']]], 7, True)
```

['Halkomelem', 'Language', 'codes', 'ISO', '639-2', 'sal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Halkomelem', 'Language', 'codes', 'ISO', '639-3', 'hur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tawang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tawang', 'Region', 'Lhoka', ',', 'Tibet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Tawang', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Tawang', 'Language', 'codes', 'ISO', '639-3', 'twm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oneida', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']], 7, True)

['Oneida', 'Region', 'Six', 'Nations', 'Reserve']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Oneida', 'Region', 'Wisconsin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Oneida', 'Language', 'codes', 'ISO', '639-3', 'one']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Võro', 'Dialects', 'Seto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Võro', 'Region', 'Southern', 'Estonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['V\xc3\xb5ro', 'Spoken', 'in', 'Estonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['V\xc3\xb5ro', 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['V\xc3\xb5ro', 'Regulated', 'by', 'V', '\xc3', '\xb5', 'ro', 'Institute', '(', 'semi-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'V', '\xc3', '\xb5', 'ro', 'Institute', '\\(', 'semi-official', '\\)']]], 11, True)

['V\xc3\xb5ro', 'Language', 'codes', 'ISO', '639-3', 'vro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Supyire', 'Spoken', 'in', 'Mali']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Supyire', 'Region', 'Sikasso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Supyire', 'Language', 'codes', 'ISO', '639-3', 'spp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sucite', 'Spoken', 'in', 'Mali']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sucite', 'Region', 'Sikasso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sucite', 'Language', 'codes', 'ISO', '639-3', 'sep']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Newfoundland', 'Spoken', 'in', 'Canada']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Newfoundland', 'Region', 'Port', 'au', 'Port', 'Peninsula', ',', 'Newfoundland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Port', 'au', 'Port', 'Peninsula', ',', 'Type: ANY']]], 8, True)

['Newfoundland', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Jutlandic', 'Total', 'speakers', 'unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jutlandic', 'Spoken', 'in', 'Denmark']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jutlandic', 'Language', 'codes', 'ISO', '639-2', 'gem']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Jutlandic', 'Language', 'codes', 'ISO', '639-3', 'jut']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chaldean', 'Region', 'Iraq', ';', 'Mosul']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chaldean', 'Region', 'Ninawa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chaldean', 'Spoken', 'in', 'Assyrian', 'diaspora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chaldean', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chaldean', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chaldean', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chaldean', 'Pronunciation', 'kal', '\xcb', '\x88', 'd', '\xc9', '\x91', 'j', '\xc9', '\x91', ',', 'sor', '\xc9', '\x9b', '\xce', '\xb8']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kal', '\xcb', '\x88', 'd', '\xc9', '\x91', 'j', '\xc9', '\x91', ',', 'sor', '\xc9', '\x9b', '\xce', '\xb8']]], 17, True)

['Chaldean', 'Language', 'codes', 'ISO', '639-2', 'syr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Chaldean', 'Language', 'codes', 'ISO', '639-3', 'cld']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Assyrian', 'Spoken', 'in', 'Assyrian', 'diaspora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Assyrian', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Assyrian', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Assyrian', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Assyrian', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Assyrian', 'Region', 'Hakkari', 'Turkey', ',', 'Urmia', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Hakkari', 'Turkey', ',', 'Urmia', 'Iran']]], 7, True)

['Assyrian', 'Language', 'codes', 'ISO', '639-2', 'syr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Assyrian', 'Language', 'codes', 'ISO', '639-3', 'aii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yurats', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yurats', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yurats', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Senaya', 'Total', 'speakers', '500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Senaya', 'Region', 'Tehran', 'and', 'Qazvin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Senaya', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Senaya', 'Spoken', 'in', 'western', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Senaya', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Senaya', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Senaya', 'Pronunciation', 's', '\xc9', '\x9b', 'n', '\xc9', '\x91', 'j', '\xc9', '\x91', ',', 'soraj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's',

'\xc9', '\x9b', 'n', '\xc9', '\x91', 'j', '\xc9', '\x91', ',', 'soraj']]],  
13, True)

['Senaya', 'Language', 'codes', 'ISO', '639-3', 'syn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minyaka', 'Spoken', 'in', 'Mali']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Minyaka', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Minyaka', 'Language', 'codes', 'ISO', '639-3', 'myk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaska', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kaska', 'Total', 'speakers', '450']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kaska', 'Language', 'codes', 'ISO', '639-3', 'kkz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dagbani', 'Spoken', 'in', 'Ghana']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Dagbani', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Dagbani', 'Language', 'codes', 'ISO', '639-3', 'dag']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Homshetsi', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Homshetsi', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Homshetsi', 'Spoken', 'in', 'Georgia', '(', 'Abkhazia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Homshetsi', 'Spoken', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Homshetsi', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']], 6, True)

['Homshetsi', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Homshetsi', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Surzhyk', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Surzhyk', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Surzhyk', 'Spoken', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Surzhyk', 'Total', 'speakers', 'spread', 'in', 'the', 'whole', 'population']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'spread', 'in', 'the', 'whole', 'population']], 8, True)

['Surzhyk', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Lishana', 'Spoken', 'in', 'Israel', ',', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Lishana', 'Pronunciation', 'li', '\xcb', '\x88', '\xca', '\x83', '\xc9', '\x91', 'n', '\xc9', '\x91', '\xcb', '\x88', 'd', '\xc9', '\x9b', 'ni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'li', '\xcb', '\x88', '\xca', '\x83', '\xc9', '\x91', 'n', '\xc9', '\x91', '\xcb', '\x88', 'd', '\xc9', '\x9b', 'ni']]], 18, True)

['Lishana', 'Language', 'codes', 'ISO', '639-3', 'lsd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hiri', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Hiri', 'Language', 'codes', 'ISO', '639-2', 'hmo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Hiri', 'Language', 'codes', 'ISO', '639-3', 'hmo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hiri', 'Language', 'codes', 'ISO', '639-1', 'ho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Lish\xc3\xa1n', 'Region', 'Jerusalem', 'and', 'Tel', 'Aviv', ',', 'originally', 'from', 'Iranian', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jerusalem', 'and', 'Tel', 'Aviv', ',', 'originally', 'from', 'Iranian', 'Azerbaijan']]], 11, True)

['Lish\xc3\xa1n', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lish\xc3\xa1n', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lish\xc3\xa1n', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Lish\xc3\xa1n', 'Spoken', 'in', 'originally', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lish\xc3\xa1n', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lish\xc3\xa1n', 'Pronunciation', 'li', '\xcb', '\x88', '\xca', '\x83', '\xc9', '\x91', 'n', 'di', '\xcb', '\x88', 'd', '\xc9', '\x91', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'li', '\xcb', '\x88', '\xca', '\x83', '\xc9', '\x91', 'n', 'di', '\xcb', '\x88', 'd', '\xc9', '\x91', 'n']]], 17, True)

['Lish\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'trg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lishanid', 'Pronunciation', 'li', '\xca', '\x83', '\xc9', '\x91', '\xcb', '\x88', 'nid', 'no', '\xcb', '\x88', '\xca', '\x83', '\xc9', '\x91', '\xcc', '\x83']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'li', '\xca', '\x83', '\xc9', '\x91', '\xcb', '\x88', 'nid', 'no', '\xcb', '\x88', '\xca', '\x83', '\xc9', '\x91', '\xcc', '\x83']]], 19, True)

['Lishanid', 'Spoken', 'in', 'Jerusalem', ',', 'originally', 'from', 'eastern', 'and', 'southern', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Jerusalem', ',', 'originally', 'from', 'eastern', 'and', 'southern', 'Iraq']]], 11, True)

['Lishanid', 'Language', 'codes', 'ISO', '639-3', 'aij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karachay-Balkar', 'Region', 'Kabardino-Balkaria', ',', 'Karachay-Cherkessia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Karachay-Balkar', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Karachay-Balkar', 'Writing', 'system', 'Cyrillic', 'alphabet', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Karachay-Balkar', 'Language', 'codes', 'ISO', '639-2', 'krc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Karachay-Balkar', 'Language', 'codes', 'ISO', '639-3', 'krc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Saramaccan', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saramaccan', 'Spoken', 'in', 'French', 'Guiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Saramaccan', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Saramaccan', 'Language', 'codes', 'ISO', '639-3', 'srm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Saramaccan', 'Language', 'codes', 'Linguasphere', '52-ABB-ax']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Patois', 'Spoken', 'in', 'Jamaica', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Patois', 'Spoken', 'in', '665']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Patois', 'Spoken', 'in', '636']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Patois', 'Language', 'codes', 'ISO', '639-3', 'jam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Patois', 'Language', 'codes', 'Linguasphere', '52-ABB-am']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]]], 5, True)

['Median', 'Spoken', 'in', 'ancient', 'Iran']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Median', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]]], 6, True)

['Median', 'Language', 'codes', 'ISO', '639-3', 'xme']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mon\xc3\xa9gasque', 'Spoken', 'in', 'Monaco']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Mon\xc3\xa9gasque', 'Language', 'codes', 'ISO', '639-3', 'lij']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mon\xc3\xa9gasque', 'Language', 'codes', 'Linguasphere', '51-AAA-cha']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]]], 5, True)

['Hulaul\xc3\xa1', 'Region', 'Israel', ',', 'originally', 'form',  
'Iranian', 'Kurdistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Israel', ',',  
'originally', 'form', 'Iranian', 'Kurdistan']]]], 8, True)

['Hulaul\xc3\xa1', 'Spoken', 'in', 'Israel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Hulaul\xc3\xa1', 'Spoken', 'in', 'Iran']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Hulaul\xc3\xa1', 'Spoken', 'in', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Hulaul\xc3\xa1', 'Pronunciation', '\xcb', '\x8c', 'hula', '\xca', '\x94', 'u', '\xcb', '\x88', 'l', '\xc9', '\x91']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x8c', 'hula', '\xca', '\x94', 'u', '\xcb', '\x88', 'l', '\xc9', '\x91']]], 13, True)

['Hulaul\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'huy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Celtiberian', 'Spoken', 'in', 'Spain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Celtiberian', 'Language', 'codes', 'ISO', '639-3', 'xce']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dazaga', 'Spoken', 'in', 'Chad', ',', 'Niger']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Dazaga', 'Region', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dazaga', 'Region', 'BET']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dazaga', 'Region', 'eastern', 'Niger']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dazaga', 'Language', 'codes', 'ISO', '639-3', 'dzg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khitan', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khitan', 'Region', 'northern']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northern']]], 3, True)

['Khitan', 'Writing', 'system', 'Khitan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Khitan', 'Language', 'codes', 'ISO', '639-3', 'zkt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Jurchen', 'Writing', 'system', 'Jurchen', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Jurchen', 'Language', 'codes', 'ISO', '639-3', 'juc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Miao', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miao', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miao', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miao', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miao', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Miao', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Miao', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miao', 'Writing', 'system', 'Pollard', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Miao', 'Writing', 'system', 'RPA', '(', 'USA', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Miao', 'Language', 'codes', 'ISO', '639-2', 'hmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Miao', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Miao', 'Language', 'codes', 'ISO', '639-3', 'hnj', '\\xe2', '\\x80', '\\x93', ']']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hnj', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)

['Bukusu', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bukusu', 'Region', 'East', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bukusu', 'Language', 'codes', 'ISO', '639-3', 'bxk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hakuchi', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakuchi', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Hakuchi', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Nobiin', 'Spoken', 'in', 'Egypt', ', ', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Nobiin,', 'Region', 'Along', 'the', 'banks', 'of', 'the', 'Nile', 'in', 'southern', 'Egypt', 'and', 'northern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Along', 'the', 'banks', 'of', 'the', 'Nile', 'in', 'southern', 'Egypt', 'and', 'northern', 'Sudan']]], 14, True)

['Nobiin,', 'Language', 'codes', 'ISO', '639-2', 'nub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Nobiin,', 'Language', 'codes', 'ISO', '639-3', 'fia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Rioplatense', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rioplatense', 'Spoken', 'in', 'Uruguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rioplatense', 'Total', 'speakers', '20', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '20']], 5, True)

['Rioplatense', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Rioplatense', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Ibibio', 'Dialects', 'Ibibio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ibibio', 'Dialects', 'Efik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ibibio', 'Dialects', 'Anaang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ibibio', 'Dialects', 'Ukwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ibibio', 'Region', 'Akwa', 'Ibom', 'State', ',', 'Cross', 'River', 'State', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Akwa', 'Ibom', 'State', ',', 'Cross', 'River', 'State', 'Ethnicity']], 10, True)

['Ibibio', 'Region', 'Annang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ibibio', 'Region', 'Efik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ibibio', 'Region', 'Ibibio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ibibio', 'Region', 'Oron']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ibibio', 'Spoken', 'in', 'Southern', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ibibio', 'Standard', 'forms', 'Efik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ibibio', 'Language', 'codes', 'ISO', '639-2', 'efi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ibibio', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ibibio', 'Language', 'codes', 'ISO', '639-3', ':', 'anw', '\xe2', '\x80', '\x93', 'Anaang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'anw', '\xe2', '\x80', '\x93', 'Anaang']], ['NUM', '639-3']], 11, True)

['Ibibio', 'Language', 'codes', 'ISO', '639-3', 'efi', '\xe2', '\x80', '\x93', 'Efik']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'efi', '\xe2', '\x80', '\x93', 'Efik']], ['NUM', '639-3']], 10, True)

['Ibibio', 'Language', 'codes', 'ISO', '639-3', 'ukq', '\xe2', '\x80', '\x93', 'Ukwa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ukq', '\xe2', '\x80', '\x93', 'Ukwa']], ['NUM', '639-3']], 10, True)

['Iraqi', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iraqi', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iraqi', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iraqi', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iraqi', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Iraqi', 'Official', 'language', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Iraqi', 'Language', 'codes', 'ISO', '639-3', 'acm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Saek', 'Spoken', 'in', 'Laos', ',', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Saek', 'Language', 'codes', 'ISO', '639-3', 'skb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Modern', 'Spoken', 'in', 'Arab', 'world']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Modern', 'Total', 'speakers', 'ca', '.', '300', 'million', ';', 'up',  
'to', '1', 'billion', '(', 'mainly', 'Muslims', 'around', 'the', 'globe',  
)', 'with', 'a', 'significant', 'or', 'working', 'knowledge', 'of', 'the',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'ca', '.', 'NUM', 'million', ';', 'up', 'to', 'NUM', 'billion', '\\\\(',  
'mainly', 'Muslims', 'around', 'the', 'globe', '\\\\)', 'with', 'a',  
'significant', 'or', 'working', 'knowledge', 'of', 'the', 'language']]],  
['NUM', '300'], ['NUM', '1']], 28, True)
```

```
['Modern', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Modern', 'Official', 'language', 'in', 'Arab', 'world']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Modern', 'Official', 'language', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Modern', 'Official', 'language', 'in', 'Djibouti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Modern', 'Official', 'language', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Modern', 'Official', 'language', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```



```
['Modern', 'Official', 'language', 'in', 'Somalia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]]], 5, True)
```

```
['Modern', 'Official', 'language', 'in', 'Comoros']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]]], 5, True)
```

```
['Modern', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Modern', 'Language', 'codes', 'ISO', '639-3', 'arb']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Modern', 'Language', 'codes', 'ISO', '639-1', 'ar']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Urum', 'Spoken', 'in', 'Ukraine', ',', 'Greece']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]]], 6, True)
```

```
['Urum', 'Writing', 'system', 'Cyrillic', 'alphabet', ',', 'Greek',
'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]]], 8, True)
```

```
['Urum', 'Pronunciation', 'u', '\xcb', '\x88', 'rum']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'u',
'\xcb', '\x88', 'rum']]]], 6, True)
```

```
['Urum', 'Language', 'codes', 'ISO', '639-3', 'uum']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'Isle', 'of', 'Man']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle',
'of', 'Man']]]], 6, True)
```

```
['Old', 'Spoken', 'in', 'western', 'coast', 'of', 'Great', 'Britain']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'western', 'coast', 'of', 'Great', 'Britain']]]], 8, True)
```

```
['Old', 'Writing', 'system', 'Latin']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Pronunciation', '\xcb', '\x88', '\xc9', '\xa1', 'oi', '\xc3', '\xb0', '\xca', '\xb2', 'el', '\xc9', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', '\xc9', '\xa1', 'oi', '\xc3', '\xb0', '\xca', '\xb2', 'el', '\xc9', '\xa1']]], 14, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'sga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'sga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Primitive', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Primitive', 'Spoken', 'in', 'Isle', 'of', 'Man']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle', 'of', 'Man']]], 6, True)

['Primitive', 'Spoken', 'in', 'western', 'coast', 'of', 'Great', 'Britain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'western', 'coast', 'of', 'Great', 'Britain']]], 8, True)

['Primitive', 'Writing', 'system', 'Ogham']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Primitive', 'Language', 'codes', 'ISO', '639-3', 'pgl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Niellim', 'Spoken', 'in', 'southwestern', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Niellim', 'Language', 'codes', 'ISO', '639-3', 'nie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Achumawi', 'Spoken', 'in', 'California']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Achumawi', 'Total', 'speakers', '8']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Achumawi', 'Language', 'codes', 'ISO', '639-3', 'acv']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Atsugewi', 'Spoken', 'in', 'California']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Atsugewi', 'Total', 'speakers', 'Extinct']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Atsugewi', 'Language', 'codes', 'ISO', '639-3', 'atw']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ahtna', 'Total', 'speakers', '80']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ahtna', 'Region', 'Alaska', '(', 'Copper', 'River', 'region', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\
(', 'Copper', 'River', 'region', '\\)']]], 8, True)
```

```
['Ahtna', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Ahtna', 'Writing', 'system', 'Latin', '(', 'Ahtna', 'variant', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Ahtna', 'Language', 'codes', 'ISO', '639-3', 'aht']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nanerige', 'Spoken', 'in', 'Burkina', 'Faso']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'Type: ANY']]], 5, True)

['Nanerige', 'Language', 'codes', 'ISO', '639-3', 'sen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Palaka', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palaka', 'Region', 'Sikasso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Palaka', 'Language', 'codes', 'ISO', '639-3', 'plr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gula', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gula', 'Language', 'codes', 'ISO', '639-3', 'glj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buryat', 'Spoken', 'in', 'Buryat', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Buryat', 'Spoken', 'in', 'northern', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Buryat', 'Spoken', 'in', 'Ust-Orda', 'Buryatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Buryat', 'Spoken', 'in', 'Aga', 'Buryatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Buryat', 'Language', 'codes', 'ISO', '639-2', 'bua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Buryat', 'Language', 'codes', 'ISO', '639-3', 'bua', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bua', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Buryat', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Buryat', 'Language', 'codes', 'ISO', '639-3', ':', 'bxu', '\xe2', '\x80', '\x93', 'China', 'Buriat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'bxu', '\xe2', '\x80', '\x93', 'China', 'Buriat']], ['NUM', '639-3']], 12, True)

['Buryat', 'Language', 'codes', 'ISO', '639-3', 'bxm', '\xe2', '\x80', '\x93', 'Mongolia', 'Buriat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bxm', '\xe2', '\x80', '\x93', 'Mongolia', 'Buriat']], ['NUM', '639-3']], 11, True)

['Buryat', 'Language', 'codes', 'ISO', '639-3', 'bxr', '\xe2', '\x80', '\x93', 'Russia', 'Buriat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bxr', '\xe2', '\x80', '\x93', 'Russia', 'Buriat']], ['NUM', '639-3']], 11, True)

['Bua', 'Spoken', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bua', 'Language', 'codes', 'ISO', '639-3', 'bub']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hmar', 'Region', 'Southeast', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hmar', 'Spoken', 'in', 'Mizoram']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hmar', 'Spoken', 'in', 'Manipur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Hmar', 'Spoken', 'in', 'and', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hmar', 'Spoken', 'in', 'India', ';', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
';', 'Type: ANY']]], 6, True)

['Hmar', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Hmar', 'Pronunciation', 'm', '\xcc', '\xa5', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm',  
'\xcc', '\xa5', 'a']]], 6, True)

['Hmar', 'Official', 'language', 'in', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Hmar', 'Official', 'language', 'in', 'recognized', 'as', 'a', 'minority',  
'language', 'in', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'recognized', 'as', 'a', 'minority', 'language', 'in', 'Assam']]],  
11, True)

['Hmar', 'Official', 'language', 'in', 'Manipur', 'and', 'Mizoram']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Manipur', 'and', 'Mizoram']]], 7, True)

['Hmar', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Hmar', 'Language', 'codes', 'ISO', '639-3', 'hmr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'amw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marquesan', 'Spoken', 'in', 'Marquesas', 'Islands', ',', 'Tahiti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Marquesas', 'Islands', ',', 'Tahiti']]], 7, True)

['Marquesan', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rapa', 'Region', 'Easter', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rapa', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rapa', 'Language', 'codes', 'ISO', '639-2', 'rap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Rapa', 'Language', 'codes', 'ISO', '639-3', 'rap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rapan', 'Spoken', 'in', 'French', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Rapan', 'Total', 'speakers', '521']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rapan', 'Language', 'codes', 'ISO', '639-3', 'ray']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dakota', 'Region', 'but', 'also', 'northern', 'Nebraska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'but', 'also',  
'northern', 'Nebraska']]], 6, True)

['Dakota', 'Region', 'southern', 'Minnesota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Dakota', 'Spoken', 'in', 'United', 'States', ',', 'with', 'some',  
'speakers', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', ',', 'with', 'some', 'speakers', 'in', 'Canada']]]], 11, True)
```

```
['Dakota', 'Pronunciation', 'da', '\xcb', '\x88', 'k', '\xca', '\xb0',  
'otijapi', ',', 'da', '\xcb', '\x88', 'q', '\xcb', '\xa3', 'otijapi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'da',  
'\xcb', '\x88', 'k', '\xca', '\xb0', 'otijapi', ',', 'da', '\xcb', '\x88',  
'q', '\xcb', '\xa3', 'otijapi']]]], 17, True)
```

```
['Dakota', 'Language', 'codes', 'ISO', '639-2', 'dak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Dakota', 'Language', 'codes', 'ISO', '639-3', 'dak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sioux', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]]], 7, True)
```

```
['Sioux', 'Region', 'Northern', 'Nebraska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Sioux', 'Region', 'southern', 'Minnesota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Sioux', 'Region', 'North', 'Dakota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Sioux', 'Region', 'South', 'Dakota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Sioux', 'Region', 'northeastern', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Sioux', 'Region', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)
```



['Sioux', 'Language', 'codes', 'ISO', '639-2', 'dak', '\xe2', '\x80', '\x94', 'Dakota']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'dak', '\xe2', '\x80', '\x94', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Sioux', 'Language', 'codes', 'ISO', '639-2', 'sio', '\xe2', '\x80', '\x94', 'Siouan', '(', 'collective', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sio', '\xe2', '\x80', '\x94', 'Siouan', '\\(', 'collective', '\\)']], ['NUM', '639-2']], 13, True)

['Sioux', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sioux', 'Language', 'codes', 'ISO', '639-3', ':', 'dak', '\xe2', '\x80', '\x93', 'Dakota']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'dak', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Sioux', 'Language', 'codes', 'ISO', '639-3', 'lkt', '\xe2', '\x80', '\x93', 'Lakota']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lkt', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Pukapukan', 'Language', 'codes', 'ISO', '639-2', 'map']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pukapukan', 'Language', 'codes', 'ISO', '639-3', 'pkp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Barzani', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Barzani', 'Region', 'Jerusalem', ',', 'originally', 'from', 'Bijil', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jerusalem', ',', 'originally', 'from', 'Bijil', 'in', 'Iraq']], 9, True)

```
['Barzani', 'Language', 'codes', 'ISO', '639-3', 'bjf']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gusii', 'Total', 'speakers', '1', ',', '5', 'million']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'million']], ['NUM', '1'], ['NUM', '5']], 7, True)
```

```
['Gusii', 'Spoken', 'in', 'Kenya']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Gusii', 'Language', 'codes', 'ISO', '639-3', 'guz']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['H\u03a9rtevin', 'Region', 'Siirt', 'Province']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['H\u03a9rtevin', 'Spoken', 'in', 'Turkey']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['H\u03a9rtevin', 'Pronunciation', '\xcb', '\x88', 'h', '\xc9', '\x9b',
'rt', '\xc9', '\x99', 'v', '\xc9', '\x99', 'n', ',', '\xcb', '\x88', 'sor',
'\xc9', '\x9b', '\xce', '\xb8']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',
'\x88', 'h', '\xc9', '\x9b', 'rt', '\xc9', '\x99', 'v', '\xc9', '\x99',
'n', ',', '\xcb', '\x88', 'sor', '\xc9', '\x9b', '\xce', '\xb8']], 22,
True)
```

```
['H\u03a9rtevin', 'Language', 'codes', 'ISO', '639-3', 'hrt']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Koy', 'Spoken', 'in', 'Iraq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Koy', 'Language', 'codes', 'ISO', '639-3', 'kqd']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Luo', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', ',', 'Type: ANY']]], 6, True)

['Luo', 'Language', 'codes', 'ISO', '639-2', 'luo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Luo', 'Language', 'codes', 'ISO', '639-3', 'luo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adangme', 'Spoken', 'in', 'South-eastern', 'Ghana', ',', 'east', 'of', 'Accra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'South-eastern', 'Ghana', ',', 'east', 'of', 'Accra']]], 9, True)

['Adangme', 'Official', 'language', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Adangme', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Adangme', 'Language', 'codes', 'ISO', '639-2', 'ada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Adangme', 'Language', 'codes', 'ISO', '639-3', 'ada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Thai', 'Total', 'speakers', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', 'NUM', 'million']], ['NUM', '5']], 6, True)

['Southern', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Region', 'Southern', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'tai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'sou']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ho', 'Spoken', 'in', 'India', ',', 'Bangladesh']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Ho', 'Writing', 'system', 'Devanagari', ',', 'Varang', 'Kshiti']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devanagari', ',', 'Varang', 'Kshiti']], 7, True)

['Ho', 'Language', 'codes', 'ISO', '639-3', 'hoc']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Twi/Fante', 'Spoken', 'in', 'Ghana', ',', 'Suriname']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Twi/Fante', 'Total', 'speakers', '20', 'million']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '20']], 5, True)

['Twi/Fante', 'Official', 'language', 'in', 'None']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Twi/Fante', 'Official', 'language', 'in', '\xe2', '\x80', '\x94', 'Government-sponsored', 'languages', 'of', 'Ghana']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', '\xe2', '\x80', '\x94', 'Government-sponsored', 'languages', 'of', 'Ghana']], 11, True)

['Twi/Fante', 'Regulated', 'by', 'Akan', 'Orthography', 'Committee']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Akan', 'Orthography', 'Committee']], 6, True)

['Twi/Fante', 'Language', 'codes', 'ISO', '639-2', 'aka']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Twi/Fante', 'Language', 'codes', 'ISO', '639-3', 'aka', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'aka', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Twi/Fante', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Twi/Fante', 'Language', 'codes', 'ISO', '639-3', ':', 'fat', '\xe2', '\x80', '\x93', 'Fanti', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'fat', '\xe2', '\x80', '\x93', 'Fanti', 'language']], ['NUM', '639-3']], 12, True)

['Twi/Fante', 'Language', 'codes', 'ISO', '639-3', 'twi', '\xe2', '\x80', '\x93', 'Twi', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'twi', '\xe2', '\x80', '\x93', 'Twi', 'language']], ['NUM', '639-3']], 11, True)

['Twi/Fante', 'Language', 'codes', 'ISO', '639-1', 'ak']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Timbisha', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Timbisha', 'Region', 'California', ',', 'Nevada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Timbisha', 'Language', 'codes', 'ISO', '639-3', 'par']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kipchak', 'Total', 'speakers', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kipchak', 'Spoken', 'in']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']], 3, True)

['Kipchak', 'Region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']], 2, True)

['Kipchak', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Gronings', 'Total', 'speakers', '320', ',', '000', '(', 'est', '.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'est', '.', '\\)']], ['NUM', '320'], ['NUM',  
'000']], 10, True)

['Gronings', 'Spoken', 'in', 'Netherlands', ':', 'Groningen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Netherlands', ':', 'Groningen']], 6, True)

['Gronings', 'Spoken', 'in', 'parts', 'in', 'the', 'north', 'and', 'east',  
'of', 'Drenthe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',  
'in', 'the', 'north', 'and', 'east', 'of', 'Drenthe']], 11, True)

['Gronings', 'Region', 'Groningen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Gronings', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Gronings', 'Language', 'codes', 'ISO', '639-3', 'gos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koryak', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Koryak', 'Region', 'Koryak', 'Okrug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Koryak', 'Language', 'codes', 'ISO', '639-3', 'kpy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mbum', 'Dialects', 'Gbete']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Mbum', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Mbum', 'Language', 'codes', 'ISO', '639-3', 'mdd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hawaii', 'Region', 'limited', 'to', 'the', 'islands', 'of', 'Hawaii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'limited',  
'to', 'the', 'islands', 'of', 'Type: ANY']], 8, True)

['Hawaii', 'Signed', 'in', 'Hawaii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Hawaii', 'Total', 'signers', 'virtually', 'extinct', ';', 'a', 'few',  
'elderly', 'signers', 'are', 'bilingual', 'with', 'the', 'dominant',  
'ASL']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'virtually', 'extinct', ';', 'a', 'few', 'elderly', 'signers', 'are',  
'bilingual', 'with', 'the', 'dominant', 'ASL']], 16, True)

['Hawaii', 'Language', 'codes', 'ISO', '639-2', 'sgn-US-HI']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mbabaram', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Mbabaram', 'Language', 'codes', 'ISO', '639-3', 'vmb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yanyuwa', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Yanyuwa', 'Language', 'codes', 'ISO', '639-3', 'jao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Muslim', 'Tat', 'Total', 'speakers', '26', ',', '000', 'excluding',  
'Judeo-Tat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Tat', 'Total',

'speakers', 'NUM', ',', 'NUM', 'excluding', 'Judeo-Tat']], ['NUM', '26'],  
['NUM', '000']], 9, True)

['Muslim', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Muslim', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Muslim', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Muslim', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Muslim', 'Region', 'Northeastern', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Muslim', 'Language', 'codes', 'ISO', '639-3', 'ttt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Muslim', 'Language', 'codes', 'Linguasphere', '58-AAC-g']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Miami-Illinois', 'Spoken', 'in', 'Illinois']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miami-Illinois', 'Spoken', 'in', 'Indiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miami-Illinois', 'Spoken', 'in', 'Kansas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Miami-Illinois', 'Spoken', 'in', 'Michigan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



['Miami-Illinois', 'Spoken', 'in', 'Ohio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Miami-Illinois', 'Spoken', 'in', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Miami-Illinois', 'Language', 'codes', 'ISO', '639-2', 'alg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Miami-Illinois', 'Language', 'codes', 'ISO', '639-3', 'mia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tugen', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tugen', 'Region', 'West', 'Central', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tugen', 'Language', 'codes', 'ISO', '639-3', 'tuy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kobon', 'Region', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kobon', 'Region', 'Middle', 'Ramu', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kobon', 'Language', 'codes', 'ISO', '639-3', 'kpw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['P\xc3\xb6koot', 'Spoken', 'in', 'Kenya', ',', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['P\xc3\xb6koot', 'Language', 'codes', 'ISO', '639-3', 'pko']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ogiek', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Ogiek', 'Region', 'Akiek', ':', 'Tanzania', ',', 'southern', 'part', 'of', 'Arusha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Akiek', ':', 'Tanzania', ',', 'southern', 'part', 'of', 'Arusha']], 10, True)

['Ogiek', 'Region', 'Kinare', ':', 'Kinare']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Ogiek', 'Region', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ogiek', 'Region', 'on', 'the', 'eastern', 'slope', 'of', 'the', 'Rift', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'on', 'the', 'eastern', 'slope', 'of', 'the', 'Rift', 'Valley']], 10, True)

['Ogiek', 'Language', 'codes', 'ISO', '639-3', 'oki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Spoken', 'in', 'Armenian', 'Highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Classical', 'Writing', 'system', 'Armenian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'xcl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Language', 'codes', 'Linguasphere', '57-AAA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Lendu', 'Spoken', 'in', 'Congo', '(', 'DRC', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)

['Lendu', 'Language', 'codes', 'ISO', '639-3', 'led']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koyra', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Koyra', 'Region', 'Niger', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Koyra', 'Language', 'codes', 'ISO', '639-3', 'khq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bactrian', 'Spoken', 'in', 'Bactria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bactrian', 'Writing', 'system', 'Greek', 'alphabet', ',', 'Manichean', 'abjad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Bactrian', 'Language', 'codes', 'ISO', '639-3', 'xbc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Walliser', 'Language', 'codes', 'ISO', '639-2', 'gsw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Walliser', 'Language', 'codes', 'ISO', '639-3', 'gsw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vaeakau-Taumako', 'Spoken', 'in', 'Reef', 'Islands', 'and', 'Taumako', ',', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Reef', 'Islands', 'and', 'Taumako', ',', 'Solomon', 'Islands']], 10, True)

['Vaeakau-Taumako', 'Language', 'codes', 'ISO', '639-3', 'piv']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wiyot', 'Total', 'speakers', 'extinct']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Wiyot', 'Language', 'codes', 'ISO', '639-3', 'wiy']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yurok', 'Language', 'codes', 'ISO', '639-3', 'yur']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Trinidadian', 'Spoken', 'in', 'Trinidad']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Trinidadian', 'Total', 'speakers']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)
```

```
['Trinidadian', 'Language', 'codes', 'ISO', '639-3', 'trf']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Trinidadian', 'Language', 'codes', 'Linguasphere', '52-ABB-au']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Tobagonian', 'Spoken', 'in', 'Tobago']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tobagonian', 'Total', 'speakers']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)
```

```
['Tobagonian', 'Language', 'codes', 'ISO', '639-3', 'tgh']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tobagonian', 'Language', 'codes', 'Linguasphere', '52-ABB-at']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)
```

['Xiang', 'Total', 'speakers', '30-36', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '30-36']], 5, True)

['Xiang', 'Dialects', 'Old', 'Xiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xiang', 'Dialects', 'New', 'Xiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xiang', 'Dialects', 'Chen-Xu', 'Xiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xiang', 'Region', 'Sichuan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Xiang', 'Region', 'and', 'northern', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Xiang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xiang', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Xiang', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Xiang', 'Language', 'codes', 'ISO', '639-3', 'hsn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xiang', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tunica', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tunica', 'Region', 'Central', 'Louisiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tunica', 'Language', 'codes', 'ISO', '639-3', 'tun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Pinghua', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pinghua', 'Region', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Pinghua', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Palenquero', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Palenquero', 'Region', 'South', 'America']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Palenquero', 'Language', 'codes', 'ISO', '639-3', 'pln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Palenquero', 'Language', 'codes', 'Linguasphere', '51-AAC-bc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Klallam', 'Total', 'speakers', '4', 'native', ';', '5', 'linguists']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'native', ';', 'NUM', 'linguists']]], ['NUM', '4'], ['NUM', '5']], 8, True)

['Klallam', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Klallam', 'Region', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Klallam', 'Language', 'codes', 'ISO', '639-3', 'clm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Atikamekw', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Atikamekw', 'Region', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Atikamekw', 'Language', 'codes', 'ISO', '639-3', 'atj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Atikamekw', 'Language', 'codes', 'Linguasphere', '62-ADA-c']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Berta', 'Total', 'speakers', '205', ',', '732', 'in', 'Ethiopia', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Ethiopia', ',', '']]], ['NUM', '205'], ['NUM', '732']], 9, True)

['Berta', 'Total', 'speakers', '22', ',', '000', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '22'], ['NUM', '000']], 8, True)

['Berta', 'Spoken', 'in', 'Sudan', ',', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Berta', 'Region', 'Benishangul-Gumuz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Berta', 'Language', 'codes', 'ISO', '639-3', 'wti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tolomako', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tolomako', 'Language', 'codes', 'ISO', '639-3', 'tlm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kurmanji', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kurmanji', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kurmanji', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kurmanji', 'Language', 'codes', 'ISO', '639-2', 'kur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kurmanji', 'Language', 'codes', 'ISO', '639-3', 'kmr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Macedonian', 'Spoken', 'in', 'Macedon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Macedonian', 'Language', 'codes', 'ISO', '639-3', 'xmk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ch'orti'", 'Spoken', 'in', 'Guatemala', ',', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

["Ch'orti'", 'Region', 'Cop', '\xc3', '\xa1', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Cop', '\xc3', '\xa1', 'n']], 6, True)



```
["Ch'orti'", 'Language', 'codes', 'ISO', '639-3', 'caa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nigerian', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Nigerian', 'Total', 'speakers', '80', 'million', 'in', 'total']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'in', 'total']], ['NUM', '80']], 7, True)
```

```
['Nigerian', 'Language', 'codes', 'ISO', '639-3', 'pcm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Alsean', 'Language', 'extinction', '1940s']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Alsean', 'Spoken', 'in', 'Oregon', 'Ethnicity']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Alsean', 'Spoken', 'in', 'Alsea', 'people', ',', 'Yaquina', 'people']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Alsea',
'people', ',', 'Yaquina', 'people']]], 8, True)
```

```
['Alsean', 'Dialects', 'Alsea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Alsean', 'Dialects', 'Yaquina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Alsean', 'Language', 'codes', 'ISO', '639-3', 'aes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Adai', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Adai', 'Region', 'Louisiana', 'Ethnicity']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Adai', 'Region', 'Adai', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Adai', 'Language', 'codes', 'ISO', '639-3', 'xad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Adai', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'Adai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Pre-contact', 'distribution', 'of', 'Type: ANY']]], ['NUM', '639-3']], 9, True)

['Lusitanian', 'Spoken', 'in', 'Beira', 'Alta', 'and', 'Alto', 'Alentejo', 'Portugal', 'and', 'adjacent', 'areas', 'of', 'Extramadura', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Beira', 'Alta', 'and', 'Alto', 'Alentejo', 'Portugal', 'and', 'adjacent', 'areas', 'of', 'Extramadura', 'Spain']]], 15, True)

['Lusitanian', 'Language', 'codes', 'ISO', '639-3', 'xls']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lisu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lisu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lisu', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lisu', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lisu', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

```

['Lisu', 'Language', 'codes', 'ISO', '639-3', 'lis']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Krio', 'Total', 'speakers', 'ca']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Krio', 'Spoken', 'in', 'Sierra', 'Leone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Krio', 'Region', 'West', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Krio', 'Language', 'codes', 'ISO', '639-3', 'kri']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sekani', 'Total', 'speakers', '25', 'fluent', ',', '61', 'partial']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'fluent', ',', 'NUM', 'partial']], ['NUM', '25'], ['NUM', '61']], 8,
True)

['Sekani', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sekani', 'Region', 'British', 'Columbia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sekani', 'Language', 'codes', 'ISO', '639-3', 'sek']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jud\u03\xa6o-Tat', 'Total', 'speakers', '106', ',', '000', '(', 'est',
', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'est', '.', '\\)']], ['NUM', '106'], ['NUM',
'000']], 10, True)

['Jud\u03\xa6o-Tat', 'Region', 'Caucasus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

```

```
['Jud\xc3\xa6o-Tat', 'Spoken', 'in', 'Azerbaijan', '(', 'Baku', ',', 'Quba', ',', 'Q', '\xc4', '\xb1', 'rm', '\xc4', '\xb1', 'z', '\xc4', '\xb1', 'Q', '\xc9', '\x99', 's', '\xc9', '\x99', 'b', '\xc9', '\x99', ',', 'O', '\xc4', '\x9f', 'uz', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Azerbaijan', '\\(', 'Baku', ',', 'Quba', ',', 'Q', '\xc4', '\xb1', 'rm', '\xc4', '\xb1', 'z', '\xc4', '\xb1', 'Q', '\xc9', '\x99', 's', '\xc9', '\x99', 'b', '\xc9', '\x99', ',', 'O', '\xc4', '\x9f', 'uz', '\\)']]], 33, True)
```

```
['Jud\xc3\xa6o-Tat', 'Spoken', 'in', 'Russia', '(', 'Derbent', ',', 'Makhachkala', ',', 'Nalchik', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Russia', '\\(', 'Derbent', ',', 'Makhachkala', ',', 'Nalchik', '\\)']]], 11, True)
```

```
['Jud\xc3\xa6o-Tat', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Jud\xc3\xa6o-Tat', 'Writing', 'system', 'Roman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Jud\xc3\xa6o-Tat', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Jud\xc3\xa6o-Tat', 'Writing', 'system', 'Hebrew']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Jud\xc3\xa6o-Tat', 'Official', 'language', 'in', 'no', 'official', 'status']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'no', 'official', 'status']]], 7, True)
```

```
['Jud\xc3\xa6o-Tat', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Jud\xc3\xa6o-Tat', 'Language', 'codes', 'ISO', '639-3', 'jdt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zarphatic', 'Spoken', 'in', 'France']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Zarphatic', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Zarphatic', 'Language', 'codes', 'ISO', '639-3', 'zrp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Nitinaht', 'Total', 'speakers', '8']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Nitinaht', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Nitinaht', 'Language', 'codes', 'ISO', '639-2', 'wak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Nitinaht', 'Language', 'codes', 'ISO', '639-3', 'noo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Goemai', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Goemai', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Goemai', 'Language', 'codes', 'ISO', '639-3', 'ank']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Haisla', 'Spoken', 'in', 'Douglas', 'Channel', 'head']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Douglas', 'Channel', 'head']]], 6, True)
```

```
['Haisla', 'Spoken', 'in', 'near', 'Kitimat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

['Haisla', 'Total', 'speakers', '250']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Haisla', 'Dialects', 'Kitamaat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Haisla', 'Dialects', 'Kitlope']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Haisla', 'Language', 'codes', 'ISO', '639-3', 'has']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Heiltsuk', '(', 'Bella', 'Bella', ')', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', '\\(', 'Bella', 'Bella', '\\)', 'Total', 'speakers']], 7, True)

['Heiltsuk', 'Region', 'al', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Heiltsuk', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Heiltsuk', 'Region', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Heiltsuk', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Heiltsuk', 'Language', 'codes', 'ISO', '639-3', 'hei-hai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['She', 'Region', 'Zengcheng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['She', 'Region', 'Boluo', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['She', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['She', 'Pronunciation', 'h', '\xc9', '\x94', '\xcb', '\xa8', 'ne', '\xcb', '\xa5', '\xcb', '\xa7']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'h', '\xc9', '\x94', '\xcb', '\xa8', 'ne', '\xcb', '\xa5', '\xcb', '\xa7']], 12, True)

['She', 'Pronunciation', 'sa', '\xc5', '\x8b', '\xcb', '\xa6', 'xa', '\xca', '\x94', '\xcb', '\xa5', 'u', '\xc9', '\x99', '\xcb', '\xa8', '\xcb', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sa', '\xc5', '\x8b', '\xcb', '\xa6', 'xa', '\xca', '\x94', '\xcb', '\xa5', 'u', '\xc9', '\x99', '\xcb', '\xa8', '\xcb', '\xa9']], 19, True)

['She', 'Language', 'codes', 'ISO', '639-3', 'shx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kayla', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kayla', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kayla', 'Region', 'Amhara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kayla', 'Language', 'codes', 'ISO', '639-3', 'ahg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bukhori', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bukhori', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bukhori', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bukhori', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bukhori', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bukhori', 'Region', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bukhori', 'Language', 'codes', 'ISO', '639-3', 'bhh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Krymchak', 'Spoken', 'in', 'Crimea', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Krymchak', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Krymchak', 'Language', 'codes', 'ISO', '639-3', 'jct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bhili', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhili', 'Region', 'Madhya', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhili', 'Region', 'Gujarat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



ANY']]], 3, True)

['Bhili', 'Region', 'Rajasthan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bhili', 'Region', 'Maharashtra']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bhili', 'Language', 'codes', 'ISO', '639-3', 'bhb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Basay', 'Spoken', 'in', 'Taiwan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Basay', 'Total', 'speakers', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Basay', 'Language', 'codes', 'ISO', '639-3', 'byq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mochica', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']]], 6, True)

['Mochica', 'Language', 'codes', 'ISO', '639-3', 'omc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ithkuil', 'Total', 'speakers', 'presumably', 'none']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'presumably', 'none']]], 5, True)

['Ithkuil', 'Writing', 'system', 'original', '(', 'I', '\xc3', '\xa7', 'ta', '\xc3', '\xae', 'l', ')', ', ', 'with', 'romanized', 'transcription']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'original', '\\(', 'I', '\xc3', '\xa7', 'ta', '\xc3', '\xae', 'l', '\\)', ', ', 'with', 'romanized', 'transcription']]], 17, True)

['Ithkuil', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',

'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Ithkuil', 'Category', '(', 'purpose', ')', 'Ithkuil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Ithkuil', 'Created', 'by', 'John', 'Quijada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ithkuil', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ithkuil', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Temne', 'Official', 'language', 'in', 'Sierra', 'Leone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Temne', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Temne', 'Language', 'codes', 'ISO', '639-2', 'tem']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Temne', 'Language', 'codes', 'ISO', '639-3', 'tem']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fula', 'Spoken', 'in', 'Mauritania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Fula', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Gambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Sierra', 'Leone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Fula', 'Spoken', 'in', 'Benin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Guinea-Bissau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Central', 'African', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY', 'Republic']]], 6, True)

['Fula', 'Spoken', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Togo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Liberia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Eritrea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Spoken', 'in', 'Gabon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fula', 'Total', 'speakers', '40', '\xe2', '\x80', '\x93', '65', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']]], ['NUM', '40'], ['NUM', '65']], 9, True)

['Fula', 'Language', 'codes', 'ISO', '639-2', 'ful']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Fula', 'Language', 'codes', 'ISO', '639-3', 'ful', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ful', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)

['Fula', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

```
['Fula', 'Language', 'codes', 'ISO', '639-3', ':', 'fub', '\xe2', '\x80',  
\x93', 'Adamawa', 'Fulfulde']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'fub', '\xe2', '\x80', '\x93', 'Adamawa', 'Fulfulde']],  
['NUM', '639-3']], 12, True)
```

```
['Fula', 'Language', 'codes', 'ISO', '639-3', 'fue', '\xe2', '\x80',  
\x93', 'Borgu', 'Fulfulde']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'fue', '\xe2', '\x80', '\x93', 'Borgu', 'Fulfulde']], ['NUM',  
'639-3']], 11, True)
```

```
['Fula', 'Language', 'codes', 'ISO', '639-3', 'fuq', '\xe2', '\x80',  
\x93', 'Central-Eastern', 'Niger', 'Fulfulde']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'fuq', '\xe2', '\x80', '\x93', 'Central-Eastern', 'Niger',  
'Fulfulde']], ['NUM', '639-3']], 12, True)
```

```
['Fula', 'Language', 'codes', 'ISO', '639-3', 'fuc', '\xe2', '\x80',  
\x93', 'Pulaar']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'fuc', '\xe2', '\x80', '\x93', 'Pulaar']], ['NUM', '639-3']],  
10, True)
```

```
['Fula', 'Language', 'codes', 'ISO', '639-1', 'ff']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Kinnauri', 'Spoken', 'in', 'Himachal', 'Pradesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Kinnauri', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']], 6, True)
```

```
['Kinnauri', 'Dialects', 'Sunam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kinnauri', 'Language', 'codes', 'ISO', '639-3', 'kfk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bulloom', 'Language', 'codes', 'ISO', '639-3', 'buy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bom', 'Total', 'speakers', '250']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bom', 'Spoken', 'in', 'Sierra', 'Leone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Bom', 'Region', 'West', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bom', 'Language', 'codes', 'ISO', '639-3', 'bmf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Krim', 'Spoken', 'in', 'Sierra', 'Leone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Krim', 'Language', 'codes', 'ISO', '639-3', 'krm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sherbro', 'Spoken', 'in', 'Sierra', 'Leone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Sherbro', 'Language', 'codes', 'ISO', '639-3', 'bun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karon', 'Spoken', 'in', 'Senegal', ',', 'Gambia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Karon', 'Region', 'Southwest', 'Senegal', 'coast']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Karon', 'Language', 'codes', 'ISO', '639-3', 'krx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Mlomp', 'Spoken', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mlomp', 'Region', 'Southwest', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mlomp', 'Language', 'codes', 'ISO', '639-3', 'mlo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mende', 'Language', 'codes', 'ISO', '639-2', 'men']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Mende', 'Language', 'codes', 'ISO', '639-3', 'men']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tedaga', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tedaga', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tedaga', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tedaga', 'Spoken', 'in', 'Libya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tedaga', 'Region', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tedaga', 'Region', 'BET']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tedaga', 'Region', 'Eastern', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Tedaga', 'Language', 'codes', 'ISO', '639-3', 'tuq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Volscian', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Volscian', 'Spoken', 'in', 'Latium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Volscian', 'Region', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Volscian', 'Writing', 'system', 'Old', 'Italic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Volscian', 'Language', 'codes', 'ISO', '639-3', 'xvo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Volscian', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of', 'the',  
'Volsci']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Location', 'of', 'the', 'Volsci']], ['NUM', '639-3']], 9,  
True)

['Kuy', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kuy', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kuy', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kuy', 'Total', 'speakers', '490', ',', '000', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',



'NUM', ',', 'NUM', 'speakers']], ['NUM', '490'], ['NUM', '000']], 7, True)

['Kuy', 'Language', 'codes', 'ISO', '639-3', 'kdt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fala', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Fala', 'Region', 'autonomous', 'community', 'of', 'Extremadura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Fala', 'Language', 'codes', 'ISO', '639-3', 'fax']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kiowa', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kiowa', 'Region', 'western', 'Oklahoma', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Kiowa', 'Region', 'Kiowa', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kiowa', 'Language', 'codes', 'ISO', '639-3', 'kio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gagauz', 'Spoken', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gagauz', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gagauz', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Gagauz', 'Region', 'Gagauzia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gagauz', 'Official', 'language', 'in', 'Gagauzia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Gagauz', 'Regulated', 'by', 'No', 'official', 'regulation']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Gagauz', 'Recognised', 'minority', 'language', 'in', 'Moldova']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Gagauz', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Gagauz', 'Language', 'codes', 'ISO', '639-3', 'gag']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chulym', 'Spoken', 'in', 'Russia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chulym', 'Total', 'speakers', '35-40']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chulym', 'Language', 'codes', 'ISO', '639-3', 'clw']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sicani', 'Language', 'extinction', 'BCE']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicani', 'Spoken', 'in', 'Sicily']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicani', 'Language', 'codes', 'ISO', '639-3', 'sxc']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shuadit', 'Spoken', 'in', 'formerly', 'southern', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'formerly', 'southern', 'France']]]], 6, True)

['Shuadit', 'Region', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['Shuadit', 'Language', 'codes', 'ISO', '639-3', 'sdt']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Judaeo-Portuguese', 'Spoken', 'in', 'Netherlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Judaeo-Portuguese', 'Spoken', 'in', 'England']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Judaeo-Portuguese', 'Spoken', 'in', 'North', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Judaeo-Portuguese', 'Total', 'speakers', 'less', 'than', '2', ',', '000',  
'users', 'in', 'a', 'very', 'limited', 'liturgical', 'context']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'less', 'than', 'NUM', ',', 'NUM', 'users', 'in', 'a', 'very', 'limited',  
'liturgical', 'context']]], ['NUM', '2'], ['NUM', '000']], 15, True)

['Judaeo-Portuguese', 'Total', 'speakers', 'unknown', 'numbers',  
'elsewhere', ',', '']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'unknown', 'numbers', 'elsewhere', ',', '']]]], 7, True)

['Judaeo-Portuguese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Kankana-ey', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

```
['Kankana-ey', 'Region', 'Northern', 'Luzon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kankana-ey', 'Language', 'codes', 'ISO', '639-3', 'kne']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Proto-Norse', 'Spoken', 'in', 'Scandinavia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Proto-Norse', 'Writing', 'system', 'Elder', 'Futhark']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Proto-Norse', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)
```

```
['Kumaoni', 'Spoken', 'in', 'Kumaon', '(', 'Uttarakhand', ',', 'India',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kumaon',  
'\(', 'Uttarakhand', ',', 'India', '\)']]], 9, True)
```

```
['Kumaoni', 'Writing', 'system', 'Devanagari', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kumaoni', 'Official', 'language', 'in', 'Uttarakhand', '(', 'India',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\(', 'India', '\)']]], 8, True)
```

```
['Kumaoni', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Kumaoni', 'Language', 'codes', 'ISO', '639-3', 'kfy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kalenjin', 'Spoken', 'in', 'Kenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```

['Kalenjin', 'Language', 'codes', 'ISO', '639-3', 'kln']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yokutsan', 'Spoken', 'in', 'Yokut', 'people']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Yokutsan', 'Language', 'codes', 'ISO', '639-3', 'yok']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Logba', 'Spoken', 'in', 'Ghana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Logba', 'Region', 'Volta', ',', 'north-west', 'of', 'Ho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
',', 'Type: ANY', 'of', 'Type: ANY']]], 7, True)

['Logba', 'Language', 'codes', 'ISO', '639-3', 'lgq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Coahuilteco', 'Spoken', 'in', 'Mexico', ',', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Coahuilteco', 'Region', 'Coahuila', ',', 'Texas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Coahuilteco', 'Language', 'codes', 'ISO', '639-3', 'xcw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Coahuilteco', 'Language', 'codes', 'ISO', '639-3', 'Coahuilteco',
'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Coahuilteco', 'language']]], ['NUM', '639-3']], 7, True)

['Northern', 'Region', 'Northern', 'Ireland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

```
['Northern', 'Signed', 'in', 'Ireland', ',', 'United', 'Kingdom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in',  
'Ireland', ',', 'United', 'Kingdom']]]], 7, True)
```

```
['Northern', 'Total', 'signers', '3', ',', '500', '(', 'recognised', 'as',  
'BSL', 'signers', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'NUM', ',', 'NUM', '\\(', 'recognised', 'as', 'BSL', 'signers', '\\)']]],  
['NUM', '3'], ['NUM', '500']]], 12, True)
```

```
['Maya', 'Region', 'Isolated', 'villages', 'in', 'south-central', 'Yucat',  
'\xc3', '\xa1', 'n', ',', 'Guatemalan', 'Highlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Isolated',  
'villages', 'in', 'south-central', 'Yucat', '\xc3', '\xa1', 'n', ',',  
'Guatemalan', 'Highlands']]]], 13, True)
```

```
['Maya', 'Signed', 'in', 'Mexico', ',', 'Guatemala']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'Mexico',  
, ',', 'Guatemala']]]], 6, True)
```

```
['Maya', 'Total', 'signers', 'Unknown', '.', '16', 'deaf', 'signers',  
'and', '400', '\xe2', '\x80', '\x93', '500', 'hearing', 'signers', 'in',  
'primary', 'village']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'Unknown', '.', 'NUM', 'deaf', 'signers', 'and', 'NUM', '\xe2', '\x80',  
'\x93', 'NUM', 'hearing', 'signers', 'in', 'primary', 'village']]], ['NUM',  
'16'], ['NUM', '400'], ['NUM', '500']]], 19, True)
```

```
['Maya', 'Dialects', 'Nohya', 'Sign']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Dogri', 'Region', 'Jammu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)
```

```
['Dogri', 'Region', 'Kashmir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)
```

```
['Dogri', 'Region', 'Himachal', 'Pradesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Dogri', 'Region', 'Gurdaspur', '/', 'Pathankot', 'Punjab']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Gurdaspur',
```

['/', 'Pathankot', 'Punjab']]], 6, True)

['Dogri', 'Spoken', 'in', 'India', ',', 'Pakistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Dogri', 'Writing', 'system', 'Devanagari']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dogri', 'Writing', 'system', 'Takri']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dogri', 'Writing', 'system', 'Perso-Arabic', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Dogri', 'Language', 'codes', 'ISO', '639-2', 'doi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dogri', 'Language', 'codes', 'ISO', '639-3', 'doi', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'doi', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Dogri', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Dogri', 'Language', 'codes', 'ISO', '639-3', 'xnr', '\xe2', '\x80',  
'\x93', 'Kangri']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'xnr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Anfillo', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Anfillo', 'Total', 'speakers', '25', 'monolinguals']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'monolinguals']], ['NUM', '25']], 5, True)

```
['Anfillo', 'Language', 'codes', 'ISO', '639-3', 'myo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Konso', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Konso', 'Writing', 'system', 'Ethiopic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Konso', 'Language', 'codes', 'ISO', '639-3', 'kxc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bussa', 'Total', 'speakers', '8', ',', '537', 'of', 'Mashile',
'dialect']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'of', 'Mashile', 'dialect']], ['NUM', '8'], ['NUM',
'537']], 9, True)
```

```
['Bussa', 'Total', 'speakers', '9', ',', '552', 'of', 'Mosiye', 'dialect']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'of', 'Mosiye', 'dialect']], ['NUM', '9'], ['NUM',
'552']], 9, True)
```

```
['Bussa', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bussa', 'Language', 'codes', 'ISO', '639-3', 'dox']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Eblaite', 'Spoken', 'in', 'Ebla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Eblaite', 'Language', 'codes', 'ISO', '639-3', 'xeb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Quebec', 'Signed', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```



```

['Quebec', 'Total', 'signers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Modern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Old', 'Spoken', 'in', 'Sweden', ',', 'Finland', 'and', '\xc3', '\x85',
'land']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Sweden',
',', 'Finland', 'and', '\xc3', '\x85', 'land']]], 10, True)

['Old', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'non-swe', '\xe2', '\x80',
'\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'non-swe', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,
True)

['Khazar', 'Spoken', 'in', 'Southern', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Khazar', 'Spoken', 'in', 'the', 'northern', 'Caucasus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the',
'northern', 'Caucasus']]], 6, True)

['Khazar', 'Language', 'codes', 'ISO', '639-3', 'zkh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sawai', 'Dialects', 'Weda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Sawai', 'Dialects', 'Sawai']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

```

ANY'[]], 3, True)

['Sawai', 'Dialects', 'Kobe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Sawai', 'Dialects', 'Faya-Mafa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Sawai', 'Dialects', 'Messa-Dote']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Sawai', 'Region', 'North', 'Maluku', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 5, True)

['Sawai', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Sawai', 'Language', 'codes', 'ISO', '639-3', 'szw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'[]], ['NUM', '639-3'[]]], 6, True)

['Plains', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY'[]]], 5, True)

['Plains', 'Region', 'Primarily', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Plains', 'Language', 'codes', 'ISO', '639-3', 'apk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'[]], ['NUM', '639-3'[]]], 6, True)

['Kipsigis', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Kipsigis', 'Language', 'codes', 'ISO', '639-3', 'sgc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'[]], ['NUM', '639-3'[]]], 6, True)

```
['Bolgo', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bolgo', 'Language', 'codes', 'ISO', '639-3', 'bvo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Mescalero-Chiricahua', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mescalero-Chiricahua', 'Language', 'codes', 'ISO', '639-2', 'apa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Mescalero-Chiricahua', 'Language', 'codes', 'ISO', '639-3', 'apm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Lipan', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lipan', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lipan', 'Region', 'New', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lipan', 'Region', 'Texas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Lipan', 'Region', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Lipan', 'Language', 'codes', 'ISO', '639-2', 'apa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Lipan', 'Language', 'codes', 'ISO', '639-3', 'apl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['South', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Judaean-Georgian', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judaean-Georgian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judaean-Georgian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judaean-Georgian', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judaean-Georgian', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Judaean-Georgian', 'Language', 'codes', 'ISO', '639-3', 'jge']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Old', 'Writing', 'system', 'Old', 'Turkic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Old', 'Writing', 'system', 'Brahmi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Writing', 'system', 'Aramaic-derived']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Old', 'Writing', 'system', 'Uyghur', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'otk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Brabantian', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Brabantian', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Brabantian', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Brabantian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Newfoundland', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Newfoundland', 'Language', 'codes', 'ISO', '639-2', 'gle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Newfoundland', 'Language', 'codes', 'ISO', '639-1', 'ga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Old', 'Signed', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['R\xc4\x81jasth\xc4\x81n\xc4\xab', 'Total', 'speakers', '80', 'million',

('(', 'approx', '.', ')')]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', '\\(', 'approx', '.', '\\)']], ['NUM', '80']], 9, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Region', 'Rajasthan', 'and', 'its',  
'adjacent', 'areas', 'in', 'India', '.', 'Also', 'in', 'some', 'parts',  
'of', 'Sindh', 'and', 'Punjab', 'of', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rajasthan',  
'and', 'its', 'adjacent', 'areas', 'in', 'India', '.', 'Also', 'in',  
'some', 'parts', 'of', 'Sindh', 'and', 'Punjab', 'of', 'Pakistan']], 20,  
True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-2',  
'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3',  
'raj', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'raj', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3',  
'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3',  
':', 'bgq', '\\xe2', '\\x80', '\\x93', 'Bagri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'bgq', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3',  
'gda', '\\xe2', '\\x80', '\\x93', 'Gade', 'Lohar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'gda', '\\xe2', '\\x80', '\\x93', 'Gade', 'Lohar']], ['NUM',

'639-3']], 11, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'gju', '\\xe2', '\\x80', '\\x93', 'Gujari']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gju', '\\xe2', '\\x80', '\\x93', 'Gujari']], ['NUM', '639-3']], 10, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'mup', '\\xe2', '\\x80', '\\x93', 'Malvi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mup', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'wbr', '\\xe2', '\\x80', '\\x93', 'Wagdi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wbr', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'hoj', '\\xe2', '\\x80', '\\x93', 'Harauti']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hoj', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['R\\xc4\\x81jasth\\xc4\\x81n\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'lmn', '\\xe2', '\\x80', '\\x93', 'Lambadi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lmn', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Dahalo', 'Total', 'speakers', '400']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dahalo', 'Spoken', 'in', 'Kenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dahalo', 'Region', 'Coast', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dahalo', 'Language', 'codes', 'ISO', '639-2', 'cus']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dahalo', 'Language', 'codes', 'ISO', '639-3', 'dal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thompson', 'Total', 'speakers', '720']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Thompson', 'Language', 'codes', 'ISO', '639-3', 'thp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tonkawa', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tonkawa', 'Language', 'codes', 'ISO', '639-3', 'tqw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Brigasc', 'Spoken', 'in', 'Italy', ',', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Brigasc', 'Language', 'codes', 'ISO', '639-3', 'lij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Brigasc', 'Language', 'codes', 'Linguasphere', '51-AAA-og']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Gheg', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gheg', 'Spoken', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gheg', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Macedonia']]], 6, True)



['Gheg', 'Spoken', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gheg', 'Spoken', 'in', 'Serbia', '(', 'Pre', '\xc5', '\xa1', 'evo', 'Valley', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Serbia', '\\(', 'Pre', '\xc5', '\xa1', 'evo', 'Valley', '\\)']]], 11, True)

['Gheg', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gheg', 'Spoken', 'in', 'Western', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gheg', 'Spoken', 'in', 'USA', 'and', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Gheg', 'Language', 'codes', 'ISO', '639-2', 'alb', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'alb', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Gheg', 'Language', 'codes', 'ISO', '639-2', 'sqi', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sqi', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Gheg', 'Language', 'codes', 'ISO', '639-3', 'aln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gheg', 'Language', 'codes', 'ISO', '639-1', 'sq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Tuscan', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tuscan', 'Region', 'Tuscany', 'except', 'the', 'Province', 'of', 'Massa-Carrara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Tuscany', 'except', 'the', 'Province', 'of', 'Massa-Carrara']]], 8, True)

```
['Tuscan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Tuscan', 'Language', 'codes', 'ISO', '639-3', 'ita-tus']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tuscan', 'Language', 'codes', 'Linguasphere', '51-AAA-qa']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)
```

```
['aUI', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']], 6, True)
```

```
['aUI', 'Setting', 'and', 'usage', 'Designed', 'so', 'that', 'ideally',  
,',', 'the', 'meaning', 'of', 'each', 'phoneme', 'would', 'tie', 'into',  
'its', 'properties']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Designed', 'so', 'that', 'ideally', ',,', 'the', 'meaning', 'of',  
'each', 'phoneme', 'would', 'tie', 'into', 'its', 'properties']], 19,  
True)
```

```
['aUI', 'Created', 'by', 'John', 'W', '.', 'Weilgart']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'John',  
'W', '.', 'Weilgart']], 7, True)
```

```
['aUI', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']], 7, True)
```

```
['aUI', 'Category', '(', 'purpose', ')', 'logical', 'or', 'philosophical',  
'language']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'logical', 'or', 'philosophical', 'language']], 9, True)
```

```
['aUI', 'Category', '(', 'purpose', ')', 'aUI']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']], 6, True)
```

```
['aUI', 'Date', 'founded', '1962']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['aUI', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['aUI', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Luise\xc3\xb1o', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Luise\xc3\xb1o', 'Region', 'Southern', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Luise\xc3\xb1o', 'Language', 'codes', 'ISO', '639-2', 'lui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Luise\xc3\xb1o', 'Language', 'codes', 'ISO', '639-3', 'lui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Makhuwa', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Makhuwa', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Makhuwa', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Makhuwa', 'Language', 'codes', 'ISO', '639-3', 'vmw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Latgalian', 'Spoken', 'in', 'Latvia', ',', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Latgalian', 'Region', 'Latgalia']:
```

```
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Latgalian', 'Region', 'Selonia']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Latgalian', 'Region', 'Vidzeme']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Latgalian', 'Region', 'Siberia']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Latgalian', 'Region', 'Bashkiriya']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Latgalian', 'Language', 'codes', 'ISO', '639-3', 'ltg']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Boholano', 'Spoken', 'in', 'Philippines']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Boholano', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)
```

```
['Boholano', 'Language', 'codes', 'ISO', '639-3', 'ceb']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chibcha', 'Language', 'codes', 'ISO', '639-2', 'chb']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Chibcha', 'Language', 'codes', 'ISO', '639-3', 'chb']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Caddo', 'Total', 'speakers', '25']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Caddo', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Caddo', 'Language', 'codes', 'ISO', '639-2', 'cad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Caddo', 'Language', 'codes', 'ISO', '639-3', 'cad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Classical', 'Spoken', 'in', 'Sikkim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Classical', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Classical', 'Spoken', 'in', 'Tibet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Classical', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Classical', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Classical', 'Writing', 'system', 'Nepal', 'Scripts', ':', 'Ranjana',  
'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Classical', 'Writing', 'system', 'Prachalit', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Classical', 'Writing', 'system', 'Brahmi', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```

['Classical', 'Writing', 'system', 'Kutilla', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Classical', 'Writing', 'system', 'Bhujimol', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Classical', 'Writing', 'system', 'Golmol', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Classical', 'Writing', 'system', 'Litumol', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Classical', 'Writing', 'system', 'Other', ':', 'Devnagari', 'Script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Other', ':', 'Devnagari', 'Script']]], 7, True)

['Classical', 'Writing', 'system', 'Brahmi', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'nwc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'nwc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Emiliano-Romagnolo', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Emiliano-Romagnolo', 'Spoken', 'in', 'San', 'Marino']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Emiliano-Romagnolo', 'Total', 'speakers', '2', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '2']], 5, True)

['Emiliano-Romagnolo', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Emiliano-Romagnolo', 'Language', 'codes', 'ISO', '639-3', ':', 'egl',  
'\xe2', '\x80', '\x93', 'Emilian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'egl', '\xe2', '\x80', '\x93', 'Emilian']], ['NUM',  
'639-3']], 11, True)

['Emiliano-Romagnolo', 'Language', 'codes', 'ISO', '639-3', 'rgn', '\xe2',  
'\x80', '\x93', 'Romagnol']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'rgn', '\xe2', '\x80', '\x93', 'Romagnol']], ['NUM',  
'639-3']], 10, True)

['Emiliano-Romagnolo', 'Language', 'codes', 'Linguasphere', '51-AAA-ok']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Lillooet', 'Total', 'speakers', '200', '(', 'native', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'native', '\\)']], ['NUM', '200']], 7, True)

['Lillooet', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lillooet', 'Region', 'British', 'Columbia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lillooet', 'Language', 'codes', 'ISO', '639-2', 'sal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lillooet', 'Language', 'codes', 'ISO', '639-3', 'lil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yapese', 'Spoken', 'in', 'Federated', 'States', 'of', 'Micronesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']], 7, True)

['Yapese', 'Region', 'Island', 'of', 'Yap']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

```

['Yapese', 'Language', 'codes', 'ISO', '639-2', 'yap']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Yapese', 'Language', 'codes', 'ISO', '639-3', 'yap']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Argobba', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Argobba', 'Language', 'codes', 'ISO', '639-3', 'agj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gafat,', 'Spoken', 'in', 'Ethiopia', ',', 'Eritrea', 'and', 'Israel']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Ethiopia', ',', 'Eritrea', 'and', 'Israel']]], 8, True)

['Gafat,', 'Language', 'codes', 'ISO', '639-3', 'gft']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vunjo', 'Spoken', 'in', 'Tanzania', ',', 'Kenya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Vunjo', 'Region', 'Kilimanjaro', ',', 'Chaga', 'area']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kilimanjaro',
',', 'Chaga', 'area']]], 6, True)

['Vunjo', 'Language', 'codes', 'ISO', '639-2', 'bnt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Vunjo', 'Language', 'codes', 'ISO', '639-3', 'vun']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cape', 'Language', 'codes', 'ISO', '639-2', 'cpp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cape', 'Language', 'codes', 'ISO', '639-3', 'kea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cape', 'Language', 'codes', 'Linguasphere', '51-AAC-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Goaria', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Goaria', 'Region', 'Sindh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Goaria', 'Language', 'codes', 'ISO', '639-3', 'gig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ni\xc3\xa7ard', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Ni\xc3\xa7ard', 'Region', 'County', 'of', 'Nice', ',', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'County', 'of',  
'Nice', ',', 'Monaco']]], 7, True)

['Ni\xc3\xa7ard', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ni\xc3\xa7ard', 'Spoken', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ni\xc3\xa7ard', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ni\xc3\xa7ard', 'Pronunciation', 'ni', '\xcb', '\x88', 'sa', '\xca',  
'\x80', 't']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ni',  
'\xcb', '\x88', 'sa', '\xca', '\x80', 't']]], 9, True)

['Ni\xc3\xa7ard', 'Official', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Ni\xc3\xa7ard', 'Regulated', 'by', 'Conselh', 'de', 'la', 'Lenga',  
'Occitana', '(', 'norme', 'classique', ')', '/', 'F', '\xc3', '\xa9',  
'librige', '(', 'norme', 'mistralienne', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Conselh', 'de', 'la', 'Lenga', 'Occitana', '\\(', 'norme', 'classique', '\\  
)', '/', 'F', '\xc3', '\xa9', 'librige', '\\(', 'norme', 'mistralienne',  
'\\)']]], 21, True)
```

```
['Ni\xc3\xa7ard', 'Language', 'codes', 'ISO', '639-3', 'oci-nic', '\xe2',  
'\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'oci-nic', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,  
True)
```

```
['Ni\xc3\xa7ard', 'Language', 'codes', 'Linguasphere', '51-AAA-gd']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['German', 'Signed', 'in', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['German', 'Language', 'codes', 'ISO', '639-2', 'sgn-DE']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Northern', 'Spoken', 'in', 'Zimbabwe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Northern', 'Spoken', 'in', 'Botswana']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Northern', 'Spoken', 'in', 'South', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-2', 'nde']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-3', 'nde']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-1', 'nd']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Akkala', 'Spoken', 'in', 'Southwest', 'Kola', 'Peninsula', ',',  
'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Southwest', 'Kola', 'Peninsula', ',', 'Russia']]], 8, True)
```

```
['Akkala', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Akkala', 'Language', 'codes', 'ISO', '639-3', 'sia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Oorlans', 'Spoken', 'in', 'South', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Oorlans', 'Language', 'codes', 'ISO', '639-3', 'oor']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Rohingya', 'Spoken', 'in', 'Burma', ',', 'Bangladesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Rohingya', 'Region', 'Arakan', 'region', 'of', 'Burma', ',',  
'southeastern', 'Chittagong', 'region', 'of', 'Bangladesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Arakan',  
'region', 'of', 'Burma', ',', 'southeastern', 'Chittagong', 'region', 'of',  
'Bangladesh']]], 12, True)
```

```
['Rohingya', 'Language', 'codes', 'ISO', '639-3', 'rhg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Fataluku', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Fataluku', 'Language', 'codes', 'ISO', '639-3', 'ddg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Garhwali', 'Spoken', 'in', 'Garhwal', '(', 'Uttarakhand', ',', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Garhwal', '\\(', 'Uttarakhand', ',', 'India', '\\)']], 9, True)

['Garhwali', 'Writing', 'system', 'Devanagari', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Garhwali', 'Official', 'language', 'in', 'Uttarakhand', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'India', '\\)']], 8, True)

['Garhwali', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Garhwali', 'Language', 'codes', 'ISO', '639-3', 'gbm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baga', 'Spoken', 'in', 'Guinea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Baga', 'Region', 'Along', 'the', 'coast', ',', 'east', 'of', 'the', 'Nunez', 'River', 'in', 'the', 'Bok', '\\xc3', '\\xa9']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Along', 'the', 'coast', ',', 'east', 'of', 'the', 'Nunez', 'River', 'in', 'the', 'Bok', '\\xc3', '\\xa9']], 16, True)

['Baga', 'Language', 'codes', 'ISO', '639-3', 'bcg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maasai', 'Dialects', 'Purko']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Maasai', 'Dialects', 'Kwavi', '(', 'Baraguyu', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Maasai', 'Dialects', 'Kaputiei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Keekonyokie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Matapo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Laitokitok']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Iloodokilani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Damat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Loitai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Siria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Moitanik', '(', 'Wuasinkishu', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 6, True)

['Maasai', 'Dialects', 'Kore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Arusha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maasai', 'Dialects', 'Kisonko']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

ANY']]], 3, True)

['Maasai', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Maasai', 'Language', 'codes', 'ISO', '639-2', 'mas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Maasai', 'Language', 'codes', 'ISO', '639-3', 'mas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tetuaní', 'Spoken', 'in', 'now', 'mostly', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'now',  
'mostly', 'Israel']]], 6, True)

['Tetuaní', 'Total', 'speakers', 'a', 'few', 'thousand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a',  
'few', 'thousand']]], 6, True)

['Tetuaní', 'Language', 'codes', 'ISO', '639-2', 'lad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Tetuaní', 'Language', 'codes', 'ISO', '639-3', 'lad', '\xe2', '\x80',  
'\x93', 'Ladino']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lad', '\xe2', '\x80', '\x93', 'Ladino']]], ['NUM', '639-3']],  
10, True)

['Rotuman', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rotuman', 'Region', 'Rotuma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Rotuman', 'Language', 'codes', 'ISO', '639-3', 'rtm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Edo', 'Total', 'speakers', '1', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'NUM', 'million']], ['NUM', '1']], 5, True)

['Edo', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Edo', 'Region', 'Edo', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Edo', 'Language', 'codes', 'ISO', '639-2', 'bin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Edo', 'Language', 'codes', 'ISO', '639-3', 'bin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tiv', 'Spoken', 'in', 'southeast', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tiv', 'Language', 'codes', 'ISO', '639-3', 'tiv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warlpiri', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)

['Warlpiri', 'Language', 'codes', 'ISO', '639-3', 'wbp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cook', 'Spoken', 'in', 'Cook', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cook', 'Official', 'language', 'in', 'Cook', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Cook', 'Regulated', 'by', 'Kopapa', 'Reo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

```

['Cook', 'Language', 'codes', 'ISO', '639-2', 'rar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cook', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cook', 'Language', 'codes', 'ISO', '639-3', ':', 'rar', '\xe2', '\x80',
'\x93', 'Rarotonga', 'dialect']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'rar', '\xe2', '\x80', '\x93', 'Rarotonga', 'dialect']],
['NUM', '639-3']], 12, True)

['Kamayo', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kamayo', 'Region', 'Mindanao']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Kamayo', 'Language', 'codes', 'ISO', '639-3', 'kyk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gela', 'Region', 'Big', 'Nggela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Gela', 'Region', 'Small', 'Nggela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Gela', 'Region', 'Sandfly', 'and', 'Buenavista', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sandfly',
'and', 'Buenavista', 'Islands']], 6, True)

['Gela', 'Language', 'codes', 'ISO', '639-3', 'nlg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Damin', 'Language', 'extinction', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

```



```
['Damin', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Damin', 'Category', '(', 'purpose', ')', 'ritual', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Damin', 'Category', '(', 'purpose', ')', 'Damin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']]], 6, True)
```

```
['Damin', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Afrikaans', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Afrikaans', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-2', 'afr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-3', 'afr', '\\xe2', '\\x80',  
'\\x93', 'Afrikaans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'afr', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-1', 'af']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Afrikaans', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Afrikaans', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-2', 'afr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-3', 'afr', '\xe2', '\x80',  
\x93', 'Afrikaans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'afr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-1', 'af']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Afrikaans', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Afrikaans', 'Region', 'Western', 'Cape', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-2', 'afr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-3', 'afr', '\xe2', '\x80',  
\x93', 'Afrikaans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'afr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Afrikaans', 'Language', 'codes', 'ISO', '639-1', 'af']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Ndau', 'Spoken', 'in', 'Mozambique', ',', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Ndau', 'Language', 'codes', 'ISO', '639-3', 'ndc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Haya', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Haya', 'Language', 'codes', 'ISO', '639-3', 'hay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Samburu', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Samburu', 'Language', 'codes', 'ISO', '639-3', 'saq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ongamo', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ongamo', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ongamo', 'Language', 'codes', 'ISO', '639-3', 'nsg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yaaku', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yaaku', 'Region', 'Laikipia', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yaaku', 'Language', 'codes', 'ISO', '639-3', 'muu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khmu\xe2\x80\x99', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Khmu\xe2\x80\x99', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Khmu\xe2\x80\x99', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Khmu\xe2\x80\x99', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Khmu\xe2\x80\x99', 'Language', 'codes', 'ISO', '639-3', 'kjp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chumash', 'Spoken', 'in', 'southern', 'coastal', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'southern', 'coastal', 'California']], 6, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', ':', 'boi', '\xe2',  
'\x80', '\x93', 'Barbare', '\xc3', '\xb1', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'boi', '\xe2', '\x80', '\x93', 'Barbare', '\xc3',  
'\xb1', 'o']], ['NUM', '639-3']], 14, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', 'crz', '\xe2', '\x80',  
'\x93', 'Cruze', '\xc3', '\xb1', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'crz', '\xe2', '\x80', '\x93', 'Cruze', '\xc3', '\xb1',  
'o']], ['NUM', '639-3']], 13, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', 'inz', '\xe2', '\x80',  
'\x93', 'Inese', '\xc3', '\xb1', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'inz', '\xe2', '\x80', '\x93', 'Inese', '\xc3', '\xb1',  
'o']], ['NUM', '639-3']], 13, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', 'obi', '\xe2', '\x80',  
'\x93', 'Obispe', '\xc3', '\xb1', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'obi', '\xe2', '\x80', '\x93', 'Obispe', '\xc3', '\xb1',  
'o']], ['NUM', '639-3']], 13, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', 'puy', '\xe2', '\x80',  
'\x93', 'Purisime', '\xc3', '\xb1', 'o']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'puy', '\xe2', '\x80', '\x93', 'Purissime', '\xc3', '\xb1', 'o']], ['NUM', '639-3']], 13, True)
```

```
['Chumash', 'Language', 'codes', 'ISO', '639-3', 'veo', '\xe2', '\x80', '\x93', 'Venture', '\xc3', '\xb1', 'o']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'veo', '\xe2', '\x80', '\x93', 'Venture', '\xc3', '\xb1', 'o']], ['NUM', '639-3']], 13, True)
```

```
['Qiong', 'Region', 'Hainan']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Qiong', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Qiong', 'Language', 'codes', 'ISO', '639-3', 'nan-hai']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Qiong', 'Language', 'codes', 'ISO', '639-3', 'nan-han', '(', 'redundant', ')']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nan-han', '\\(', 'redundant', '\\)']], ['NUM', '639-3']], 9, True)
```

```
['Arcaicam', 'Created', 'by', 'Manuel', 'Halvelik']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Arcaicam', 'Writing', 'system', 'Latin']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Arcaicam', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)
```

```
['Arcaicam', 'Category', '(', 'purpose', ')', 'Esperanto']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)
```

```
['Arcaicam', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Arcaicam', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Zeelandic', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Zeelandic', 'Language', 'codes', 'ISO', '639-3', 'zea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zeelandic', 'Language', 'codes', 'Linguasphere', '52-ACB-af']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Ticinese', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Ticinese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['oluLuhya', 'Spoken', 'in', 'Kenya', ',', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)
```

```
['oluLuhya', 'Region', 'East', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['oluLuhya', 'Language', 'codes', 'ISO', '639-3', 'luy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Defaka', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Defaka', 'Region', 'Rivers', 'State', ',', 'Bonny', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rivers',  
'State', ',', 'Bonny', 'District']]], 7, True)
```

```
['Defaka', 'Language', 'codes', 'ISO', '639-3', 'afn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Malvi', 'Spoken', 'in', 'India', '(', 'Malva', 'region', 'of', 'Madhya',  
'Pradesh', ',', 'Rajasthan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
'\\(', 'Malva', 'region', 'of', 'Madhya', 'Pradesh', ',', 'Rajasthan', '\\  
)']]], 13, True)
```

```
['Malvi', 'Language', 'codes', 'ISO', '639-2', 'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Malvi', 'Language', 'codes', 'ISO', '639-3', 'mup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aquitanian', 'Spoken', 'in', 'France', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Aquitanian', 'Language', 'codes', 'ISO', '639-3', 'xaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mecklenburgisch-Vorpommersch', 'Total', 'speakers', '\\xe2', '\\x80',  
'\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']]], 6, True)
```

```
['Mecklenburgisch-Vorpommersch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mecklenburgisch-Vorpommersch', 'Region', 'Mecklenburg-Vorpommern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Mecklenburgisch-Vorpommersch', 'Language', 'codes', 'ISO', '639-3',  
'\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Mecklenburgisch-Vorpommersch', 'Language', 'codes', 'ISO', '639-3', 'nds-mec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Germany', '(', 'Baden-W', '\xc3', '\xbc', 'rttemberg', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Germany', '\\(', 'Baden-W', '\xc3', '\xbc', 'rttemberg', '\\)']], 10, True)

['South', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['South', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['South', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'goh-sfr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tangut', 'Spoken', 'in', 'Western', 'Xia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tangut', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Tangut', 'Writing', 'system', 'Tangut', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Tangut', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:



ANY', 'of', 'Type: ANY']]], 6, True)

['Tangut', 'Language', 'codes', 'ISO', '639-3', 'txg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'Italy', '(', 'Province', 'of', 'Milan', ',',  
'Province', 'of', 'Monza', ',', 'Province', 'of', 'Como', ',', 'Province',  
'of', 'Lecco', ',', 'Province', 'of', 'Lodi', ',', 'a', 'small', 'part',  
'of', 'Province', 'of', 'Cremona', ',', 'Province', 'of', 'Novara', ',',  
'Province', 'of', 'Pavia', ',', 'Province', 'of', 'Sondrio', ',',  
'Province', 'of', 'Varese', ',', 'Province', 'of', 'Verbano', 'Cusio',  
'Ossola', ',', 'a', 'small', 'part', 'of', 'Province', 'of', 'Vercelli',  
)', 'and', 'Switzerland', '(', 'Canton', 'Ticino', 'and', 'some',  
'valleys', 'of', 'Canton', 'Grigioni', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy',  
'\\(', 'Province', 'of', 'Milan', ',', 'Province', 'of', 'Monza', ',',  
'Province', 'of', 'Como', ',', 'Province', 'of', 'Lecco', ',', 'Province',  
'of', 'Lodi', ',', 'a', 'small', 'part', 'of', 'Province', 'of', 'Cremona',  
, ',', 'Province', 'of', 'Novara', ',', 'Province', 'of', 'Pavia', ',',  
'Province', 'of', 'Sondrio', ',', 'Province', 'of', 'Varese', ',',  
'Province', 'of', 'Verbano', 'Cusio', 'Ossola', ',', 'a', 'small', 'part',  
'of', 'Province', 'of', 'Vercelli', '\\)')', 'and', 'Switzerland', '\\(',  
'Canton', 'Ticino', 'and', 'some', 'valleys', 'of', 'Canton', 'Grigioni',  
'\\)']]], 75, True)

['Western', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tabasaran', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tabasaran', 'Region', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tabasaran', 'Language', 'codes', 'ISO', '639-3', 'tab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tosk', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Tosk', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Macedonia']]], 6, True)
```

```
['Tosk', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tosk', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Tosk', 'Language', 'codes', 'ISO', '639-2', 'alb', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'alb', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Tosk', 'Language', 'codes', 'ISO', '639-2', 'sqi', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'sqi', '\\(', 'T', '\\)'], ['NUM', '639-2']], 9, True)

['Tosk', 'Language', 'codes', 'ISO', '639-3', 'als']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tosk', 'Language', 'codes', 'ISO', '639-1', 'sq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Highest', 'Spoken', 'in', 'the', 'Alps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Highest', 'Language', 'codes', 'ISO', '639-2', 'gsw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Highest', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Upper', 'Saxon', 'Total', 'speakers', '\\e2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\\e2', '\\x80', '\\x93']], 7, True)

['Upper', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Upper', 'Region', 'Saxony']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Upper', 'Language', 'codes', 'ISO', '639-3', 'sxu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Moselle', 'Franconian', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Moselle', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Moselle', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Moselle', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Moselle', 'Region', 'North', 'Rhine-Westphalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Moselle', 'Region', 'Rhineland-Palatinate']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Moselle', 'Region', 'Saarland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Moselle', 'Region', 'Lorraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Moselle', 'Region', 'Li', '\xc3', '\xa8', 'ge']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Li', '\xc3',  
'\xa8', 'ge']]], 6, True)
```

```
['Moselle', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Tukang', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tukang', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taba', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Taba', 'Region', 'North', 'Maluku', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['Taba', 'Language', 'codes', 'ISO', '639-3', 'mky']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kurukh', 'Spoken', 'in', 'India', ',', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Kurukh', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Kurukh', 'Language', 'codes', 'ISO', '639-2', 'kru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kurukh', 'Language', 'codes', 'ISO', '639-3', 'kru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tai', 'Region', 'Yunnan', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Writing', 'system', 'Tai', 'Tham', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Tai', 'Writing', 'system', 'Thai', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Tai', 'Language', 'codes', 'ISO', '639-3', 'khb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tai', 'Region', 'Southwest', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Pronunciation', 'tai', 'ta', '\xc9', '\xaf', 'xo', '\xc5', '\x8b',  
'need', 'tone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'tai',  
'ta', '\xc9', '\xaf', 'xo', '\xc5', '\x8b', 'need', 'tone']]], 11, True)

['Tai', 'Official', 'language', 'in', 'co-official', 'in', 'Dehong', ',',  
'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'co-official', 'in', 'Dehong', ',', 'China']]], 9, True)

['Tai', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tai', 'Language', 'codes', 'ISO', '639-3', 'tdd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Stellingwarfs', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Stellingwarfs', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Stellingwarfs', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Stellingwarfs', 'Language', 'codes', 'ISO', '639-3', 'stl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pawnee', 'Total', 'speakers', '20']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pawnee', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Pawnee', 'Region', 'North-central', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pawnee', 'Language', 'codes', 'ISO', '639-3', 'paw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pawnee', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact',  
'distribution', 'of', 'Pawnee']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Pre-contact', 'distribution', 'of', 'Type: ANY']], ['NUM',  
'639-3']], 9, True)
```

```
['Lemnian', 'Spoken', 'in', 'Lemnos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lemnian', 'Language', 'codes', 'ISO', '639-3', 'xle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Lemnian', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of',  
'Lemnos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY', 'of', 'Type: ANY']], ['NUM', '639-3']], 8, True)

['Lule', 'Spoken', 'in', 'Norway', 'and', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Lule', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Lule', 'Language', 'codes', 'ISO', '639-2', 'smj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Turkana', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Turkana', 'Region', 'Northwest', 'Kenya', ',', 'west', 'of', 'Lake',  
'Turkana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Northwest',  
'Kenya', ',', 'west', 'of', 'Lake', 'Type: ANY']]], 9, True)

['Turkana', 'Language', 'codes', 'ISO', '639-3', 'tuv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Albay', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Albay', 'Region', 'Bicol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Albay', 'Language', 'codes', 'ISO', '639-3', 'bhk', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bhk', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Albay', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Albay', 'Language', 'codes', 'ISO', '639-3', 'lbl', '\xe2', '\x80', '\x93', 'Libon', 'Bikol']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lbl', '\xe2', '\x80', '\x93', 'Libon', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Albay', 'Language', 'codes', 'ISO', '639-3', 'rbl', '\xe2', '\x80', '\x93', 'Miraya', 'Bikol']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rbl', '\xe2', '\x80', '\x93', 'Miraya', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Mussau-Emira', 'Language', 'codes', 'ISO', '639-3', 'emi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Veluws', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Veluws', 'Region', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Veluws', 'Spoken', 'in', 'Netherlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Veluws', 'Pronunciation', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Veluws', 'Language', 'codes', 'ISO', '639-3', 'vel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nheengatu', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nheengatu', 'Spoken', 'in', 'Colombia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Nheengatu', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Nheengatu', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nheengatu', 'Official', 'language', 'in', 'S', '\xc3', '\xa3', 'o',
'Gabriel', 'da', 'Cachoeira', '(', 'Brazil', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'S', '\xc3', '\xa3', 'o', 'Gabriel', 'da', 'Cachoeira', '\\(',
'Brazil', '\\)']]], 14, True)
```

```
['Nheengatu', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Nheengatu', 'Language', 'codes', 'ISO', '639-2', 'tup']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Nheengatu', 'Language', 'codes', 'ISO', '639-3', 'yrl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['East', 'Spoken', 'in', 'Germany', '(', 'Bavaria', ',', 'Thuringia', ',',
'Saxony', ',', 'Baden-W', '\xc3', '\xbc', 'rttemberg', ',', 'Hesse', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Germany', '\\(', 'Bavaria', ',', 'Thuringia', ',', 'Saxony', ',', 'Baden-
W', '\xc3', '\xbc', 'rttemberg', ',', 'Hesse', '\\)']]], 18, True)
```

```
['East', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['East', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['East', 'Language', 'codes', 'ISO', '639-3', 'goh-eas', '\xe2', '\x80',
'\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'goh-eas', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,
True)
```

```
['Shasta', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Shasta', 'Region', 'primarily', 'northern', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Shasta', 'Language', 'codes', 'ISO', '639-3', 'sht']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Quechan', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Quechan', 'Region', 'California', ',', 'Arizona', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
,', 'Arizona', 'Ethnicity']]], 6, True)
```

```
['Quechan', 'Region', 'Quechan', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Quechan', 'Language', 'codes', 'ISO', '639-3', 'yum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Utopian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
, '\xe2', '\x80', '\x93']]], 6, True)
```

```
['Utopian', 'Created', 'by', 'Thomas', 'More', ',', 'Peter', 'Giles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
, 'Thomas', 'More', ',', 'Peter', 'Giles']]], 8, True)
```

```
['Utopian', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
, 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Utopian', 'Category', '(', 'purpose', ')', 'Utopian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
, 'purpose', '\\)', 'Type: ANY']]], 6, True)
```

```
['Utopian', 'Date', 'founded', '1516']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Utopian', 'Official', 'language', 'in', 'Utopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Utopian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Utopian', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Utopian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Ulithian', 'Total', 'speakers', '3', ',', '000', 'in', '1987']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '3'], ['NUM', '000']], 8,  
True)

['Ulithian', 'Spoken', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Ulithian', 'Region', 'Ulithi', 'atoll']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ulithian', 'Official', 'language', 'in', 'Federated', 'States', 'of',  
'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Federated', 'States', 'of', 'Micronesia']]], 8, True)

['Ulithian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Ulithian', 'Language', 'codes', 'ISO', '639-3', 'uli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Savoyard', 'Region', 'Aosta', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Savoyard', 'Region', 'Savoy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Savoyard', 'Region', 'Valais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Savoyard', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Savoyard', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Savoyard', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Savoyard', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Savoyard', 'Official', 'language', 'in', 'Italy', '(', 'protected', 'by', 'statute', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Italy', '\\(', 'protected', 'by', 'statute', '\\)']]], 10, True)

['Savoyard', 'Official', 'language', 'in', 'France', '(', 'region', 'language', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'France', '\\(', 'region', 'language', '\\)']]], 9, True)

['Savoyard', 'Language', 'codes', 'ISO', '639-3', 'frp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Region', 'Kent']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Old', 'Signed', 'in', 'formerly', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in',

'formerly', 'England']]], 5, True)

['Old', 'Total', 'signers', 'now', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'now', 'extinct']]], 5, True)

['Aluku', 'Spoken', 'in', 'French', 'Guiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Aluku', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aluku', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Aluku', 'Language', 'codes', 'ISO', '639-3', 'srm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ndyuka', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ndyuka', 'Language', 'codes', 'ISO', '639-3', 'djk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aku', 'Total', 'speakers', '481', ',', '600', 'in', 'all', 'countries',  
'spoken']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'all', 'countries', 'spoken']], ['NUM', '481'],  
'NUM', '600']], 10, True)

['Aku', 'Spoken', 'in', 'Gambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aku', 'Region', 'West', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aku', 'Official', 'language', 'in', 'n', '/', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'n', '/', 'a']]], 7, True)

['Aku', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Aku', 'Language', 'codes', 'ISO', '639-3', 'kri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chinese', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Total', 'speakers', 'Extinct', '(', 'in', 'China', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Extinct', '\\(', 'in', 'China', '\\)']]], 8, True)

['Chinese', 'Language', 'codes', 'ISO', '639-2', 'cpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Chinese', 'Language', 'codes', 'ISO', '639-3', 'cpi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chinese', 'Language', 'codes', 'Linguasphere', '52-ABB-da']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Cameroon', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Cameroon', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Cameroon', 'Language', 'codes', 'ISO', '639-3', 'wes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cameroon', 'Language', 'codes', 'Linguasphere', '52-ABB-bg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

```
['Grenadian', 'Spoken', 'in', 'Grenada']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Grenadian', 'Language', 'codes', 'ISO', '639-3', 'gcl']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Grenadian', 'Language', 'codes', 'Linguasphere', '52-ABB-as']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Vincentian', 'Spoken', 'in', 'Saint', 'Vincent']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Vincentian', 'Total', 'speakers']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Vincentian', 'Language', 'codes', 'ISO', '639-3', 'svc']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Vincentian', 'Language', 'codes', 'Linguasphere', '52-ABB-aq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Montserrat', 'Spoken', 'in', 'Montserrat']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Montserrat', 'Total', 'speakers']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Montserrat', 'Language', 'codes', 'ISO', '639-3', 'aig']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Montserrat', 'Language', 'codes', 'Linguasphere', '52-ABB-apn']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Antiguan', 'Spoken', 'in', 'Antigua', 'and', 'Barbuda']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```



ANY', 'and', 'Type: ANY']]], 6, True)

['Antiguan', 'Language', 'codes', 'ISO', '639-3', 'aig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saint', 'Language', 'codes', 'ISO', '639-3', 'aig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turks', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Turks', 'Language', 'codes', 'ISO', '639-3', 'tch']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turks', 'Language', 'codes', 'Linguasphere', '52-ABB-ao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Virgin', 'Spoken', 'in', 'Saba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Virgin', 'Spoken', 'in', 'Saint', 'Eustatius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Virgin', 'Spoken', 'in', 'Saint', 'Martin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Virgin', 'Language', 'codes', 'ISO', '639-3', 'vic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anguillan', 'Spoken', 'in', 'Anguilla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Anguillan', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

```
['Anguillan', 'Language', 'codes', 'ISO', '639-3', 'aig']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Anguillan', 'Language', 'codes', 'Linguasphere', '52-ABB-apf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Miskito', 'Spoken', 'in', 'Nicaragua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Miskito', 'Language', 'codes', 'ISO', '639-3', 'bzk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Miskito', 'Language', 'codes', 'Linguasphere', '52-ABB-af']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['San', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['San', 'Language', 'codes', 'ISO', '639-3', 'icr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['San', 'Language', 'codes', 'Linguasphere', '52-ABB-ah']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Afro-Seminole', 'Spoken', 'in', 'United', 'States', ',', 'Mexico']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'S', '.', 'Type: ANY']]], 7, True)
```

```
['Afro-Seminole', 'Total', 'speakers', '200']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Afro-Seminole', 'Language', 'codes', 'ISO', '639-3', 'afs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Afro-Seminole', 'Language', 'codes', 'Linguasphere', '52-ABB-ac']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'Linguasphere', 'Type: ANY']]], 5, True)

['Nauruan', 'Spoken', 'in', 'Nauru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nauruan', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Nauruan', 'Language', 'codes', 'ISO', '639-2', 'cpe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Nauruan', 'Language', 'codes', 'ISO', '639-3', 'cpi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nauruan', 'Language', 'codes', 'Linguasphere', '52-ABB-db']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Guyanese', 'Spoken', 'in', 'Guyana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Guyanese', 'Spoken', 'in', 'Overseas', 'communities']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Guyanese', 'Spoken', 'in', ':', 'Guyanese', 'British']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':', 'Type: ANY', 'Type: ANY']]], 6, True)

['Guyanese', 'Spoken', 'in', 'Guyanese', 'Canadians']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Guyanese', 'Spoken', 'in', 'Guyanese', 'Americans']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Guyanese', 'Language', 'codes', 'ISO', '639-3', 'gyn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Guyanese', 'Language', 'codes', 'Linguasphere', '52-ABB-av']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Language', 'Type: ANY']]], 5, True)
```

```
['Waima', 'Language', 'codes', 'ISO', '639-3', 'rro']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Lebanese', 'Spoken', 'in', 'Lebanon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Lebanese', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Lebanese', 'Writing', 'system', 'Arabic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Lebanese', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Lebanese', 'Language', 'codes', 'ISO', '639-3', 'apc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tsakonian', 'Total', 'speakers', '3', ',', '000', '\xe2', '\x80', '\x94',
'4', ',', '000', 'fluent']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\xe2', '\x80', '\x94', 'NUM', ',', 'NUM', 'fluent']]],
['NUM', '3'], ['NUM', '000'], ['NUM', '4'], ['NUM', '000']], 13, True)
```

```
['Tsakonian', 'Spoken', 'in', 'Greece']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tsakonian', 'Language', 'codes', 'ISO', '639-3', 'tsd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tsakonian', 'Language', 'codes', 'Linguasphere', '56-AAA-b']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Language', 'Type: ANY']]], 5, True)
```

```
['Gelao', 'Spoken', 'in', 'H', '\xc3', '\xa0', 'Giang', 'Province', ',', 'northern', 'Vietnam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'H', '\xc3', '\xa0', 'Giang', 'Province', ',', 'northern', 'Vietnam']]], 11, True)
```

```
['Gelao', 'Spoken', 'in', 'Western', 'Guizhou']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Gelao', 'Spoken', 'in', 'Wenshan', 'Prefecture', 'in', 'Yunnan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Wenshan', 'Prefecture', 'in', 'Yunnan']]], 7, True)
```

```
['Gelao', 'Spoken', 'in', 'Longlin', 'County', 'in', 'Guangxi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Longlin', 'County', 'in', 'Guangxi']]], 7, True)
```

```
['Gelao', 'Spoken', 'in', 'China', ';']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Gelao', 'Language', 'codes', 'ISO', '639-3', 'gio']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sri', 'Spoken', 'in', 'Sri', 'Lanka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sri', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Sri', 'Language', 'codes', 'ISO', '639-3', 'idb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sri', 'Language', 'codes', 'Linguasphere', '51-AAC-age']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Icelandic', 'Signed', 'in', 'Iceland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Icelandic', 'Total', 'signers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Fanagalo', 'Spoken', 'in', 'Southern', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Fanagalo', 'Dialects', 'Chilapalapa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Fanagalo', 'Dialects', 'Cikabanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Fanagalo', 'Language', 'codes', 'ISO', '639-3', 'fng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Japanese', 'Signed', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Japanese', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Japanese', 'Language', 'codes', 'ISO', '639-2', 'sgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Khanty', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Khanty', 'Region', 'Khantia-Mansia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Khanty', 'Language', 'codes', 'ISO', '639-3', 'kca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mansi', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Mansi', 'Region', 'Khantia-Mansia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mansi', 'Language', 'codes', 'ISO', '639-3', 'mns']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ter', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ter', 'Total', 'speakers', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ter', 'Writing', 'system', 'no', 'writing', 'system']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Ume', 'Spoken', 'in', 'Norway', ',', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Ume', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Pite', 'Spoken', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pite', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pite', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bunun', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bunun', 'Official', 'language', 'in', 'officially', 'recognized',  
'minority', 'language', 'of', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'officially', 'recognized', 'minority', 'language', 'of',  
'Taiwan']]], 10, True)

['Bunun', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Bunun', 'Language', 'codes', 'ISO', '639-3', 'bnn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Negerhollands', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Negerhollands', 'Language', 'codes', 'ISO', '639-3', 'dcr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Angloromani', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Angloromani', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Angloromani', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Angloromani', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Angloromani', 'Language', 'codes', 'ISO', '639-3', 'rme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bolaang', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



```
['Bolaang', 'Region', 'North', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bolaang', 'Language', 'codes', 'ISO', '639-3', 'mog']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Ayacucho', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ayacucho', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Ayacucho', 'Language', 'codes', 'ISO', '639-3', 'quy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Cusco-Collao', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cusco-Collao', 'Dialects', 'Cuzco', 'Quechua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cusco-Collao', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Cusco-Collao', 'Language', 'codes', 'ISO', '639-3', 'qxp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Motu', 'Region', 'Central', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Motu', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)
```

```
['Motu', 'Language', 'codes', 'ISO', '639-3', 'meu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Makhuwa-Meetto', 'Spoken', 'in', 'Mozambique', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Makhuwa-Meetto', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Makhuwa-Meetto', 'Language', 'codes', 'ISO', '639-3', 'mgh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tommo', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tommo', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tommo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Acholi', 'Spoken', 'in', 'Uganda', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Acholi', 'Region', 'Acholi', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Acholi', 'Region', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Acholi', 'Region', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Acholi', 'Language', 'codes', 'ISO', '639-2', 'ach']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```

['Acholi', 'Language', 'codes', 'ISO', '639-3', 'ach']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adjuvilo', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Adjuvilo', 'Date', 'founded', '1908']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Adjuvilo', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Adjuvilo', 'Category', '(', 'purpose', ')', 'Ido']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Adjuvilo', 'Category', '(', 'purpose', ')', 'Adjuvilo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Adjuvilo', 'Created', 'by', 'Claudius', 'Colas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Adjuvilo', 'Category', '(', 'sources', ')', 'a', 'posteriori',
'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)', 'a', 'posteriori', 'language']]], 8, True)

['Adjuvilo', 'Language', 'codes', 'ISO', '639-2', 'art']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Adjuvilo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Maharashtri', 'Language', 'extinction', 'developed', 'into', 'Marathi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'developed', 'into', 'Type: ANY']]], 6, True)

['Maharashtri', 'Language', 'extinction', 'Konkani']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maharashtri', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maharashtri', 'Language', 'codes', 'ISO', '639-3', 'pmh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shauraseni', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shauraseni', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Shauraseni', 'Language', 'codes', 'ISO', '639-3', 'psu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jain', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jain', 'Language', 'codes', 'ISO', '639-3', 'pka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lihir', 'Language', 'codes', 'ISO', '639-3', 'lih']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paeonian', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']]], 6, True)

['Paeonian', 'Spoken', 'in', 'Kosovo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Paeonian', 'Spoken', 'in', 'northern', 'Greece']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Paeonian', 'Spoken', 'in', 'southwestern', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Paeonian', 'Language', 'codes', 'ISO', '639-3', 'xli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mewari', 'Spoken', 'in', 'India', '(', 'Mewar', 'region', 'of',  
'Rajasthan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
'\\(', 'Mewar', 'region', 'of', 'Rajasthan', '\\)']]], 10, True)

['Mewari', 'Language', 'codes', 'ISO', '639-2', 'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Mewari', 'Language', 'codes', 'ISO', '639-3', 'mtr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ligurian', 'Spoken', 'in', 'Italy', '(', 'Liguria', ',', 'Piedmont', ',',  
'Tuscany', ',', 'Lombardy', ',', 'Emilia-Romagna', ',', 'Sardinia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy',  
'\\(', 'Liguria', ',', 'Piedmont', ',', 'Tuscany', ',', 'Lombardy', ',',  
'Emilia-Romagna', ',', 'Sardinia', '\\)']]], 17, True)

['Ligurian', 'Spoken', 'in', 'France', '(', 'Alpes', 'Maritimes', 'and',  
'Corsica', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'France',  
'\\(', 'Alpes', 'Maritimes', 'and', 'Corsica', '\\)']]], 10, True)

['Ligurian', 'Spoken', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ligurian', 'Spoken', 'in', 'Argentina', '(', 'in', 'the', 'neighborhood',  
'of', 'La', 'Boca', 'in', 'Buenos', 'Aires', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Argentina', '\\(', 'in', 'the', 'neighborhood', 'of', 'La', 'Boca', 'in',  
'Buenos', 'Aires', '\\)']]], 15, True)

['Ligurian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

```
['Ligurian', 'Language', 'codes', 'ISO', '639-3', 'lij']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Carian', 'Writing', 'system', 'Carian', 'script']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Carian', 'Language', 'codes', 'ISO', '639-3', 'xcr']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arapaho', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Arapaho', 'Region', 'Wind', 'River', 'Indian', 'Reservation', ',',
'Wyoming', ';', 'Oklahoma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Wind',
'River', 'Indian', 'Reservation', ',', 'Wyoming', ';', 'Oklahoma']], 10,
True)
```

```
['Arapaho', 'Language', 'codes', 'ISO', '639-2', 'arp']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Arapaho', 'Language', 'codes', 'ISO', '639-3', 'arp']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sercquiais', 'Spoken', 'in', 'Sark']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Sercquiais', 'Total', 'speakers', 'approx', '.', '15', ',', 'with',
'several', 'dozen', 'more', 'understanding', 'it']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'approx', '.', 'NUM', ',', 'with', 'several', 'dozen', 'more',
'understanding', 'it']], ['NUM', '15']], 13, True)
```

```
['Sercquiais', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Sercquiais', 'Language', 'codes', 'Linguasphere', '51-AAA-hcf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Auregnais', 'Spoken', 'in', 'Alderney']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Auregnais', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Auregnais', 'Language', 'codes', 'Linguasphere', '51-AAA-hcg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Skepi', 'Spoken', 'in', 'Guyana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Skepi', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Skepi', 'Language', 'codes', 'ISO', '639-3', 'skw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mundolinco', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Mundolinco', 'Created', 'by', 'J', '.', 'Braakman']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',
'.', 'Braakman']]], 6, True)
```

```
['Mundolinco', 'Category', '(', 'purpose', ')', 'Constructed', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Mundolinco', 'Category', '(', 'purpose', ')', 'Esperanto']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY']]], 6, True)
```

```
['Mundolinco', 'Category', '(', 'purpose', ')', 'Mundolinco']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
```

'purpose', '\\)', 'Type: ANY']]], 6, True)

['Mundolinco', 'Date', 'founded', '1888']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mundolinco', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Selkup', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Selkup', 'Region', 'Yamalo-Nenets', 'Autonomous', 'Okrug']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Selkup', 'Language', 'codes', 'ISO', '639-2', 'sel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Selkup', 'Language', 'codes', 'ISO', '639-3', 'sel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nganasan', 'Region', 'Taymyr', 'Autonomous', 'Okrug']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nganasan', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nganasan', 'Pronunciation', '/', 'n', '\\xc3', '\\xa6', '\\xca', '\\x94', '/']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', 'n', '\\xc3', '\\xa6', '\\xca', '\\x94', '/']], 9, True)

['Nganasan', 'Language', 'codes', 'ISO', '639-3', 'nio']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'Wales']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',



'Type: ANY']]], 4, True)

['Middle', 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'wlm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Banyjima', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Banyjima', 'Language', 'codes', 'ISO', '639-3', 'pnw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mazatec', 'Spoken', 'in', 'Oaxaca', ',', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Mazatec', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'maq', '\xe2', '\x80', '\x93', 'Chiquihuitl', '\xc3', '\xa1', 'n', 'Mazatec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'maq', '\xe2', '\x80', '\x93', 'Chiquihuitl', '\xc3', '\xa1', 'n', 'Type: ANY']]], ['NUM', '639-3']], 14, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'mau', '\xe2', '\x80', '\x93', 'Huautla', 'Mazatec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mau', '\xe2', '\x80', '\x93', 'Huautla', 'Type: ANY']]], ['NUM', '639-3']], 11, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'mzi', '\xe2', '\x80', '\x93', 'Ixcatl', '\xc3', '\xa1', 'n', 'Mazatec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mzi', '\xe2', '\x80', '\x93', 'Ixcatl', '\xc3', '\xa1', 'n',

'Type: ANY']], ['NUM', '639-3']], 14, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'vmp', '\xe2', '\x80', '\x93', 'Soyaltepec', 'Mazatec']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'vmp', '\xe2', '\x80', '\x93', 'Soyaltepec', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'vmy', '\xe2', '\x80', '\x93', 'Ayautla', 'Mazatec']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'vmy', '\xe2', '\x80', '\x93', 'Ayautla', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'vmz', '\xe2', '\x80', '\x93', 'Mazatl', '\xc3', '\xa1', 'n', 'Mazatec']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'vmz', '\xe2', '\x80', '\x93', 'Mazatl', '\xc3', '\xa1', 'n', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Itelmen', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Itelmen', 'Region', 'Kamchatka', 'Peninsula']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Itelmen', 'Language', 'codes', 'ISO', '639-3', 'itl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kapingamarangi', 'Total', 'speakers', '1500']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kapingamarangi', 'Spoken', 'in', 'Micronesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kapingamarangi', 'Language', 'codes', 'ISO', '639-3', 'kpg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kangri', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']],

3, True)

```
['Kangri', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kangri', 'Region', 'Himachal', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kangri', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kangri', 'Language', 'codes', 'ISO', '639-3', 'xnr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kerek', 'Total', 'speakers', '0', '(', 'extinct', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'extinct', '\\)']], ['NUM', '0']], 7, True)
```

```
['Kerek', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kerek', 'Region', 'Chukotka', 'Autonomous', 'Okrug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kerek', 'Language', 'codes', 'ISO', '639-3', 'krk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Are', 'Language', 'codes', 'ISO', '639-3', 'mwc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Qwara', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Qwara', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Qwara', 'Region', 'Amhara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Qwara', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Qwara', 'Language', 'codes', 'ISO', '639-3', 'ahg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Megleno-Romanian', 'Spoken', 'in', 'Greece']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Megleno-Romanian', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']]], 6, True)

['Megleno-Romanian', 'Spoken', 'in', 'Romania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Megleno-Romanian', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Megleno-Romanian', 'Region', 'Balkans']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Megleno-Romanian', 'Language', 'codes', 'ISO', '639-3', 'ruq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Megleno-Romanian', 'Language', 'codes', 'Linguasphere', '51-AAD-bb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Tartessian', 'Language', 'codes', 'ISO', '639-3', 'txr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Min', 'Spoken', 'in', 'Southern', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Min', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Min', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Min', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Min', 'Language', 'codes', 'ISO', '639-3', 'cdo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Min', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Phuthi', 'Region', 'Southern', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Phuthi', 'Language', 'codes', 'ISO', '639-3', 'ssw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chilcotin', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chilcotin', 'Language', 'codes', 'ISO', '639-3', 'clc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Miskito', 'Spoken', 'in', 'Nicaragua', ',', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Miskito', 'Language', 'codes', 'ISO', '639-3', 'miq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Sumo', 'Spoken', 'in', 'Nicaragua', ',', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Sumo', 'Region', 'Huaspuc', 'River', 'and', 'its', 'tributaries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Huaspuc',  
'River', 'and', 'its', 'tributaries']]], 7, True)
```

```
['Sumo', 'Language', 'codes', 'ISO', '639-3', 'sum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cacaopera', 'Spoken', 'in', 'El', 'Salvador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Cacaopera', 'Region', 'Moraz', '\xc3', '\xa1', 'n', 'Department']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Moraz',  
'\xc3', '\xa1', 'n', 'Department']]], 7, True)
```

```
['Cacaopera', 'Language', 'codes', 'ISO', '639-3', 'ccr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Matagalpa', 'Spoken', 'in', 'Nicaragua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Matagalpa', 'Region', 'Central', 'Highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Matagalpa', 'Language', 'codes', 'ISO', '639-3', 'mtn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Karuk', 'Language', 'codes', 'ISO', '639-3', 'kyh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kusaal', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kusaal', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['Kusaal', 'Language', 'codes', 'ISO', '639-3', 'kus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koba\xc3\xafan', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Koba\xc3\xafan', 'Date', 'founded', '1969']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koba\xc3\xafan', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Koba\xc3\xafan', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Koba\xc3\xafan', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Koba\xc3\xafan', 'Category', '(', 'purpose', ')', 'Koba', '\xc3', '\xaf', 'an']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Koba', '\xc3', '\xaf', 'an']]], 9, True)

['Koba\xc3\xafan', 'Created', 'by', 'Christian', 'Vander']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Koba\xc3\xafan', 'Category', '(', 'sources', ')', 'Based', 'on', 'elements', 'of', 'Slavic-Germanic', 'languages', 'and', 'the', 'scat-yodelling', 'vocal', 'style', 'of', 'Leon', 'Thomas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'Based', 'on', 'elements', 'of', 'Slavic-Germanic', 'languages', 'and', 'the', 'scat-yodelling', 'vocal', 'style', 'of', 'Leon', 'Thomas']]], 19, True)

['Koba\xc3\xafan', 'Setting', 'and', 'usage', 'Lyrics', 'for', 'a',  
'musical', 'group']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Lyrics', 'for', 'a', 'musical', 'group']]], 9, True)

['Koba\xc3\xafan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Sidamo', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sidamo', 'Language', 'codes', 'ISO', '639-2', 'sid']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sidamo', 'Language', 'codes', 'ISO', '639-3', 'sid']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Deg', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Deg', 'Region', 'Alaska', '(', 'lower', 'Yukon', 'River', ',', 'Anvik',  
'River', ',', 'Innoko', 'River', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\  
(', 'lower', 'Yukon', 'River', ',', 'Anvik', 'River', ',', 'Innoko',  
'River', '\\)']]], 14, True)

['Deg', 'Language', 'codes', 'ISO', '639-3', 'ing']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nahuatl', 'Spoken', 'in', 'Mexico', ':', 'M', '\xc3', '\xa9', 'xico',  
'(', 'state', ')', ',', 'Distrito', 'Federal', ',', 'Puebla', ',',  
'Veracruz', ',', 'Hidalgo', ',', 'Guerrero', ',', 'Morelos', ',', 'Oaxaca',  
'', 'Michoac', '\xc3', '\xa1', 'n', 'and', 'Durango']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  
'Federal', ',', 'Puebla', ',', 'Veracruz', ',', 'Hidalgo', ',', 'Guerrero',  
'', 'Morelos', ',', 'Oaxaca', ',', 'Michoac', '\xc3', '\xa1', 'n', 'and',  
'Durango']]], 34, True)



```
['Nahuatl', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Nahuatl', 'Regulated', 'by', 'Instituto', 'Nacional', 'de', 'Lenguas',  
'Ind', '\xc3', '\xad', 'genas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Instituto', 'Nacional', 'de', 'Lenguas', 'Ind', '\xc3', '\xad',  
'genas']]], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'naz', '\xe2', '\x80',  
'\x93', 'Coatepec', 'Nahuatl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'naz', '\xe2', '\x80', '\x93', 'Coatepec', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nln', '\xe2', '\x80',  
'\x93', 'Durango', 'Nahuatl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nln', '\xe2', '\x80', '\x93', 'Durango', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'ngu', '\xe2', '\x80',  
'\x93', 'Guerrero', 'Nahuatl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ngu', '\xe2', '\x80', '\x93', 'Guerrero', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhq', '\xe2', '\x80',  
'\x93', 'Huaxcaleca', 'Nahuatl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nhq', '\xe2', '\x80', '\x93', 'Huaxcaleca', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhk', '\xe2', '\x80',  
'\x93', 'Isthmus-Cosoleacaque', 'Nahuatl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nhk', '\xe2', '\x80', '\x93', 'Isthmus-Cosoleacaque', 'Type:  
ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhx', '\xe2', '\x80',  
'\x93', 'Isthmus-Mecayapan', 'Nahuatl']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nhx', '\xe2', '\x80', '\x93', 'Isthmus-Mecayapan', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhp', '\xe2', '\x80', '\x93', 'Isthmus-Pajapan', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nhp', '\xe2', '\x80', '\x93', 'Isthmus-Pajapan', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'ncl', '\xe2', '\x80', '\x93', 'Michoac', '\xc3', '\xa1', 'n', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ncl', '\xe2', '\x80', '\x93', 'Michoac', '\xc3', '\xa1', 'n', 'Type: ANY']], ['NUM', '639-3']], 14, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhm', '\xe2', '\x80', '\x93', 'Morelos', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nhm', '\xe2', '\x80', '\x93', 'Morelos', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nht', '\xe2', '\x80', '\x93', 'Ometepec', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nht', '\xe2', '\x80', '\x93', 'Ometepec', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhc', '\xe2', '\x80', '\x93', 'Tabasco', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nhc', '\xe2', '\x80', '\x93', 'Tabasco', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhv', '\xe2', '\x80', '\x93', 'Temascaltepec', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nhv', '\xe2', '\x80', '\x93', 'Temascaltepec', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhi', '\xe2', '\x80', '\x93', 'Tenango', 'Nahuatl']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nhi', '\xe2', '\x80', '\x93', 'Tenango', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhj', '\xe2', '\x80',  
\x93', 'Tlalitzlipa', 'Nahuatl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nhj', '\xe2', '\x80', '\x93', 'Tlalitzlipa', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nuz', '\xe2', '\x80',  
\x93', 'Tlamacazapa', 'Nahuatl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nuz', '\xe2', '\x80', '\x93', 'Tlamacazapa', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Pattani', 'Spoken', 'in', 'Thailand']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pattani', 'Region', 'Pattani', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pattani', 'Language', 'codes', 'ISO', '639-2', 'may', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'may', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Pattani', 'Language', 'codes', 'ISO', '639-2', 'msa', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'msa', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Pattani', 'Language', 'codes', 'ISO', '639-3', 'mfa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pattani', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Seri', 'Pronunciation', 'kw', '\xcc', '\x83', '\xc4', '\xa9', 'k',  
\xcb', '\x88', 'i', '\xcb', '\x90', 'tom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kw',  
'\xcc', '\x83', '\xc4', '\xa9', 'k', '\xcb', '\x88', 'i', '\xcb', '\x90',  
'tom']]], 14, True)
```

```
['Seri', 'Spoken', 'in', 'Sonora', ',', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```

['Seri', 'Language', 'codes', 'ISO', '639-3', 'sei']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Spoken', 'in', 'Mexico']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Classical', 'Region', 'Aztec', 'Empire']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'nah']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'nci']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Moriari', 'Total', 'speakers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Moriari', 'Spoken', 'in', 'New', 'Zealand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Moriari', 'Region', 'Polynesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Moriari', 'Writing', 'system', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Moriari', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Moriari', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Teonaht', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

```

'\xe2', '\x80', '\x93']]], 6, True)

['Teonaht', 'Date', 'founded', '1962']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Teonaht', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Teonaht', 'Category', '(', 'purpose', ')', 'artistic', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Teonaht', 'Category', '(', 'purpose', ')', 'fictional', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Teonaht', 'Category', '(', 'purpose', ')', 'Teonaht']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)

['Teonaht', 'Created', 'by', 'Sally', 'Caves']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Teonaht', 'Category', '(', 'sources', ')', 'draws', 'on', 'Indo-European', 'languages', ':', 'Romance', ',', 'Germanic', 'and', 'Celtic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'draws', 'on', 'Indo-European', 'languages', ':', 'Romance', ',', 'Germanic', 'and', 'Celtic']], 15, True)

['Teonaht', 'Setting', 'and', 'usage', 'Fantasy', 'setting', 'of', 'the', 'Teonim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Fantasy', 'setting', 'of', 'the', 'Teonim']], 9, True)

['Teonaht', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Monguor', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Monguor', 'Region', 'Qinghai', ',', 'Gansu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Monguor', 'Language', 'codes', 'ISO', '639-3', 'mjb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yagua', 'Spoken', 'in', 'western', 'Amazon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Yagua', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Yagua', 'Language', 'codes', 'ISO', '639-3', 'yad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kangjia', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kangjia', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kangjia', 'Region', 'Qinghai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kangjia', 'Language', 'codes', 'ISO', '639-3', 'kxs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khowar', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khowar', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khowar', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khowar', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khowar', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khowar', 'Official', 'language', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Khowar', 'Official', 'language', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Khowar', 'Regulated', 'by', 'Khowar', 'Academy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Khowar', 'Regulated', 'by', 'Chitral', '(', 'Pakistan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', '\\\\(', 'Pakistan', '\\\\)']], 7, True)

['Khowar', 'Language', 'codes', 'ISO', '639-3', 'khw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khowar', 'Language', 'codes', 'Linguasphere', '59-AAB-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Fuzhou', 'Spoken', 'in', 'Miri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fuzhou', 'Spoken', 'in', 'Sarikei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fuzhou', 'Spoken', 'in', 'Bintulu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fuzhou', 'Spoken', 'in', 'Yong', 'Peng']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Fuzhou', 'Total', 'speakers', 'Less', 'than', 'ten', 'million']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Less', 'than', 'ten', 'million']]], 7, True)

['Fuzhou', 'Writing', 'system', 'Chinese', 'characters', 'and', 'Foochow', 'Romanized']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Fuzhou', 'Official', 'language', 'in', 'none', ';', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public', 'transport', 'announcements', 'in', 'the', 'Matsu', 'Islands', ',', 'Republic', 'of', 'China']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'none', ';', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public', 'transport', 'announcements', 'in', 'the', 'Matsu', 'Islands', ',', 'Republic', 'of', 'China']]], 23, True)

['Fuzhou', 'Regulated', 'by', 'No', 'official', 'regulation']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', 'cdo-fuz', 'Foochowese', 'in', 'Fujian', 'Province', ',', 'regions', 'where', 'the', 'standard', 'form', 'is', 'spoken', 'are', 'deep', 'blue']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cdo-fuz', 'Foochowese', 'in', 'Fujian', 'Province', ',', 'regions', 'where', 'the', 'standard', 'form', 'is', 'spoken', 'are', 'deep', 'blue']]], ['NUM', '639-3']], 21, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '15', ':', 'Regions', 'in', 'Nanping', ',', '16', ':', 'Regions', 'in', 'Youxi']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Regions', 'in', 'Nanping', ',', 'NUM', ':', 'Regions', 'in', 'Youxi']]], ['NUM', '639-3'], ['NUM', '15'], ['NUM', '16']], 16, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '2', ':', 'Minhou']:



```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '2']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '3', ':', 'Fuqing']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '3']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '4', ':', 'Lianjiang']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '4']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '5', ':', 'Pingnan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '5']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '6', ':', 'Luoyuan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '6']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '7', ':', 'Gutian']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '7']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '8', ':', 'Minqing']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '8']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '9', ':', 'Changle']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '9']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '10', ':', 'Yongtai']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '10']], 8, True)
```

```
['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '11', ':', 'Pingtan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'NUM', ':', 'Type: ANY']], ['NUM', '639-3'], ['NUM', '11']], 8, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '12', ':', 'Regions', 'in', 'Fuding']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Regions', 'in', 'Fuding']], ['NUM', '639-3'], ['NUM', '12']], 10, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '13', ':', 'Regions', 'in', 'Xiapu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Regions', 'in', 'Xiapu']], ['NUM', '639-3'], ['NUM', '13']], 10, True)

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '14', ':', 'Regions', 'in', 'Ningde']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'NUM', ':', 'Regions', 'in', 'Ningde']], ['NUM', '639-3'], ['NUM', '14']], 10, True)

['Dameli', 'Spoken', 'in', 'Pakistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dameli', 'Region', 'Khyber', 'Pakhtunkhwa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dameli', 'Language', 'codes', 'ISO', '639-3', 'dml']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalasha', 'Spoken', 'in', 'Pakistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalasha', 'Region', 'Pakistan', ':', 'Chitral']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kalasha', 'Region', 'Bumburet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kalasha', 'Region', 'Rumbur', 'and', 'Birir']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kalasha', 'Region', 'Khyber', 'Pakhtunkhwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalasha', 'Language', 'codes', 'ISO', '639-3', 'kls']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalasha', 'Language', 'codes', 'Linguasphere', '59-AAB-ab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Palula', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Palula', 'Language', 'codes', 'ISO', '639-3', 'phl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gawar-Bati', 'Spoken', 'in', 'Afghanistan', ',', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Gawar-Bati', 'Region', 'Kunar', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gawar-Bati', 'Language', 'codes', 'ISO', '639-3', 'gwt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yidgha', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yidgha', 'Writing', 'system', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yidgha', 'Language', 'codes', 'ISO', '639-3', 'ydg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Yidgha', 'Language', 'codes', 'Linguasphere', '58-ABD-bb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Munji', 'Spoken', 'in', 'Badakhshan', 'Province', ',', 'Afghanistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Badakhshan', 'Province', ',', 'Afghanistan']]], 7, True)
```

```
['Munji', 'Total', 'speakers', 'Not', 'known']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Not', 'known']]], 5, True)
```

```
['Munji', 'Writing', 'system', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Munji', 'Language', 'codes', 'ISO', '639-3', 'mnj']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Munji', 'Language', 'codes', 'Linguasphere', '58-ABD-ba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Panare', 'Spoken', 'in', 'Venezuela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Panare', 'Region', 'just', 'south', 'of', 'the', 'Orinoco', 'River', ',',  
'Estado', 'Bol', '\xc3', '\xad', 'var']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'just',  
'south', 'of', 'the', 'Orinoco', 'River', ',', 'Estado', 'Bol', '\xc3',  
'\xad', 'var']]], 14, True)
```

```
['Panare', 'Language', 'codes', 'ISO', '639-3', 'pbh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Brij', 'Region', 'Uttar', 'Pradesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Brij', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```

['Brij', 'Writing', 'system', 'Devanagari', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Brij', 'Language', 'codes', 'ISO', '639-2', 'bra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Brij', 'Language', 'codes', 'ISO', '639-3', 'bra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamassian', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kamassian', 'Total', 'speakers', '0']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kamassian', 'Language', 'codes', 'ISO', '639-3', 'xas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mator', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mator', 'Total', 'speakers', '0']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mator', 'Dialects', 'Mator']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mator', 'Dialects', 'Taygi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mator', 'Dialects', 'Karagas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mator', 'Language', 'codes', 'ISO', '639-3', 'mtm']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Angika', 'Spoken', 'in', 'India', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Angika', 'Region', 'Bihar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Angika', 'Region', 'Jharkhand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Angika', 'Region', 'West', 'Bengal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Angika', 'Language', 'codes', 'ISO', '639-2', 'anp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Angika', 'Language', 'codes', 'ISO', '639-3', 'anp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Angika', 'Language', 'codes', 'ISO', '639-1', 'bh', '(', 'Bihari', 'languages', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bh', '\\\\(', 'Bihari', 'languages', '\\\\)']], ['NUM', '639-1']], 10, True)

['Kambera', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']], 8, True)

['Kambera', 'Dialects', 'Rua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kambera', 'Language', 'codes', 'ISO', '639-3', 'wnk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Algerian', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Algerian', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Algerian', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Algerian', 'Language', 'codes', 'ISO', '639-3', 'arq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Fuy\xc3\xbc', 'Pronunciation', 'q', '\xc9', '\x99', 'r', '\xca', '\x81',  
'\xc9', '\x99', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'q',  
'\xc9', '\x99', 'r', '\xca', '\x81', '\xc9', '\x99', 's']]], 11, True)
```

```
['Fuy\xc3\xbc', 'Region', 'Heilongjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Fuy\xc3\xbc', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Fuy\xc3\xbc', 'Language', 'codes', 'ISO', '639-3', 'kjh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nyoro', 'Standard', 'forms', 'Kitara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nyoro', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nyoro', 'Language', 'codes', 'ISO', '639-3', 'nyo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Babm', 'Total', 'speakers', 'none']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Babm', 'Date', 'founded', '1962']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Babm', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Babm', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'int', '.', 'auxiliary', 'languages']]], 9, True)

['Babm', 'Category', '(', 'purpose', ')', 'Babm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Babm', 'Created', 'by', 'Rikichi', '(', 'Fuishiki', ')', 'Okamoto']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Rikichi', '\\(', 'Fuishiki', '\\)', 'Okamoto']]], 8, True)

['Babm', 'Category', '(', 'sources', ')', 'a', 'priori', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'a', 'priori', 'language']]], 8, True)

['Babm', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Babm', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Kutenai', 'Total', 'speakers', '12']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kutenai', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kutenai', 'Region', 'British', 'Columbia']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kutenai', 'Language', 'codes', 'ISO', '639-3', 'kut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Carrier', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Carrier', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Carrier', 'Language', 'codes', 'ISO', '639-3', ':', 'crx', '\xe2', '\x80', '\x93', 'Carrier']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'crx', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM', '639-3']], 11, True)

['Kaqchikel', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kaqchikel', 'Region', 'Central', 'Highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kaqchikel', 'Language', 'codes', 'ISO', '639-3', 'cak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Atakapa', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Atakapa', 'Region', 'Louisiana', ',', 'Texas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Atakapa', 'Language', 'codes', 'ISO', '639-3', 'aqp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chitimacha', 'Language', 'extinction', '1940']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chitimacha', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chitimacha', 'Region', 'Southern', 'Louisiana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chitimacha', 'Language', 'codes', 'ISO', '639-3', 'ctm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Esselen', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Esselen', 'Language', 'codes', 'ISO', '639-3', 'esq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalapuya', 'Spoken', 'in', 'Northwest', 'Oregon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kalapuya', 'Total', 'speakers', '1', 'or', '2', ';', 'effectively', 'extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'or', 'NUM', ';', 'effectively', 'extinct']], ['NUM', '1'], ['NUM', '2']], 9, True)

['Kalapuya', 'Language', 'codes', 'ISO', '639-3', 'kyl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karankawa', 'Language', 'extinction', '1858']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Karankawa', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Karankawa', 'Region', 'Karankawa', 'people']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Karankawa', 'Language', 'codes', 'ISO', '639-3', 'zkk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kaw\xc3\xa9sqr', 'Total', 'speakers', '22', 'or', 'fewer']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'or', 'fewer']]], ['NUM', '22']], 6, True)

['Kaw\xc3\xa9sqr', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kaw\xc3\xa9sqr', 'Region', 'Channel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kaw\xc3\xa9sqr', 'Region', 'western', 'Patagonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kaw\xc3\xa9sqr', 'Region', '49', '\xc2', '\xb0', 'south']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'NUM', '\xc2', '\xb0', 'south']]], ['NUM', '49']], 6, True)

['Kaw\xc3\xa9sqr', 'Region', 'with', 'centre', 'in', 'Puerto', 'Ed', '\xc3', '\xa9', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'with', 'centre', 'in', 'Puerto', 'Ed', '\xc3', '\xa9', 'n']]], 10, True)

['Kaw\xc3\xa9sqr', 'Language', 'codes', 'ISO', '639-3', 'alc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Indo-Pakistani', 'Region', 'al', 'dialects']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Indo-Pakistani', 'Signed', 'in', 'Pakistan', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'Pakistan', ',', 'India']]], 6, True)

['Indo-Pakistani', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Indo-Pakistani', 'Language', 'codes', 'ISO', '639-3', ':', 'ins', '\xe2', '\x80', '\x93', 'Indo-Bengali', 'Sign', 'Language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ins', '\xe2', '\x80', '\x93', 'Indo-Bengali', 'Sign', 'Type: ANY']], ['NUM', '639-3']], 13, True)

['Leonese', 'Spoken', 'in', 'Spain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Leonese', 'Spoken', 'in', 'Portugal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Leonese', 'Region', 'Provinces', 'of', 'Le', '\xc3', '\xb3', 'n', '(', 'north', 'and', 'west', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Provinces', 'of', 'Le', '\xc3', '\xb3', 'n', '\\(', 'north', 'and', 'west', '\\)']], 13, True)

['Leonese', 'Region', 'Zamora', '(', 'north-west', ')', 'in', 'Spain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Zamora', '\\(', 'north-west', '\\)', 'in', 'Spain']], 8, True)

['Leonese', 'Region', 'and', 'the', 'towns', 'of', 'Rionor', 'and', 'Guadramil', 'in', 'northeastern', 'Portugal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'the', 'towns', 'of', 'Rionor', 'and', 'Guadramil', 'in', 'northeastern', 'Portugal']], 12, True)

['Leonese', 'Official', 'language', 'in', 'Special', 'protection', 'status', 'in', 'Castile', 'and', 'Leon', ',', 'Spain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Special', 'protection', 'status', 'in', 'Castile', 'and', 'Leon', ',', 'Spain']], 13, True)

['Leonese', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Leonese', 'Language', 'codes', 'ISO', '639-2', 'ast']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Leonese', 'Language', 'codes', 'ISO', '639-3', 'ast']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Leonese', 'Language', 'codes', 'Linguasphere', '51-AAA-cc']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]]], 5, True)
```

```
['Toda', 'Spoken', 'in', 'India']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['Toda', 'Region', 'Nilgiri', 'Hills']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['Toda', 'Language', 'codes', 'ISO', '639-3', 'tcx']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Malvani', 'Region', 'Konkan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]]], 3, True)
```

```
['Malvani', 'Spoken', 'in', 'India']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['Malvani', 'Writing', 'system', 'Devanagari']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['Malvani', 'Pronunciation', 'malva', '\xc9', '\xb3', 'i', '(', 'standard',
')', ',', 'malvani', '(', 'popular', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'malva',
'\xc9', '\xb3', 'i', '\\(', 'standard', '\\)', ',', 'malvani', '\\(',
'popular', '\\)']]]], 14, True)
```

```
['Malvani', 'Language', 'codes', 'ISO', '639-2', 'kok']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Malvani', 'Language', 'codes', 'ISO', '639-3', 'gom']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wari\xe2\x80\x99', 'Total', 'speakers', '1300-1800']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Wari\xe2\x80\x99', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Wari\xe2\x80\x99', 'Region', 'Amazon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Wari\xe2\x80\x99', 'Language', 'codes', 'ISO', '639-3', 'pav']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Neo', 'Total', 'speakers', 'Unknown']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Neo', 'Date', 'founded', '1961']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Neo', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Neo', 'Category', '(', 'purpose', ')', 'Neo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)
```

```
['Neo', 'Created', 'by', 'Arturo', 'Alfandari']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Neo', 'Category', '(', 'sources', ')', 'Most', 'of', 'the', 'vocabulary',  
'from', 'Romance', 'and', 'some', 'from', 'Germanic', 'languages', ';',  
'phonology', 'from', 'Romance', 'and', 'Slavic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Most', 'of', 'the', 'vocabulary', 'from', 'Romance',  
'and', 'some', 'from', 'Germanic', 'languages', ';', 'phonology', 'from',  
'Romance', 'and', 'Slavic', 'languages']]], 23, True)
```

```
['Neo', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Neo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Oirata', 'Spoken', 'in', 'Kisar', ',', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Oirata', 'Total', 'speakers', '1200']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Oirata', 'Language', 'codes', 'ISO', '639-3', 'oia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Martuthunira', 'Spoken', 'in', 'Western', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Martuthunira', 'Language', 'codes', 'ISO', '639-3', 'vma']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gutob', 'Spoken', 'in', 'India', '(', 'Orissa', ',', 'Andhra', 'Pradesh',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
'\\(', 'Orissa', ',', 'Andhra', 'Pradesh', '\\)']]], 10, True)

['Gutob', 'Language', 'codes', 'ISO', '639-3', 'gbj']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paic\xc3\xae', 'Spoken', 'in', 'New', 'Caledonia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Paic\xc3\xae', 'Language', 'codes', 'ISO', '639-3', 'pri']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tagish', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tagish', 'Total', 'speakers', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tagish', 'Language', 'codes', 'ISO', '639-3', 'tgx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shor', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shor', 'Region', 'Kemerovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Shor', 'Language', 'codes', 'ISO', '639-3', 'cjs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Afrihili', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Afrihili', 'Date', 'founded', '1970']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afrihili', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Afrihili', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'int', '.', 'auxiliary', 'languages']], 9, True)

['Afrihili', 'Category', '(', 'purpose', ')', 'Afrihili']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)

['Afrihili', 'Created', 'by', 'K', '.', 'A', '.', 'Kumi', 'Attobrah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'K', '.', 'A', '.', 'Kumi', 'Attobrah']], 9, True)



['Afrihili', 'Category', '(', 'sources', ')', 'a', 'posteriori',  
'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'a', 'posteriori', 'language']]], 8, True)

['Afrihili', 'Category', '(', 'sources', ')', 'notably', 'Swahili']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Type: ANY', 'Type: ANY']]], 7, True)

['Afrihili', 'Setting', 'and', 'usage', 'intended', 'for', 'use',  
'throughout', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'intended', 'for', 'use', 'throughout', 'Africa']]], 9, True)

['Afrihili', 'Language', 'codes', 'ISO', '639-2', 'afh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Afrihili', 'Language', 'codes', 'ISO', '639-3', 'afh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shina', 'Total', 'speakers', '321', ',', '000', '(', 'total', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\\\(', 'total', '\\\\)']], ['NUM', '321'], ['NUM',  
'000']], 9, True)

['Shina', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Shina', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Shina', 'Region', 'Jammu', 'and', 'Kashmir', ',', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jammu', 'and',  
'Kashmir', ',', 'India']]], 7, True)

['Shina', 'Writing', 'system', 'Arabic', 'Script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Shina', 'Language', 'codes', 'ISO', '639-3', 'scl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bungi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bungi', 'Region', 'Red', 'River', 'Colony', 'and', 'Assiniboia', ',', 'present-day', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Red', 'River', 'Colony', 'and', 'Assiniboia', ',', 'present-day', 'Manitoba']], 10, True)

['Bungi', 'Language', 'codes', 'ISO', '639-2', 'cpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bungi', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dahalik', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dahalik', 'Region', 'Dahlak', 'Archipelago']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dahalik', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Genoese', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Genoese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Genoese', 'Region', 'Genoa', ',', 'Liguria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Genoese', 'Language', 'codes', 'ISO', '639-3', 'lij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsuut\xe2\x80\x99ina', 'Total', 'speakers', '50']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tsuut\xe2\x80\x99ina', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tsuut\xe2\x80\x99ina', 'Region', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Tsuut\xe2\x80\x99ina', 'Language', 'codes', 'ISO', '639-3', 'srs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tuscarora', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Tuscarora', 'Region', 'Six', 'Nations', 'of', 'the', 'Grand', 'River',  
'First', 'Nation', 'in', 'southern', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Six',  
'Nations', 'of', 'the', 'Grand', 'River', 'First', 'Nation', 'in',  
'southern', 'Ontario']]], 13, True)

['Tuscarora', 'Language', 'codes', 'ISO', '639-3', 'tus', 'Pre-contact',  
'distribution', 'of', 'Tuscarora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tus', 'Pre-contact', 'distribution', 'of', 'Type: ANY']],  
['NUM', '639-3']], 10, True)

['Tahltan', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tahltan', 'Total', 'speakers', '450']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tahltan', 'Language', 'codes', 'ISO', '639-3', 'tth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Seneca', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Seneca', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Seneca', 'Language', 'codes', 'ISO', '639-3', 'see']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Onondaga', 'Total', 'speakers', '65-115']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Onondaga', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Onondaga', 'Region', 'Six', 'Nations', 'Reserve']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Onondaga', 'Region', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Onondaga', 'Language', 'codes', 'ISO', '639-3', 'ono']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Canadian', 'Region', 'Nova', 'Scotia', ';', 'scattered', 'traditional', 'speakers', 'in', 'mainland', 'Nova', 'Scotia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nova', 'Scotia', ';', 'scattered', 'traditional', 'speakers', 'in', 'mainland', 'Nova', 'Scotia']]], 12, True)

['Canadian', 'Region', 'Prince', 'Edward', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Canadian', 'Region', 'Ontario', '(', 'Glengarry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Canadian', 'Region', 'Stormont']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

ANY'']], 3, True)

['Canadian', 'Region', 'The', 'Red', 'River', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'The', 'Red',  
'River', 'Valley']]], 6, True)

['Canadian', 'Region', 'Assiniboia', 'and', 'The', 'Northwest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Assiniboia',  
'and', 'The', 'Northwest']]], 6, True)

['Canadian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Canadian', 'Pronunciation', '\xc9', '\x99', '\xcb', '\x88', '\xc9',  
'\xa3', 'a', '\xcb', '\x90', 'l', '\xc9', '\xaa', 'k', '\xca', '\xb2',  
'\xcb', '\x88', 'xan', '\xc9', '\x99', 't', '\xcc', '\xaa', '\xc9', '\x99',  
'x']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',  
'\x99', '\xcb', '\x88', '\xc9', '\xa3', 'a', '\xcb', '\x90', 'l', '\xc9',  
'\xaa', 'k', '\xca', '\xb2', '\xcb', '\x88', 'xan', '\xc9', '\x99', 't',  
'\xcc', '\xaa', '\xc9', '\x99', 'x']]], 28, True)

['Canadian', 'Language', 'codes', 'ISO', '639-2', 'gl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Canadian', 'Language', 'codes', 'ISO', '639-3', 'gl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Canadian', 'Language', 'codes', 'ISO', '639-1', 'gd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Chenchu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chenchu', 'Region', 'Andhra', 'Pradesh', '(', 'highest', 'concentration',  
'in', 'Kurnool', 'district', ')', ',', 'Karnataka', 'and', 'Orissa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Andhra',  
'Pradesh', '\\(', 'highest', 'concentration', 'in', 'Kurnool', 'district',  
'\\)', ',', 'Karnataka', 'and', 'Orissa']]], 15, True)

['Chenchu', 'Writing', 'system', 'Telugu', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chenchu', 'Language', 'codes', 'ISO', '639-3', 'cde']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Khalaj', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khalaj', 'Region', 'Northeast', 'of', 'Arak', 'in', 'Markazi', 'Province', 'of', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Northeast', 'of', 'Arak', 'in', 'Markazi', 'Province', 'of', 'Iran']]], 10, True)

['Khalaj', 'Language', 'codes', 'ISO', '639-3', 'klj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['\xc7\x83Kung', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['\xc7\x83Kung', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['\xc7\x83Kung', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['\xc7\x83Kung', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', ':', 'oun', '\xe2', '\x80', '\x93', '\xc7', '\x83', '\xca', '\xbc', '0', '\xc7', '\x83', 'Kung']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'oun', '\xe2', '\x80', '\x93', '\xc7', '\x83', '\xca', '\xbc', '0', '\xc7', '\x83', 'Kung']]], ['NUM', '639-3']], 18, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', 'mwj', '\xe2',  
\x80', '\x93', 'Maligo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mwj', '\xe2', '\x80', '\x93', 'Maligo']], ['NUM', '639-3']],  
10, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', 'knw', '\xe2',  
\x80', '\x93', '\xc7', '\x80', '\xca', '\xbc', 'Akhwe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'knw', '\xe2', '\x80', '\x93', '\xc7', '\x80', '\xca',  
\xbc', 'Akhwe']], ['NUM', '639-3']], 14, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', 'vaj', '\xe2',  
\x80', '\x93', 'Vasekela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'vaj', '\xe2', '\x80', '\x93', 'Vasekela']], ['NUM',  
'639-3']], 10, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', 'ktz', '\xe2',  
\x80', '\x93', 'Ju', '\xc7', '\x80', '\xca', '\xbc', 'hoansi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ktz', '\xe2', '\x80', '\x93', 'Ju', '\xc7', '\x80', '\xca',  
\xbc', 'hoansi']], ['NUM', '639-3']], 15, True)

['\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3', 'aue', '\xe2',  
\x80', '\x93', '\xc7', '\x82', 'Kx', '\xca', '\xbc', 'au', '\xc7', '\x81',  
\xca', '\xbc', 'ein']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'aue', '\xe2', '\x80', '\x93', '\xc7', '\x82', 'Kx', '\xca',  
\xbc', 'au', '\xc7', '\x81', '\xca', '\xbc', 'ein']], ['NUM', '639-3']],  
20, True)

['Middle', 'Spoken', 'in', 'Scotland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Middle', 'Spoken', 'in', 'Isle', 'of', 'Man']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle',  
'of', 'Man']], 6, True)

['Middle', 'Writing', 'system', 'Latin', '(', 'Gaelic', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '\\(', 'Gaelic', 'variant', '\\)']], 8, True)

['Middle', 'Pronunciation', '\xcb', '\x88', '\xc9', '\xa1', '\xc9', '\xaf',

'\xcb', '\x90', '\xca', '\x9d', 'e', '\xc9', '\xab', '\xc9', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', '\xc9', '\xa1', '\xc9', '\xaf', '\xcb', '\x90', '\xca', '\x9d',  
'e', '\xc9', '\xab', '\xc9', '\xa1']]], 17, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'mga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'mga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Halang', 'Spoken', 'in', 'Kon', 'Tum', 'Province', '(', 'Vietnam', ')',  
,',', 'Attapu', 'Province', '(', 'Laos', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kon',  
'Tum', 'Province', '\\(', 'Vietnam', '\\)', ',,', 'Attapu', 'Province', '\\(  
,', 'Laos', '\\)']]], 15, True)

['Halang', 'Language', 'codes', 'ISO', '639-3', 'hal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sara', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sara', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sara', 'Region', 'near', 'Sanggau-Ledo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sara', 'Region', 'northeast', 'of', 'Ledo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sara', 'Region', 'Kalimantan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Sara', 'Language', 'codes', 'ISO', '639-3', 'sre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Sika', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Sika', 'Region', 'Flores']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Sika', 'Language', 'codes', 'ISO', '639-3', 'ski']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Udi', 'Total', 'speakers', '8', ',', '000', '(', 'est', '.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'est', '.', '\\)']], ['NUM', '8'], ['NUM',  
'000']], 10, True)

['Udi', 'Spoken', 'in', 'Azerbaijan', ',', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Udi', 'Region', 'Azerbaijan', '(', 'Qabala', 'and', 'Oguz', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Azerbaijan',  
'\\(', 'Qabala', 'and', 'Oguz', '\\)']], 8, True)

['Udi', 'Region', 'Georgia', '(', 'Kvareli', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Georgia', '\\(  
, 'Kvareli', '\\)']], 6, True)

['Udi', 'Region', 'and', 'Armenia', '(', 'Tavush', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Armenia', '\\(', 'Tavush', '\\)']], 7, True)

['Udi', 'Language', 'codes', 'ISO', '639-3', 'udi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lara\xe2\x80\x99', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lara\xe2\x80\x99', 'Region', 'Borneo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

```
['Lara\xe2\x80\x99', 'Language', 'codes', 'ISO', '639-3', 'lra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Parthian', 'Spoken', 'in', 'Ancient', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Parthian', 'Language', 'codes', 'ISO', '639-3', 'xpr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Weitou', 'Spoken', 'in', 'Southern', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Weitou', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Weitou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Swampy', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Swampy', 'Region', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Swampy', 'Language', 'codes', 'ISO', '639-2', 'cre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Swampy', 'Language', 'codes', 'ISO', '639-3', 'csw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lithuanian', 'Region', 'Limited', 'to', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Lithuanian', 'Signed', 'in', 'Lithuanian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Lithuanian', 'Total', 'signers', '(', 'unknown', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\  
(', 'unknown', '\\\')']]], 6, True)

['Sukur', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sukur', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sukur', 'Language', 'codes', 'ISO', '639-3', 'syk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Haida', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Haida', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Haida', 'Language', 'codes', 'ISO', '639-2', 'hai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Haida', 'Language', 'codes', 'ISO', '639-3', 'hai', '\\xe2', '\\x80',  
'\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hai', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Haida', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Haida', 'Language', 'codes', 'ISO', '639-3', 'Pre-European',  
'distribution', 'of', 'Haida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Pre-European', 'distribution', 'of', 'Type: ANY']], ['NUM',  
'639-3']], 9, True)

['Lari', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lari', 'Spoken', 'in', 'Kuwait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lari', 'Spoken', 'in', 'Qatar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lari', 'Spoken', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3, True)

['Lari', 'Language', 'codes', 'ISO', '639-3', 'lrl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Northeast', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['South', 'Region', '\xc3', '\x89', 'fat', '\xc3', '\xa9', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', '\xc3', '\x89', 'fat', '\xc3', '\xa9', 'Island']]], 8, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'erk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Aas\xc3\xa1x', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aas\xc3\xa1x', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aas\xc3\xa1x', 'Language', 'codes', 'ISO', '639-3', 'aas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yazgulyam', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Yazgulyam', 'Writing', 'system', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yazgulyam', 'Language', 'codes', 'ISO', '639-3', 'yah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yazgulyam', 'Language', 'codes', 'Linguasphere', '58-ABD-f']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Cotoname', 'Spoken', 'in', 'lower', 'Rio', 'Grande']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'lower', 'Rio', 'Grande']]], 6, True)

['Cotoname', 'Language', 'codes', 'ISO', '639-3', 'xcn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chuukese', 'Spoken', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Chuukese', 'Region', 'Chuuk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chuukese', 'Official', 'language', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Federated', 'States', 'of', 'Micronesia']]], 8, True)

['Chuukese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Chuukese', 'Language', 'codes', 'ISO', '639-2', 'chk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Chuukese', 'Language', 'codes', 'ISO', '639-3', 'chk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Cayuga', 'Total', 'speakers', '100-200']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Cayuga', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Cayuga', 'Region', 'Six', 'Nations', 'of', 'the', 'Grand', 'River',
'First', 'Nation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Six',
'Nations', 'of', 'the', 'Grand', 'River', 'First', 'Nation']]], 10, True)
```

```
['Cayuga', 'Language', 'codes', 'ISO', '639-3', 'cay']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tsongla', 'Spoken', 'in', 'India', ':', '11', ',', '200', '(', 'Kameng',
',', 'W', 'Siang', ',', 'and', 'Dirang', 'Districts', ',', 'Arunachal',
'Pradesh', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',
':', 'NUM', ',', 'NUM', '\\(', 'Kameng', ',', 'W', 'Siang', ',', 'and',
'Dirang', 'Districts', ',', 'Arunachal', 'Pradesh', '\\)']], ['NUM', '11'],
['NUM', '200']], 21, True)
```

```
['Tsongla', 'Spoken', 'in', '000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tsongla', 'Writing', 'system', 'none', 'official', ';', 'Tibetan',
'used']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Tsongla', 'Language', 'codes', 'ISO', '639-3', 'tsj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tai', 'Spoken', 'in', 'Vietnam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tai', 'Spoken', 'in', 'Laos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

['Tai', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tai', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tai', 'Language', 'codes', 'ISO', '639-3', 'blt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Basaa', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Basaa', 'Region', 'Centre', 'and', 'Littoral', 'Provinces']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Centre', 'and', 'Littoral', 'Provinces']], 6, True)

['Basaa', 'Language', 'codes', 'ISO', '639-2', 'bas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Basaa', 'Language', 'codes', 'ISO', '639-3', 'bas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bassa', 'Language', 'codes', 'ISO', '639-3', 'bsq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Unami', 'Total', 'speakers', 'Extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Unami', 'Language', 'codes', 'ISO', '639-3', 'unm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Munsee', 'Spoken', 'in', 'Canada', ';', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada', ';', 'United', 'States']], 7, True)

['Munsee', 'Region', 'New', 'Jersey']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Munsee', 'Region', 'Pennsylvania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Munsee', 'Language', 'codes', 'ISO', '639-3', 'umu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tanapag', 'Region', 'Saipan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Tanapag', 'Language', 'codes', 'ISO', '639-3', 'tpv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Hupa', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Hupa', 'Region', 'California', '(', 'Hoopa', 'Valley', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'California', '\\\\(', 'Hoopa', 'Valley', '\\\\)']]], 7, True)
```

```
['Hupa', 'Language', 'codes', 'ISO', '639-2', 'hup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Hupa', 'Language', 'codes', 'ISO', '639-3', 'hup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Huastec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Huastec', 'Language', 'codes', 'ISO', '639-3', 'hus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Unserdeutsch', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```



['Unserdeutsch', 'Language', 'codes', 'ISO', '639-3', 'uIn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Plains', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']], 7, True)

['Plains', 'Region', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Plains', 'Region', 'Saskatchewan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Plains', 'Region', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Plains', 'Region', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Plains', 'Language', 'codes', 'ISO', '639-2', 'cre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Plains', 'Language', 'codes', 'ISO', '639-3', 'crk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Plains', 'Language', 'codes', 'Linguasphere', '62-ADA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Plains', 'Language', 'codes', 'Linguasphere', 'Linguistic',  
'subdivisions', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Linguistic', 'subdivisions', 'in', 'Canada']], 8, True)

['Plains', 'Language', 'codes', 'ISO', '639-1', 'cr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

```
['Wintu', 'Total', 'speakers', '5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Wintu', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Wintu', 'Region', 'Shasta', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Wintu', 'Region', 'Trinity', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Wintu', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Wintu', 'Language', 'codes', 'ISO', '639-3', 'wit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kharia', 'Language', 'codes', 'ISO', '639-3', 'khr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Miwok', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Miwok', 'Pronunciation', 'm', '\xc3', '\xad', 'w', '\xcb', '\x90',  
'\xc9', '\xa8', '\xcb', '\x90', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm',  
'\xc3', '\xad', 'w', '\xcb', '\x90', '\xc9', '\xa8', '\xcb', '\x90',  
'k']]], 13, True)
```

```
['Miwok', 'Spoken', 'in', 'Miwok', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Miwok', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Dinka', 'Total', 'speakers', '2-3', 'million']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '2-3']], 5, True)
```

```
['Dinka', 'Region', 'Sudan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Dinka', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Dinka', 'Pronunciation', '/', 't', '\xcc', '\xaa', 'u', '\xc9', '\x94',
'\xc5', '\x8b', '.', '\xc9', '\x9f', 'a', '\xcc', '\xa4', '\xc5', '\x8b',
'/']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',
't', '\xcc', '\xaa', 'u', '\xc9', '\x94', '\xc5', '\x8b', '.', '\xc9',
'\x9f', 'a', '\xcc', '\xa4', '\xc5', '\x8b', '/']], 20, True)
```

```
['Dinka', 'Language', 'codes', 'ISO', '639-2', 'din']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Dinka', 'Language', 'codes', 'ISO', '639-3', 'din', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'din', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Dinka', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Massa', 'Spoken', 'in', 'Chad', ',', 'Cameroon']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Massa', 'Language', 'codes', 'ISO', '639-3', 'mcn']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Skolt', 'Spoken', 'in', 'Finland', 'and', 'Russia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']], 6, True)
```

```
['Skolt', 'Writing', 'system', 'Latin', 'alphabet']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Skolt', 'Language', 'codes', 'ISO', '639-2', 'sms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Skolt', 'Language', 'codes', 'ISO', '639-3', 'sms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kangiryuarmiutun', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Kangiryuarmiutun', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kangiryuarmiutun', 'Region', 'North', 'America']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kangiryuarmiutun', 'Language', 'codes', 'ISO', '639-3', 'ikt-cop', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ikt-cop', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 9, True)

['Siglitun', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Siglitun', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Siglitun', 'Region', 'North', 'America']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Siglitun', 'Language', 'codes', 'ISO', '639-3', 'ikt-sig', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ikt-sig', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 9, True)

['Uummarmiutun', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Uummarmiutun', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Uummarmiutun', 'Region', 'North', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Uummarmiutun', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Nunatsiavummiutut', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Nunatsiavummiutut', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nunatsiavummiutut', 'Region', 'North', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nunatsiavummiutut', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)

['Nunatsiavummiutut', 'Language', 'codes', 'ISO', '639-3', 'ike-lab']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mono', 'Spoken', 'in', 'Solomon', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Mono', 'Total', 'speakers', '3400']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Mono', 'Dialects', 'Mono']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mono', 'Dialects', 'Alu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mono', 'Dialects', 'Fauro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mono', 'Language', 'codes', 'ISO', '639-3', 'mte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngunnawal', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kosraean', 'Spoken', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Kosraean', 'Region', 'Kosrae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kosraean', 'Official', 'language', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Federated', 'States', 'of', 'Micronesia']]], 8, True)

['Kosraean', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kosraean', 'Language', 'codes', 'ISO', '639-2', 'kos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kosraean', 'Language', 'codes', 'ISO', '639-3', 'kos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Potawatomi', 'Total', 'speakers', '1300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Potawatomi', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

['Potawatomi', 'Region', 'Michigan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Potawatomi', 'Region', 'Indiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Potawatomi', 'Region', 'Wisconsin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Potawatomi', 'Region', 'Kansas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Potawatomi', 'Region', 'and', 'southern', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Potawatomi', 'Language', 'codes', 'ISO', '639-3', 'pot']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Car', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Car', 'Region', 'Nicobar', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Car', 'Language', 'codes', 'ISO', '639-3', 'caq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Al-Sayyid', 'Region', 'Negev']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Al-Sayyid', 'Signed', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Al-Sayyid', 'Total', 'signers', '150', 'Deaf', ',', 'many', 'more', 'hearing']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'NUM', 'Deaf', ',', 'many', 'more', 'hearing']], ['NUM', '150']], 9, True)

['Al-Sayyid', 'Language', 'codes', 'ISO', '639-2', 'sgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Wutun', 'Total', 'speakers', '2000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wutun', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wutun', 'Region', 'Qinghai', 'province', ',', 'mainly', 'in', 'Tongren', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Qinghai', 'province', ',', 'mainly', 'in', 'Tongren', 'County']], 9, True)

['Wutun', 'Language', 'codes', 'ISO', '639-3', 'wuh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kim', 'Mun', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']], 7, True)

['Kim', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kim', 'Region', 'South', 'East', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kim', 'Regulated', 'by', 'No', 'official', 'regulation']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kim', 'Language', 'codes', 'ISO', '639-3', 'mji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kwanyama', 'Spoken', 'in', 'Namibia', 'and', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Kwanyama', 'Region', 'Ovamboland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kwanyama', 'Language', 'codes', 'ISO', '639-2', 'kua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kwanyama', 'Language', 'codes', 'ISO', '639-3', 'kua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kwanyama', 'Language', 'codes', 'Linguasphere', '05-PEA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Kwanyama', 'Language', 'codes', 'ISO', '639-1', 'kj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Kamba', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Kamba', 'Region', 'Machakos', ',', 'Kitui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kamba', 'Language', 'codes', 'ISO', '639-2', 'kam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kamba', 'Language', 'codes', 'ISO', '639-3', 'kam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Awabakal', 'Region', 'Lake', 'Macquarie', ',', 'Newcastle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Lake',  
'Macquarie', ',', 'Newcastle']]], 6, True)

['Awabakal', 'Language', 'codes', 'ISO', '639-3', 'awk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Day', 'Spoken', 'in', 'southern', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Day', 'Language', 'codes', 'ISO', '639-3', 'dai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bathari', 'Spoken', 'in', 'Yemen', ',', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Bathari', 'Language', 'codes', 'ISO', '639-3', 'bhm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Isthmus\xe2\x80\x93Mecayapan', 'Region', 'Mecayapan', 'and',  
'Tatahuicapan', ',', 'Veracruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mecayapan',  
'and', 'Tatahuicapan', ',', 'Veracruz']]], 7, True)

['Isthmus\xe2\x80\x93Mecayapan', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Isthmus\xe2\x80\x93Mecayapan', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Isthmus\xe2\x80\x93Mecayapan', 'Language', 'codes', 'ISO', '639-2',  
'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Isthmus\xe2\x80\x93Mecayapan', 'Language', 'codes', 'ISO', '639-3',  
'nhx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yidiny', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yidiny', 'Region', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Yidiny', 'Language', 'codes', 'ISO', '639-3', 'yii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ndonga', 'Spoken', 'in', 'Namibia', 'and', 'southern', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Namibia', 'and', 'southern', 'Angola']], 7, True)

['Ndonga', 'Region', 'Ovamboland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ndonga', 'Language', 'codes', 'ISO', '639-2', 'ndo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ndonga', 'Language', 'codes', 'ISO', '639-3', 'ndo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ndonga', 'Language', 'codes', 'Linguasphere', '05-PEA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Ndonga', 'Language', 'codes', 'ISO', '639-1', 'ng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Lan-nang', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lan-nang', 'Region', 'Metro', 'Manila']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Lan-nang', 'Region', 'Angeles']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lan-nang', 'Region', 'Cebu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lan-nang', 'Region', 'Bacolod']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lan-nang', 'Region', 'Vigan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lan-nang', 'Region', 'Naga']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lan-nang', 'Region', 'Iligan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lan-nang', 'Region', 'Davao', 'City']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lan-nang', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Lan-nang', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Lan-nang', 'Language', 'codes', 'ISO', '639-3', 'nan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lan-nang', 'Language', 'codes', 'ISO', '639-1', 'zh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Providence', 'Region', 'Providence', 'Island']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Providence', 'Signed', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Providence', 'Total', 'signers', '19', 'deaf', 'signers', ',', 'but', 'known', 'by', 'the', 'majority', 'of', 'the', '2', ',', '500', '\xe2', '\x80', '\x93', '3000', 'population']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'NUM', 'deaf', 'signers', ',', 'but', 'known', 'by', 'the', 'majority', 'of', 'the', 'NUM', ',', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'population']]], ['NUM', '19'], ['NUM', '2'], ['NUM', '500'], ['NUM', '3000']]], 22, True)

['Urubu', 'Region', 'Maranh', '\xc3', '\xa3', 'o']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maranh', '\xc3', '\xa3', 'o']]], 6, True)

['Urubu', 'Signed', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Urubu', 'Total', 'signers', '7', 'deaf', 'signers', 'and', 'about', '500', 'hearing', 'signers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'NUM', 'deaf', 'signers', 'and', 'about', 'NUM', 'hearing', 'signers']]], ['NUM', '7'], ['NUM', '500']]], 11, True)

['Mono', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Mono', 'Language', 'codes', 'ISO', '639-2', 'bad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']]], 6, True)

['Mono', 'Language', 'codes', 'ISO', '639-3', 'mnh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']]], 6, True)

['Mono', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Mono', 'Region', 'California']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mono', 'Language', 'codes', 'ISO', '639-3', 'mnr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nez', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nez', 'Region', 'Idaho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Nez', 'Language', 'codes', 'ISO', '639-3', 'nez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Aji\\xc3\\xab', 'Language', 'codes', 'ISO', '639-3', 'aji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Badimaya', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Badimaya', 'Language', 'codes', 'ISO', '639-3', 'bia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Comanche', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Comanche', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Comanche', 'Pronunciation', '\\xcb', '\\x88', 'n', '\\xc9', '\\xa8', 'm', '\\xc9', '\\xa8', '\\xcb', '\\x88', 'rek', '\\xca', '\\xb7', 'ap', '\\xc9', '\\xa8', '\\xcc', '\\xa5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb', '\\x88', 'n', '\\xc9', '\\xa8', 'm', '\\xc9', '\\xa8', '\\xcb', '\\x88', 'rek', '\\xca', '\\xb7', 'ap', '\\xc9', '\\xa8', '\\xcc', '\\xa5']]], 20, True)

```

['Comanche', 'Language', 'codes', 'ISO', '639-3', 'com']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Montagnais', 'Total', 'speakers', '11', ',', '815', '(', 'Montagnais-
Naskapi', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Montagnais-Naskapi', '\\)']], ['NUM', '11'],
['NUM', '815']], 9, True)

['Montagnais', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Montagnais', 'Region', 'Quebec', ',', 'Labrador']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Montagnais', 'Language', 'codes', 'ISO', '639-3', 'moe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Montagnais', 'Language', 'codes', 'Linguasphere', '62-ADA-bb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Mehri', 'Spoken', 'in', 'Yemen', ',', 'Oman']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Mehri', 'Language', 'codes', 'ISO', '639-3', 'gdq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ngan'gikurunggurr", 'Spoken', 'in', 'Northern', 'Territory']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

["Ngan'gikurunggurr", 'Language', 'codes', 'ISO', '639-3', 'nam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vai', 'Language', 'codes', 'ISO', '639-3', 'vai']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

```
['Shawnee', 'Total', 'speakers', '200']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Shawnee', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Shawnee', 'Region', 'Oklahoma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Shawnee', 'Language', 'codes', 'ISO', '639-3', 'sjw']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bamum', 'Spoken', 'in', 'Cameroon', ',', 'Nigeria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Bamum', 'Language', 'codes', 'ISO', '639-3', 'bax']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngadha', 'Spoken', 'in', 'Indonesia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ngadha', 'Region', 'Flores']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Ngadha', 'Language', 'codes', 'ISO', '639-3', 'nxg']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kriol', 'Spoken', 'in', 'Australia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kriol', 'Region', 'Roper', 'River']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kriol', 'Region', 'Katherine', 'areas']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```



'Type: ANY']]], 4, True)

['Kriol', 'Region', 'Ngukurr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kriol', 'Region', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kriol', 'Language', 'codes', 'ISO', '639-2', 'cpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kriol', 'Language', 'codes', 'ISO', '639-3', 'rop']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngalakan', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngalakan', 'Region', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngalakan', 'Language', 'codes', 'ISO', '639-3', 'nig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lotha', 'Spoken', 'in', 'Nagaland', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Lotha', 'Region', 'West-central', 'Nagaland', ',', 'Workha', 'district']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'West-central', 'Nagaland', ',', 'Workha', 'district']]], 7, True)

['Lotha', 'Language', 'codes', 'ISO', '639-3', 'njh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Berlin-Brandenburgish', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Berlin-Brandenburgish', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Berlin-Brandenburgish', 'Region', 'Berlin', ',', 'Brandenburg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Berlin-Brandenburgish', 'Language', 'codes', 'ISO', '639-3', '\xe2',  
'\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Ramaytush', '(', 'San', '&']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ramaytush', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Ramaytush', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ramaytush', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ramaytush', 'Language', 'codes', 'ISO', '639-3', 'cst']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adhola', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adhola', 'Region', 'Tororo', 'District', ',', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Tororo',  
'District', ',', 'Uganda']]], 6, True)

['Adhola', 'Language', 'codes', 'ISO', '639-3', 'adh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bobot', 'Spoken', 'in', 'Indonesia', '(', 'Maluku', 'Islands', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Indonesia', '\\(', 'Maluku', 'Islands', '\\)']], 8, True)

['Bobot', 'Region', 'Southeast', 'Seram']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bobot', 'Region', 'Werinama', 'District']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bobot', 'Region', 'and', 'Tunsai', 'village', 'in', 'the', 'Liana', 'area']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'Tunsai', 'village', 'in', 'the', 'Liana', 'area']], 9, True)

['Bobot', 'Region', 'central', 'Maluku']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bobot', 'Language', 'codes', 'ISO', '639-3', 'bty']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wajarri', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wajarri', 'Language', 'codes', 'ISO', '639-3', 'wbv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tundra', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tundra', 'Region', 'Yakutia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tundra', 'Language', 'codes', 'ISO', '639-3', 'ykg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaurna', 'Language', 'extinction', '1931']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Kaurna', 'Spoken', 'in', 'South', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Kaurna', 'Language', 'codes', 'ISO', '639-3', 'zku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Khams', 'Spoken', 'in', 'China', ',', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Khams', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Khams', 'Language', 'codes', 'ISO', '639-3', 'khg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Flemish', 'Region', 'Flanders', '(', 'northern', 'Belgium', ')', 'and',  
'Brussels-Capital']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Flanders', '\\  
(', 'northern', 'Belgium', '\\\')', 'and', 'Brussels-Capital']]], 9, True)

['Flemish', 'Signed', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Resian', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Resian', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Resian', 'Region', 'Resia', 'valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Resian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Resian', 'Language', 'codes', 'ISO', '639-3', 'slv-res']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Igala', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Igala', 'Language', 'codes', 'ISO', '639-3', 'igl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nyaw', 'Spoken', 'in', 'Thailand', ',', 'Laos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)
```

```
['Nyaw', 'Language', 'codes', 'ISO', '639-3', 'nyw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Phu', 'Spoken', 'in', 'Thailand', ',', 'Laos', 'and', 'Vietnam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Thailand', ',', 'Laos', 'and', 'Vietnam']], 8, True)
```

```
['Phu', 'Language', 'codes', 'ISO', '639-3', 'pht']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bergamasque', 'Spoken', 'in', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Bergamasque', 'Region', 'Lombardy', '(', 'Italy', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Lombardy', '\\(', 'Italy', '\\)']], 6, True)
```

```
['Bergamasque', 'Official', 'language', 'in', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)
```

```
['Bergamasque', 'Regulated', 'by', 'Ducato', 'di', 'Piazza', 'Pontida', '(', 'unofficial', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
```

'Ducato', 'di', 'Piazza', 'Pontida', '\\(', 'unofficial', '\\)']]]], 10, True)

['Nyangumarta', 'Spoken', 'in', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Nyangumarta', 'Region', 'Western', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Nyangumarta', 'Language', 'codes', 'ISO', '639-3', 'nna']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Natchez', 'Language', 'extinction', '1930s']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Natchez', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]]], 5, True)

['Natchez', 'Region', 'Louisiana', 'Ethnicity']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Natchez', 'Region', 'Natchez', 'people']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Natchez', 'Language', 'codes', 'ISO', '639-3', 'ncz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Teso', 'Spoken', 'in', 'Kenya', 'and', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]]], 6, True)

['Teso', 'Region', 'Northwest', 'Kenya', ',', 'west', 'of', 'Lake', 'Turkana']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Northwest', 'Kenya', ',', 'west', 'of', 'Lake', 'Type: ANY']]]], 9, True)

['Teso', 'Language', 'codes', 'ISO', '639-3', 'teo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xamtanga', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xamtanga', 'Region', 'North', 'Amhara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xamtanga', 'Language', 'codes', 'ISO', '639-3', 'xan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mongsen', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mongsen', 'Region', 'Nagaland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Mongsen', 'Language', 'codes', 'ISO', '639-3', 'njo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-3', ':', 'ant', '\xe2', '\x80', '\x93', 'Antikirinya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ant', '\xe2', '\x80', '\x93', 'Antikirinya']], ['NUM', '639-3']], 11, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'ktd', '\xe2', '\x80', '\x93', 'Kokata']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ktd', '\xe2', '\x80', '\x93', 'Kokata']], ['NUM', '639-3']], 10, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'kux', '\xe2', '\x80', '\x93', 'Kukatja']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kux', '\xe2', '\x80', '\x93', 'Kukatja']], ['NUM', '639-3']], 10, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'mpj', '\xe2', '\x80', '\x93', 'Martu', 'Wangka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mpj', '\xe2', '\x80', '\x93', 'Martu', 'Wangka']], ['NUM', '639-3']], 11, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'ntj', '\xe2', '\x80', '\x93', 'Ngaanyatjarra']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ntj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'piu', '\xe2', '\x80', '\x93', 'Pintupi-Luritja']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'piu', '\xe2', '\x80', '\x93', 'Pintupi-Luritja']], ['NUM', '639-3']], 10, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'pjt', '\xe2', '\x80', '\x93', 'Pitjantjatjara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pjt', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'kdd', '\xe2', '\x80', '\x93', 'Yankunytjatjara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kdd', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Atayal', 'Spoken', 'in', 'north', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Atayal', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Atayal', 'Language', 'codes', 'ISO', '639-3', 'tay']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turkish', 'Total', 'signers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malgana', 'Language', 'codes', 'ISO', '639-3', 'vml']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nhanda', 'Total', 'speakers', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nhanda', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nhanda', 'Language', 'codes', 'ISO', '639-3', 'nha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jam', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jam', 'Language', 'codes', 'ISO', '639-3', 'djm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['East', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['East', 'Region', 'Thuringia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['East', 'Region', 'Saxony']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['East', 'Region', 'Berlin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['East', 'Region', 'Brandenburg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['East', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Kambera', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']], 8, True)

['Kambera', 'Language', 'codes', 'ISO', '639-3', 'xbr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Palu'e", 'Region', "Palu'e"]:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

["Palu'e", 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Palu'e", 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

["Palu'e", 'Language', 'codes', 'ISO', '639-3', 'ple']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Soninke', 'Spoken', 'in', 'Mali']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Soninke', 'Spoken', 'in', 'Senegal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Soninke', 'Spoken', 'in', 'Gambia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Soninke', 'Spoken', 'in', 'Mauritania']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Soninke', 'Spoken', 'in', 'Guinea-Bissau']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Soninke', 'Spoken', 'in', 'Guinea', 'and', 'Ghana']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)
```

```
['Soninke', 'Region', 'West', 'Africa']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Soninke', 'Language', 'codes', 'ISO', '639-2', 'snk']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Soninke', 'Language', 'codes', 'ISO', '639-3', 'snk']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Even', 'Spoken', 'in', 'Russia']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Even', 'Region', 'Russian', 'Far', 'East']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Even', 'Language', 'codes', 'ISO', '639-3', 'eve']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dhatki', 'Spoken', 'in', 'Pakistan', 'and', 'India', '(', 'Marwar', 'region', 'of', 'Rajasthan', 'and', 'Thar', 'region', 'of', 'Sindh', ')', '(', 'sociolect', ')']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Pakistan', 'and', 'India', '\\\\(', 'Marwar', 'region', 'of', 'Rajasthan', 'and', 'Thar', 'region', 'of', 'Sindh', '\\\\)', '\\\\(', 'sociolect', '\\\\)']]], 20, True)
```

```
['Dhatki', 'Language', 'codes', 'ISO', '639-2', 'raj']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```

['Dhatki', 'Language', 'codes', 'ISO', '639-3', 'mki']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['T\xc3\xa2y', 'Spoken', 'in', 'Vietnam', ',', 'United', 'States', '(',
'California', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Vietnam', ',', 'United', 'States', '\\(', 'California', '\\)']], 10,
True)

['T\xc3\xa2y', 'Language', 'codes', 'ISO', '639-3', 'tas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bishnupriya', 'Spoken', 'in', 'Northeast', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Bishnupriya', 'Spoken', 'in', 'Bangladesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Bishnupriya', 'Language', 'codes', 'ISO', '639-3', 'bpy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamayur\xc3\xa1', 'Total', 'speakers', '279']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kamayur\xc3\xa1', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kamayur\xc3\xa1', 'Region', 'Upper', 'Xingu', 'region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['Kamayur\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'kay']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Toto', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']], 6, True)

```

```
['Toto', 'Spoken', 'in', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2',  
'\x80', '\x93']]], 6, True)
```

```
['Toto', 'Region', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Toto', 'Language', 'codes', 'ISO', '639-3', 'txo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Puma', 'Pronunciation', '\xcb', '\x88', 'puma', '\xcb', '\x90']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'puma', '\xcb', '\x90']]], 7, True)
```

```
['Puma', 'Spoken', 'in', 'Khotang', 'district', 'and', 'Udayapur',  
'district', ',', 'Nepal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Khotang', 'district', 'and', 'Udayapur', 'district', ',', 'Type: ANY']]],  
10, True)
```

```
['Puma', 'Language', 'codes', 'ISO', '639-3', 'pum']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Akatek', 'Spoken', 'in', 'Guatemala']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Akatek', 'Spoken', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Akatek', 'Region', 'Huehuetenango']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Akatek', 'Region', 'Chiapas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Akatek', 'Language', 'codes', 'ISO', '639-3', 'knj']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Tonga', 'Spoken', 'in', 'Zambia', ',', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Tonga', 'Official', 'language', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Tonga', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tonga', 'Language', 'codes', 'ISO', '639-3', 'toi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hungarian', 'Region', 'Throughout', 'Hungary', ',', 'with', 'possibly',  
'a', 'few', 'signers', 'in', 'western', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Throughout',  
'Hungary', ',', 'with', 'possibly', 'a', 'few', 'signers', 'in', 'western',  
'Romania']]], 13, True)

['Hungarian', 'Signed', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hungarian', 'Total', 'signers', '30', '\xe2', '\x80', '\x93', '60', ',',  
'000', 'deaf', 'and', '300', ',', '000', 'hearing', 'impaired']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'NUM', '\xe2', '\x80', '\x93', 'NUM', ',', 'NUM', 'deaf', 'and', 'NUM',  
, ',', 'NUM', 'hearing', 'impaired']], ['NUM', '30'], ['NUM', '60'], ['NUM',  
'000'], ['NUM', '300'], ['NUM', '000']], 17, True)

['Torlakian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Torlakian', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Torlakian', 'Spoken', 'in', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Torlakian', 'Spoken', 'in', 'Kosovo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Torlakian', 'Spoken', 'in', 'Romania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Torlakian', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Torlakian', 'Language', 'codes', 'ISO', '639-3', 'srp-tor', '\xe2',
'\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'srp-tor', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,
True)
```

```
['Au', 'Language', 'codes', 'ISO', '639-3', 'avt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kriol', 'Language', 'codes', 'ISO', '639-3', 'bjj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kriol', 'Language', 'codes', 'Linguasphere', '52-ABB-ad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Danzhouhua', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Danzhouhua', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Gonja', 'Spoken', 'in', 'Ghana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Gonja', 'Region', 'Northern', 'Ghana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```

['Gonja', 'Language', 'codes', 'ISO', '639-3', 'gjn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tillamook', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Tillamook', 'Region', 'Northwestern', 'Oregon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tillamook', 'Language', 'codes', 'ISO', '639-2', 'sal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tillamook', 'Language', 'codes', 'ISO', '639-3', 'til']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngiyambaa', 'Spoken', 'in', 'New', 'South', 'Wales']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', 'Wales']]], 6, True)

['Ngiyambaa', 'Total', 'speakers', '12', '(', '?', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', '\\(', '?', '\\)']], ['NUM', '12']], 7, True)

['Ngiyambaa', 'Language', 'codes', 'ISO', '639-3', 'wyb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Harari', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Harari', 'Region', 'Harari']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Harari', 'Language', 'codes', 'ISO', '639-3', 'har']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Silt'e", 'Spoken', 'in', 'Southern', 'Nations']:

```



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Silt'e", 'Spoken', 'in', 'Nationalities']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
["Silt'e", 'Spoken', 'in', 'and', "People's"]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Silt'e", 'Region', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
["Silt'e", 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Silt'e", 'Language', 'codes', 'ISO', '639-3', 'wle', '\xe2', '\x80', '\x93', 'Wolane']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wle', '\xe2', '\x80', '\x93', 'Wolane']], ['NUM', '639-3']], 10, True)
```

```
['Bohtan', 'Spoken', 'in', 'Georgia', ',', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Bohtan', 'Region', 'Mainly', 'in', 'Gardabani', 'village', ',', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mainly', 'in', 'Gardabani', 'village', ',', 'Georgia']]], 8, True)
```

```
['Bohtan', 'Language', 'codes', 'ISO', '639-3', 'bhn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Maria', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Maria', 'Language', 'codes', 'ISO', '639-3', 'mrr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kwanga', 'Language', 'codes', 'ISO', '639-3', 'kwj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kam', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kam', 'Region', 'Guizhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kam', 'Region', 'Hunan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kam', 'Region', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kam', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Vlax', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']], 6, True)
```

```
['Vlax', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Vlax', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Vlax', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Vlax', 'Spoken', 'in', 'Hungary', 'and', '21', 'other', 'states',  
'worldwide']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Hungary', 'and', 'NUM', 'other', 'states', 'worldwide']], ['NUM', '21']],  
9, True)
```

['Vlax', 'Official', 'language', 'in', 'Officially-recognised', 'minority',  
'language', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Officially-recognised', 'minority', 'language', 'in', 'Hungary']]],  
9, True)

['Vlax', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Vlax', 'Official', 'language', 'in', 'Poland', 'and', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Poland', 'and', 'Serbia']]], 7, True)

['Vlax', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Vlax', 'Language', 'codes', 'ISO', '639-2', 'rom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Vlax', 'Language', 'codes', 'ISO', '639-3', 'rmy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tiwi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tiwi', 'Language', 'codes', 'ISO', '639-3', 'tiw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kwak'wala', 'Language', 'codes', 'ISO', '639-3', 'kwk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cal\xc3\xb3', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cal\xc3\xb3', 'Spoken', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ca1\xc3\xb3', 'Language', 'codes', 'ISO', '639-3', 'rmq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tetet\xc3\xa9', 'Spoken', 'in', 'Ecuador', ',', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Tetet\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'teb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nar', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nar', 'Region', 'Manang', 'district']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nar', 'Language', 'codes', 'ISO', '639-3', 'npa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amuzgo', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Amuzgo', 'Region', 'Guerrero', ',', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Amuzgo', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amuzgo', 'Language', 'codes', 'ISO', '639-3', 'azm', '\xe2', '\x80',  
'\x93', 'Ipalapa', 'Amuzgo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'azm', '\xe2', '\x80', '\x93', 'Ipalapa', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Mikasuki', 'Total', 'speakers', '500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Mikasuki', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Mikasuki', 'Region', 'Southern', 'Florida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mikasuki', 'Language', 'codes', 'ISO', '639-3', 'mik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Loxian', 'Date', 'founded', '2005']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Loxian', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Loxian', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Loxian', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Loxian', 'Category', '(', 'purpose', ')', 'Loxian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)
```

```
['Loxian', 'Created', 'by', 'Roma', 'Ryan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Loxian', 'Category', '(', 'sources', ')', 'based', 'on', 'snippets',  
'of', 'English']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'based', 'on', 'snippets', 'of', 'Type: ANY']]], 10,  
True)
```

```
['Loxian', 'Category', '(', 'sources', ')', 'Irish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Type: ANY']]], 6, True)
```

```

['Loxian', 'Category', '(', 'sources', ')', 'Old', 'English']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)', 'Type: ANY', 'Type: ANY']]], 7, True)

['Loxian', 'Category', '(', 'sources', ')', 'Hindi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)', 'Type: ANY']]], 6, True)

['Loxian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Jersey', 'Spoken', 'in', 'Jersey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Jersey', 'Official', 'language', 'in', 'Jersey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Jersey', 'Regulated', 'by']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Jersey', 'Language', 'codes', 'ISO', '639-3', 'fre']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sonsorolese', 'Total', 'speakers', '600']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sonsorolese', 'Spoken', 'in', 'Palau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sonsorolese', 'Region', 'Palau', ':', 'originally', 'Sonsorol', 'state',
'(', 'all', 'three', 'inhabited', 'islands', 'Sonsorol', ',', 'Pulo',
'Anna', 'and', 'Merir', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Palau', ':',
'originally', 'Sonsorol', 'state', '\\(', 'all', 'three', 'inhabited',
'islands', 'Sonsorol', ',', 'Pulo', 'Anna', 'and', 'Merir', '\\)']]], 19,
True)

['Sonsorolese', 'Language', 'codes', 'ISO', '639-3', 'sov']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Argentina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['South', 'Spoken', 'in', 'Bolivia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['South', 'Spoken', 'in', 'Chile', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['South', 'Language', 'codes', 'ISO', '639-2', 'que']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'quh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mayo', 'Spoken', 'in', 'Sonora', 'and', 'Sinaloa', ',', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Sonora', 'and', 'Sinaloa', ',', 'Mexico']], 8, True)

['Mayo', 'Language', 'codes', 'ISO', '639-3', 'mfy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maxakal\xc3\xad', 'Total', 'speakers', '728']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maxakal\xc3\xad', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maxakal\xc3\xad', 'Region', 'Minas', 'Gerais']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maxakal\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'mbl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Assiniboine', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']]], 7, True)
```

```
['Assiniboine', 'Language', 'codes', 'ISO', '639-3', 'asb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sasak', 'Spoken', 'in', 'Lombok', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Sasak', 'Writing', 'system', 'Balinese', 'script', ',', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'script', ',', 'Type: ANY']]], 7, True)
```

```
['Sasak', 'Language', 'codes', 'ISO', '639-2', 'sas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Sasak', 'Language', 'codes', 'ISO', '639-3', 'sas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Honduras', 'Signed', 'in', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Honduras', 'Total', 'signers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Ida\xe2\x80\x99an', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ida\xe2\x80\x99an', 'Region', 'Lahad', 'Datu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ida\xe2\x80\x99an', 'Region', 'Kinabatangan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Ida\xe2\x80\x99an', 'Region', 'and', 'Sandakan', 'districts', 'of',  
'Sabah']:
```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'Sandakan', 'districts', 'of', 'Sabah']]], 7, True)

['Ida\xe2\x80\x99an', 'Language', 'codes', 'ISO', '639-3', 'dbj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mlabri', 'Spoken', 'in', 'Thailand', ',', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Mlabri', 'Language', 'codes', 'ISO', '639-3', 'mra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tuamotuan', 'Spoken', 'in', 'French', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tuamotuan', 'Region', 'the', 'Tuamotus', ',', 'Tahiti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'the', 'Tuamotus', ',', 'Tahiti']]], 6, True)

['Tuamotuan', 'Language', 'codes', 'ISO', '639-3', 'pmt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Liburnian', 'Spoken', 'in', 'Western', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Liburnian', 'Region', 'NorthWestern', 'Balkans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Liburnian', 'Language', 'codes', 'ISO', '639-3', 'xli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lak', 'Spoken', 'in', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lak', 'Writing', 'system', 'Cyrillic', '(', 'Lak', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Lak', 'Official', 'language', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Lak', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Lak', 'Language', 'codes', 'ISO', '639-3', 'lbe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xinca', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xinca', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xinca', 'Language', 'codes', 'ISO', '639-3', 'xin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mbundu', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mbundu', 'Region', 'Malanje', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mbundu', 'Language', 'codes', 'ISO', '639-2', 'kmb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mbundu', 'Language', 'codes', 'ISO', '639-3', 'kmb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khakas', 'Dialects', 'Fuy', '\xc3', '\xbc', 'G', '\xc3', '\xaf', 'rg',  
'\xc3', '\xaf', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Fuy',  
'\xc3', '\xbc', 'G', '\xc3', '\xaf', 'rg', '\xc3', '\xaf', 's']]], 12,  
True)

['Khakas', 'Region', 'Khakassia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Khakas', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khakas', 'Language', 'codes', 'ISO', '639-3', 'kjh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Comox', 'Total', 'speakers', '400']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Comox', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Comox', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Comox', 'Language', 'codes', 'ISO', '639-3', 'coo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anus', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Anus', 'Region', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Anus', 'Language', 'codes', 'ISO', '639-3', 'auq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dolgan', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dolgan', 'Region', 'Krasnoyarsk', 'Krai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dolgan', 'Language', 'codes', 'ISO', '639-3', 'dlg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Eteocypriot', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'Cyprus']]], 7, True)

['Eteocypriot', 'Region', 'Eastern', 'Mediterranean', 'Sea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Eteocypriot', 'Language', 'codes', 'ISO', '639-3', 'ecy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Irula', 'Spoken', 'in', 'Tamil', 'Nadu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Irula', 'Language', 'codes', 'ISO', '639-3', 'iru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tadaksahak', 'Spoken', 'in', 'Mali', ',', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Tadaksahak', 'Language', 'codes', 'ISO', '639-3', 'dsq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nuu-chah-nulth', 'Region', 'from', 'Barkley', 'Sound', 'to', 'Quatsino', 'Sound']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'from', 'Barkley', 'Sound', 'to', 'Quatsino', 'Sound']]], 8, True)

['Nuu-chah-nulth', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nuu-chah-nulth', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Nuu-chah-nulth', 'Pronunciation', 'nu', '\xcb', '\x90', 't', '\xcd', '\xa1', '\xca', '\x83', 'a', '\xcb', '\x90', 'n', '\xcb', '\x80', 'u', '\xc9', '\xac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'nu', '\xcb', '\x90', 't', '\xcd', '\xa1', '\xca', '\x83', 'a', '\xcb', '\x90', 'n', '\xcb', '\x80', 'u', '\xc9', '\xac']]], 19, True)

['Nuu-chah-nulth', 'Language', 'codes', 'ISO', '639-3', 'noo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Babungo', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Babungo', 'Language', 'codes', 'ISO', '639-3', 'bav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Q'eqchi'", 'Total', 'speakers', '400', ',', '000', 'speakers', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'speakers', 'in', 'Type: ANY']], ['NUM', '400'], ['NUM', '000']], 9, True)

["Q'eqchi'", 'Total', 'speakers', '11', ',', '200', 'speakers', 'in', 'Belize']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'speakers', 'in', 'Type: ANY']], ['NUM', '11'], ['NUM', '200']], 9, True)

["Q'eqchi'", 'Region', 'Alta', 'Verapaz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Q'eqchi'", 'Region', 'Pet', '\xc3', '\xa9', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Pet', '\xc3', '\xa9', 'n']]], 6, True)

["Q'eqchi'", 'Region', 'Izabal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

["Q'eqchi'", 'Region', 'Baja', 'Verapaz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Q'eqchi'", 'Spoken', 'in', 'Guatemala']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Q'eqchi'", 'Spoken', 'in', 'Belize']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Q'eqchi'", 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Q'eqchi'", 'Spoken', 'in', 'El', 'Salvador']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

["Q'eqchi'", 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Q'eqchi'", 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

["Q'eqchi'", 'Official', 'language', 'in', 'Considered', 'a', 'national', 'language', 'of', 'Guatemala', 'and', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Considered', 'a', 'national', 'language', 'of', 'Guatemala', 'and', 'Mexico']]], 12, True)

["Q'eqchi'", 'Regulated', 'by', 'Academ', '\xc3', '\xad', 'a', 'de', 'las', 'Lenguas', 'Mayas', 'de', 'Guatemala']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Academ', '\xc3', '\xad', 'a', 'de', 'las', 'Lenguas', 'Mayas', 'de', 'Guatemala']]], 13, True)

["Q'eqchi'", 'Language', 'codes', 'ISO', '639-3', 'kek']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Groma', 'Spoken', 'in', 'Tibet', ',', 'Bhutan', 'and', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Tibet', ',', 'Bhutan', 'and', 'India']]], 8, True)

['Groma', 'Region', 'Chambi', 'Valley', 'region', 'between', 'Sikkim', 'and', 'Bhutan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Chambi', 'Valley', 'region', 'between', 'Sikkim', 'and', 'Bhutan']]], 9, True)

['Groma', 'Language', 'codes', 'ISO', '639-3', 'gro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gurung', 'Total', 'speakers', '227', ',', '918', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '227'], ['NUM', '918']], 8, True)

['Gurung', 'Region', 'South', 'Asia', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Gurung', 'Region', 'Gurung', 'people']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gurung', 'Spoken', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gurung', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gurung', 'Spoken', 'in', 'Bhutan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gurung', 'Writing', 'system', 'Tibetan', 'script', ',', 'Devanagari', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Gurung', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lepcha', 'Writing', 'system', 'Lepcha', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Type: ANY', 'Type: ANY']]], 5, True)

['Lepcha', 'Language', 'codes', 'ISO', '639-3', 'lep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Limbu', 'Spoken', 'in', 'Limbuwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Limbu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Limbu', 'Writing', 'system', 'Limbu', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Limbu', 'Official', 'language', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Limbu', 'Official', 'language', 'in', 'Sikkim', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'India', '\\)']]], 8, True)

['Limbu', 'Official', 'language', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Limbu', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Limbu', 'Language', 'codes', 'ISO', '639-3', 'lif']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Magar', 'Spoken', 'in', 'Nepal', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Magar', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)



['Magar', 'Official', 'language', 'in', 'No', 'official', 'status']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'No', 'official', 'status']]], 7, True)

['Magar', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Magar', 'Language', 'codes', 'ISO', '639-3', 'mgp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Dhanwar', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dhanwar', 'Language', 'codes', 'ISO', '639-3', 'dhw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sherpa', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sherpa', 'Spoken', 'in', 'Tibet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sherpa', 'Spoken', 'in', 'Kashmir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sherpa', 'Language', 'codes', 'ISO', '639-3', 'xsr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sunuwar', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sunuwar', 'Language', 'codes', 'ISO', '639-3', 'suz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tamang', 'Total', 'speakers', '1', 'million', 'approx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',

'NUM', 'million', 'approx']], ['NUM', '1']], 6, True)

['Tamang', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tamang', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tamang', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tamang', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tamang', 'Writing', 'system', 'Tibetan', 'script', ',', 'Devnagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'script', ',', 'Type: ANY']], 7, True)

['Tamang', 'Language', 'codes', 'ISO', '639-3', 'taj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thulung', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thulung', 'Language', 'codes', 'ISO', '639-3', 'tdh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yakkha', 'Spoken', 'in', 'Nepal', ',', 'Sikkim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Yakkha', 'Language', 'codes', 'ISO', '639-3', 'ybh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Signed', 'in', 'Norway', ',', 'Madagascar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'Norway', ',', 'Madagascar']], 6, True)

```
['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nsl', '\xe2',  
'\x80', '\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'nsl', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Rombo', 'Total', 'speakers', '1', ',', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'million']], ['NUM', '1'], ['NUM', '5']], 7, True)
```

```
['Rombo', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Rombo', 'Region', 'Kilimanjaro', ',', 'Chaga', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kilimanjaro',  
',', 'Chaga', 'area']], 6, True)
```

```
['Rombo', 'Language', 'codes', 'ISO', '639-3', 'rof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Iu', 'Region', 'South', 'East', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Iu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Iu', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Iu', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Iu', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

['Iu', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iu', 'Spoken', 'in', 'and', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Iu', 'Pronunciation', 'ju', 'mj', '\xc9', '\x9b', 'nneed', 'tone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ju', 'mj', '\xc9', '\x9b', 'nneed', 'tone']], 8, True)

['Iu', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Iu', 'Language', 'codes', 'ISO', '639-3', 'ium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Naukan', 'Spoken', 'in', 'Russian', 'Federation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Naukan', 'Region', 'Bering', 'Strait', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Naukan', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Naukan', 'Language', 'codes', 'ISO', '639-3', 'ynk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pecheneg', 'Spoken', 'in', 'Crimea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pecheneg', 'Region', 'Krym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Pecheneg', 'Language', 'codes', 'ISO', '639-3', 'xpc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['M\xc6\xb0\xe1\xbb\x9dng', 'Spoken', 'in', 'Vietnam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['M\xc6\xb0\xe1\xbb\x9dng', 'Region', 'Southeast', 'Asia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['M\xc6\xb0\xe1\xbb\x9dng', 'Language', 'codes', 'ISO', '639-3', 'mtq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tongva', 'Spoken', 'in', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tongva', 'Region', 'Southern', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tongva', 'Language', 'codes', 'ISO', '639-3', 'xgf']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Serrano', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Serrano', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Serrano', 'Region', 'Southern', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Serrano', 'Language', 'codes', 'ISO', '639-3', 'ser']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balkan', 'Spoken', 'in', 'Turkey', ',']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',']], 5, True)

```
['Balkan', 'Spoken', 'in', 'Greece', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',']], 5, True)
```

```
['Balkan', 'Spoken', 'in', 'Macedonia', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',']], 5, True)
```

```
['Balkan', 'Language', 'codes', 'ISO', '639-3', 'bgx']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Anal', 'Spoken', 'in', 'India', 'and', 'Myanmar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']], 6, True)
```

```
['Anal', 'Region', 'Southeast', 'Manipur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Anal', 'Language', 'codes', 'ISO', '639-3', 'anm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Libyan', 'Spoken', 'in', 'Libya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Libyan', 'Spoken', 'in', 'Egypt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Libyan', 'Spoken', 'in', 'Niger']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Libyan', 'Spoken', 'in', 'Algeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Libyan', 'Spoken', 'in', 'Tunisia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Libyan', 'Writing', 'system', 'Arabic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['Libyan', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Libyan', 'Language', 'codes', 'ISO', '639-3', 'ayl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gamlaraay', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Gamlaraay', 'Language', 'codes', 'ISO', '639-3', 'kld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kakauhua', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kakauhua', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kakauhua', 'Language', 'codes', 'ISO', '639-3', 'kbf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wiradjuri', 'Spoken', 'in', 'New', 'South', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', 'Wales']]], 6, True)

['Wiradjuri', 'Language', 'codes', 'ISO', '639-3', 'wrh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Old', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'owl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hill', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Hill', 'Language', 'codes', 'ISO', '639-2', 'chm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hill', 'Language', 'codes', 'ISO', '639-3', 'mrj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gao-Yang', 'Spoken', 'in', 'Southern', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Gao-Yang', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']],  
3, True)

['Gao-Yang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Gao-Yang', 'Language', 'codes', 'ISO', '639-3', 'yue-gao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kele', 'Total', 'speakers', '600']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kele', 'Region', 'Manus', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kele', 'Language', 'codes', 'ISO', '639-3', 'sbc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Munichi', 'Language', 'extinction', '1990s']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Munichi', 'Spoken', 'in', 'Loreto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Munichi', 'Language', 'codes', 'ISO', '639-3', 'myr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Garifuna', 'Spoken', 'in', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Garifuna', 'Spoken', 'in', 'Belize']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Garifuna', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Garifuna', 'Language', 'codes', 'ISO', '639-3', 'cab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tweants', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tweants', 'Region', 'Northeast', ',', 'Overijssels', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Northeast',  
,',', 'Overijssels', 'province']]], 6, True)

['Tweants', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tweants', 'Language', 'codes', 'ISO', '639-3', 'twd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sikkimese', 'Spoken', 'in', 'Sikkim', ',', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', ',', 'Type: ANY']]], 6, True)

['Sikkimese', 'Writing', 'system', 'Tibetan', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sikkimese', 'Language', 'codes', 'ISO', '639-3', 'sip']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Eastern', 'Region', 'Northern', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'peb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kashaya', 'Total', 'speakers', '45']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kashaya', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Kashaya', 'Region', 'Sonoma', 'County', ',', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'County', ',', 'California']]], 6, True)

['Kashaya', 'Language', 'codes', 'ISO', '639-3', 'kju']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mexican', 'Region', 'Cities']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mexican', 'Signed', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Yanam', 'Spoken', 'in', 'Brazil', ',', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Yanam', 'Total', 'speakers', '560']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Yanam', 'Language', 'codes', 'ISO', '639-3', 'shb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yanam', 'Language', 'codes', 'ISO', '639-3', 'location', 'of', 'Roraima',  
,, 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'location', 'of', 'Roraima', ',', 'Brazil']], ['NUM',  
'639-3']], 10, True)
```

```
['Shiriana', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Shiriana', 'Total', 'speakers', '0', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '?']], ['NUM', '0']], 5, True)
```

```
['Shiriana', 'Language', 'codes', 'ISO', '639-3', 'xir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Berbice', 'Spoken', 'in', 'Guyana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Berbice', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Berbice', 'Language', 'codes', 'ISO', '639-3', 'brc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kenyan', 'Signed', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['I\xe2\x80\x99saka', 'Total', 'speakers', '421']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['I\xe2\x80\x99saka', 'Language', 'codes', 'ISO', '639-3', 'ksi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mazandarani', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mazandarani', 'Spoken', 'in', 'province', 'of', 'Mazandaran', 'and',  
'parts', 'of', 'the', 'provinces', 'of', 'Tehran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'province', 'of', 'Mazandaran', 'and', 'parts', 'of', 'the', 'provinces',  
'of', 'Tehran']]], 13, True)

['Mazandarani', 'Spoken', 'in', 'Golestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mazandarani', 'Spoken', 'in', 'Semnan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mazandarani', 'Spoken', 'in', 'Khorasan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mazandarani', 'Writing', 'system', 'Perso-Arabic', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Mazandarani', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Mazandarani', 'Regulated', 'by', 'Linguistic', 'Faculty', 'of', 'Babol',  
'University']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Linguistic', 'Faculty', 'of', 'Babol', 'University']]], 8, True)

['Mazandarani', 'Language', 'codes', 'ISO', '639-3', 'mzn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wutung', 'Total', 'speakers', '903']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wutung', 'Language', 'codes', 'ISO', '639-3', 'wut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['R\\xc3\\xa9union', 'Spoken', 'in', 'R', '\\xc3', '\\xa9', 'union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'R', '\\xc3', '\\xa9', 'union']], 7, True)

['R\\xc3\\xa9union', 'Language', 'codes', 'ISO', '639-3', 'rcf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['R\\xc3\\xa9union', 'Language', 'codes', 'Linguasphere', '51-AAC-cf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Zay', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zay', 'Region', 'Lake', 'Zway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zay', 'Language', 'codes', 'ISO', '639-3', 'zwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shipibo-Conibo', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shipibo-Conibo', 'Official', 'language', 'in', 'Ucayali', 'Region', '(', 'Peru', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Ucayali', 'Region', '\\\\(', 'Peru', '\\\\)']], 9, True)

['Shipibo-Conibo', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Shipibo-Conibo', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Shipibo-Conibo', 'Language', 'codes', 'ISO', '639-3', 'shp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jarawara', 'Spoken', 'in', 'Amazonas', ',', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Jarawara', 'Total', 'speakers', '155']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jarawara', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Jarawara', 'Language', 'codes', 'ISO', '639-3', 'jap']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huaorani', 'Spoken', 'in', 'Ecuador', ',', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Huaorani', 'Official', 'language', 'in', 'Ecuador', ':', 'indigenous', 'languages', 'official', 'in', 'own', 'territories']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Ecuador', ':', 'indigenous', 'languages', 'official', 'in', 'own', 'territories']], 12, True)

['Huaorani', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Huaorani', 'Language', 'codes', 'ISO', '639-3', 'auc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Godoberi', 'Spoken', 'in', 'Southwestern', 'Dagestan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Godoberi', 'Language', 'codes', 'ISO', '639-3', 'gdo']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Enga', 'Regulated', 'by', 'No', 'official', 'regulation']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY' ] ] ], 6, True)

['Enga', 'Language', 'codes', 'ISO', '639-3', 'enq']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Tauya', 'Total', 'speakers', '350']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Tauya', 'Region', 'Madang', 'Province']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Tauya', 'Language', 'codes', 'ISO', '639-3', 'tya']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Libras', 'Region', 'Cities']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 3, True)

['Libras', 'Signed', 'in', 'Brazil']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Libras', 'Total', 'signers', 'unknown']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Pumi', 'Region', 'Sichuan', ',', 'Yunnan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 5, True)

['Pumi', 'Writing', 'system', 'none']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Pumi', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northern', 'Spoken', 'in', 'Sichuan', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'cng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Spoken', 'in', 'Sichuan', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'qxs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mararit', 'Spoken', 'in', 'Sudan', 'and', 'Tchad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Mararit', 'Language', 'codes', 'ISO', '639-2', 'ssa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mararit', 'Language', 'codes', 'ISO', '639-3', 'mgb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shatt', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shatt', 'Region', 'Northern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shatt', 'Language', 'codes', 'ISO', '639-3', 'shj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shatt', 'Language', 'codes', 'Linguasphere', '05-PEA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)



['Northeastern', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Northeastern', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Northeastern', 'Language', 'codes', 'ISO', '639-3', 'nmu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tucano', 'Spoken', 'in', 'Brazil', ',', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Tucano', 'Language', 'codes', 'ISO', '639-3', 'tuo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tariana', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tariana', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Tariana', 'Language', 'codes', 'ISO', '639-3', 'tae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gilaki', 'Total', 'speakers', '2-4', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '2-4']], 5, True)

['Gilaki', 'Spoken', 'in', 'Iran', ',', 'province', 'of', 'Gilan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iran', ',', 'province', 'of', 'Gilan']]], 8, True)

['Gilaki', 'Region', 'southwest', 'coast', 'of', 'the', 'Caspian', 'Sea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'southwest', 'coast', 'of', 'the', 'Caspian', 'Sea']]], 8, True)

['Gilaki', 'Language', 'codes', 'ISO', '639-3', 'glk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gilaki', 'Language', 'codes', 'Linguasphere', '58-AAC-eb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Remo', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Remo', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Remo', 'Language', 'codes', 'ISO', '639-3', 'rem']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northern', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Northern', 'Region', 'Northern', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'pej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pintupi', 'Total', 'speakers', '300-400']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pintupi', 'Language', 'codes', 'ISO', '639-3', 'piu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fox', 'Total', 'speakers', '200-1000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fox', 'Spoken', 'in', 'United', 'States', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

```
['Fox', 'Region', 'Central', 'Oklahoma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Fox', 'Region', 'Northeastern', 'Kansas']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Fox', 'Region', 'Iowa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Fox', 'Region', 'and', 'Coahuila', 'Ethnicity']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Fox', 'Region', 'Fox']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Fox', 'Region', 'Sauk']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Fox', 'Region', 'and', 'Kickapoo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Fox', 'Language', 'codes', 'ISO', '639-3', 'either']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Fox', 'Language', 'codes', 'ISO', '639-3', 'kic', '\xe2', '\x80', '\x93',
'Kickapoo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'kic', '\xe2', '\x80', '\x93', 'Kickapoo']], ['NUM',
'639-3']], 10, True)
```

```
['Tooro', 'Standard', 'forms', 'Kitara']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tooro', 'Spoken', 'in', 'Uganda']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```

['Tooro', 'Language', 'codes', 'ISO', '639-3', 'ttj']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Drehu', 'Language', 'codes', 'ISO', '639-3', 'dhv']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['American', 'Region', 'Anglophone', 'North', 'America']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['American', 'Signed', 'in', 'United', 'States', ',', 'Canada']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'United',
'States', ',', 'Canada']], 7, True)

['American', 'Total', 'signers', 'unknown', ',', 'perhaps', '500', ',',
'000', 'or', 'so', 'primary', 'users']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',
'unknown', ',', 'perhaps', 'NUM', ',', 'NUM', 'or', 'so', 'primary',
'users']], ['NUM', '500'], ['NUM', '000']], 13, True)

['American', 'Dialects', 'Black', 'ASL']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['American', 'Dialects', 'Bolivian', 'Sign']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Armenian', 'Spoken', 'in', 'Armenia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Armenian', 'Spoken', 'in', 'Nagorno-Karabakh', '(', 'not', 'recognized',
'internationally', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nagorno-
Karabakh', '\\(', 'not', 'recognized', 'internationally', '\\)']], 9,
True)

['Armenian', 'Spoken', 'in', 'Russia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Armenian', 'Spoken', 'in', 'United', 'States']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Armenian', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Armenian', 'Writing', 'system', 'Armenian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Armenian', 'Pronunciation', 'h', '\xc9', '\x91', 'j', '\xc9', '\x9b', '\xcb', '\x88', '\xc9', '\xbe', '\xc9', '\x9b', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'h', '\xc9', '\x91', 'j', '\xc9', '\x9b', '\xcb', '\x88', '\xc9', '\xbe', '\xc9', '\x9b', 'n']], 15, True)

['Armenian', 'Official', 'language', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Official', 'language', 'in', 'Nagorno-Karabakh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Official', 'language', 'in', '(', 'not', 'recognized', 'internationally', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', '\\(', 'not', 'recognized', 'internationally', '\\)']], 9, True)

['Armenian', 'Official', 'language', 'in', 'Minority', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Armenian', 'Official', 'language', 'in', ':', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Armenian', 'Official', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Armenian', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences', 'of', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'National', 'Academy', 'of', 'Sciences', 'of', 'Armenia']]], 9, True)

['Armenian', 'Language', 'codes', 'ISO', '639-2', 'arm', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'arm', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Armenian', 'Language', 'codes', 'ISO', '639-2', 'hye', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hye', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Armenian', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Armenian', 'Language', 'codes', 'ISO', '639-1', 'hy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Aramaic', 'Spoken', 'in', 'Assyrian', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Aramaic', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aramaic', 'Spoken', 'in', 'Palestine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Aramaic', 'Writing', 'system', 'Aramaic', 'abjad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Aramaic', 'Writing', 'system', 'Syriac', 'abjad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Aramaic', 'Writing', 'system', 'Hebrew', 'abjad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Aramaic', 'Writing', 'system', 'Mandaic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Aramaic', 'Pronunciation', '\xc9', '\x91', 'r', '\xc9', '\x91', 'm',  
'\xc9', '\x91', 'j', '\xc9', '\x91', ',', '\xc9', '\x94', 'r', '\xc9',  
'\x94', 'm', '\xc9', '\x94', 'j', '\xc9', '\x94']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',  
'\x91', 'r', '\xc9', '\x91', 'm', '\xc9', '\x91', 'j', '\xc9', '\x91', ',',  
'\xc9', '\x94', 'r', '\xc9', '\x94', 'm', '\xc9', '\x94', 'j', '\xc9',  
'\x94']]]], 25, True)
```

```
['Aramaic', 'Pronunciation', 'ar', '\xc9', '\x91', 'mi', '\xce', '\xb8']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ar',  
'\xc9', '\x91', 'mi', '\xce', '\xb8']]]], 8, True)
```

```
['Aramaic', 'Pronunciation', 'ar', '\xc9', '\x91', 'mit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ar',  
'\xc9', '\x91', 'mit']]]], 6, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'hrt', '\xe2', '\x80',  
'\x93', 'H', '\xc3', '\xa9', 'rtevin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'hrt', '\xe2', '\x80', '\x93', 'H', '\xc3', '\xa9',  
'rtevin']]], ['NUM', '639-3']], 13, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'huy', '\xe2', '\x80',  
'\x93', 'Hulaul', '\xc3', '\xa1']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```



```
'ISO', 'NUM', 'huy', '\xe2', '\x80', '\x93', 'Hulaul', '\xc3', '\xa1']],  
['NUM', '639-3']], 12, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'lhs', '\xe2', '\x80',  
\x93', 'Mlahs', '\xc3', '\xb4']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lhs', '\xe2', '\x80', '\x93', 'Mlahs', '\xc3', '\xb4']],  
['NUM', '639-3']], 12, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'syn', '\xe2', '\x80',  
\x93', 'Senaya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'syn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'trg', '\xe2', '\x80',  
\x93', 'Lish', '\xc3', '\xa1', 'n', 'Did', '\xc3', '\xa1', 'n']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'trg', '\xe2', '\x80', '\x93', 'Lish', '\xc3', '\xa1', 'n',  
'Did', '\xc3', '\xa1', 'n']], ['NUM', '639-3']], 17, True)
```

```
['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'tru', '\xe2', '\x80',  
\x93', 'Turoyo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tru', '\xe2', '\x80', '\x93', 'Turoyo']], ['NUM', '639-3']],  
10, True)
```

```
['Aramaic', 'Language', 'codes', 'Linguasphere', '12-AAA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Basque', 'Total', 'speakers', '665', ',', '800', 'in', '2006']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '665'], ['NUM', '800']], 8,  
True)
```

```
['Basque', 'Spoken', 'in', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Basque', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Basque', 'Region', 'Basque', 'Country']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Basque', 'Region', 'Navarre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Basque', 'Region', 'French', 'Basque', 'Country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Basque', 'Official', 'language', 'in', 'Basque', 'Country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Basque', 'Official', 'language', 'in', 'Navarre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Basque', 'Regulated', 'by', 'Euskaltzaindia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Basque', 'Language', 'codes', 'ISO', '639-2', 'baq', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'baq', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Basque', 'Language', 'codes', 'ISO', '639-2', 'eus', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'eus', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Basque', 'Language', 'codes', 'ISO', '639-3', 'eus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Basque', 'Language', 'codes', 'Linguasphere', '40-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Basque', 'Language', 'codes', 'ISO', '639-1', 'eu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Bulgarian', 'Total', 'speakers', '12', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '12']], 5, True)

['Bulgarian', 'Region', 'The', 'Balkans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bulgarian', 'Spoken', 'in', 'Kosovo', 'and', 'among', 'emigrant',  
'communities', 'worldwide']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kosovo',  
'and', 'among', 'emigrant', 'communities', 'worldwide']]], 9, True)

['Bulgarian', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Bulgarian', 'Official', 'language', 'in', 'Bulgaria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Bulgarian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bulgarian', 'Official', 'language', 'in', 'Mount', 'Athos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bulgarian', 'Recognised', 'minority', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Bulgarian', 'Recognised', 'minority', 'language', 'in', 'Romania', '(',  
'Banatian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Romania', '\\(', 'Banatian', 'variant', '\\)']]], 10,  
True)

['Bulgarian', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Bulgarian', 'Language', 'codes', 'ISO', '639-2', 'bul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Bulgarian', 'Language', 'codes', 'ISO', '639-3', 'bul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bulgarian', 'Language', 'codes', 'ISO', '639-1', 'bg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Bambara', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bambara', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bambara', 'Spoken', 'in', 'Guinea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bambara', 'Spoken', 'in', 'Senegal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bambara', 'Spoken', 'in', 'The', 'Gambia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bambara', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']]], 7, True)

['Bambara', 'Language', 'codes', 'ISO', '639-2', 'bam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Bambara', 'Language', 'codes', 'ISO', '639-3', 'bam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bambara', 'Language', 'codes', 'ISO', '639-1', 'bm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Breton', 'Region', 'Brittany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Breton', 'Spoken', 'in', 'France']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Breton', 'Pronunciation', 'b', '\xc9', '\xbe', 'e', '.', '\xcb', '\x88', 'z', '\xc3', '\xb5', '\xcb', '\x90', '.', 'n', '\xc9', '\x9b', 'k']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'b', '\xc9', '\xbe', 'e', '.', '\xcb', '\x88', 'z', '\xc3', '\xb5', '\xcb', '\x90', '.', 'n', '\xc9', '\x9b', 'k']]], 19, True)

['Breton', 'Language', 'codes', 'ISO', '639-2', 'bre']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Breton', 'Language', 'codes', 'ISO', '639-3', 'bre']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Breton', 'Language', 'codes', 'ISO', '639-1', 'br']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Blissymbols', 'Regulated', 'by', 'Blissymbolics', 'Communication', 'International']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Blissymbolics', 'Communication', 'Type: ANY']]], 6, True)

['Blissymbols', 'Total', 'speakers', 'some', 'thousands', 'of', 'users']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'some', 'thousands', 'of', 'users']]], 7, True)

['Blissymbols', 'Date', 'founded', '1949']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Blissymbols', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Blissymbols', 'Category', '(', 'purpose', ')', 'Blissymbols']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Blissymbols', 'Created', 'by', 'Charles', 'K', '.', 'Bliss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Charles', 'K', '.', 'Bliss']]], 7, True)

['Blissymbols', 'Category', '(', 'sources', ')', 'Ideographic', 'written', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(', 'sources', '\\\\)', 'Ideographic', 'written', 'language']]], 8, True)

['Blissymbols', 'Setting', 'and', 'usage', 'Augmentative', 'and', 'Alternative', 'Communication']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Augmentative', 'and', 'Alternative', 'Communication']]], 8, True)

['Blissymbols', 'Language', 'codes', 'ISO', '639-2', 'zbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```

['Blissymbols', 'Language', 'codes', 'ISO', '639-3', 'zbl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khmer', 'Spoken', 'in', 'Cambodia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'Vietnam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'Thailand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'USA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khmer', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khmer', 'Writing', 'system', 'Khmer', 'script', '(', 'abugida', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Khmer', 'Official', 'language', 'in', 'Cambodia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Khmer', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Khmer', 'Language', 'codes', 'ISO', '639-2', 'khm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Khmer', 'Language', 'codes', 'ISO', '639-3', 'either']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khmer', 'Language', 'codes', 'ISO', '639-1', 'km']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Cornish', 'Total', 'speakers', '2', ',', '000', 'fluent']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'fluent']], ['NUM', '2'], ['NUM', '000']], 7, True)

['Cornish', 'Region', 'Cornwall']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Cornish', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Cornish', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Cornish', 'Pronunciation', 'k', '\xc9', '\x99', 'r', '\xcb', '\x88', 'nu', '\xcb', '\x90', '\xc9', '\x99', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k', '\xc9', '\x99', 'r', '\xcb', '\x88', 'nu', '\xcb', '\x90', '\xc9', '\x99', 'k']], 14, True)

['Cornish', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Cornish', 'Recognised', 'minority', 'language', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'United', 'Kingdom']], 7, True)

['Cornish', 'Language', 'codes', 'ISO', '639-2', 'cor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cornish', 'Language', 'codes', 'ISO', '639-3', 'cor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Cornish', 'Language', 'codes', 'Linguasphere', '50-ABB-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Cornish', 'Language', 'codes', 'ISO', '639-1', 'kw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Czech', 'Total', 'speakers', '12', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']]], ['NUM', '12']], 5, True)

['Czech', 'Region', 'Central', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Czech', 'Spoken', 'in', 'Czech', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Czech', 'Spoken', 'in', 'Austria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Czech', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Czech', 'Spoken', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Czech', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Czech', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Czech', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Czech', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'and', 'Type: ANY']]], 6, True)

['Czech', 'Spoken', 'in', 'Romania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Czech', 'Writing', 'system', 'Czech', 'variant', 'of', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Czech', 'Official', 'language', 'in', 'Czech', 'Republic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Czech', 'Official', 'language', 'in', 'European', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Czech', 'Language', 'codes', 'ISO', '639-2', 'cze', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cze', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']]], 9, True)

['Czech', 'Language', 'codes', 'ISO', '639-2', 'ces', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ces', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']]], 9, True)

['Czech', 'Language', 'codes', 'ISO', '639-3', 'ces']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']]], 6, True)

['Czech', 'Language', 'codes', 'ISO', '639-1', 'cs']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']]], 6, True)

['Danish', 'Total', 'speakers', 'c', '.', '6', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c', '.', 'NUM', 'million']]], ['NUM', '6']]], 7, True)

['Danish', 'Spoken', 'in', 'Denmark', ',']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',']]], 5, True)

['Danish', 'Spoken', 'in', 'Faroe', 'Islands', ',']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Faroe', 'Islands', ',']]], 6, True)

```
['Danish', 'Spoken', 'in', 'Greenland', ', ']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ', ']]], 5, True)
```

```
['Danish', 'Spoken', 'in', 'Schleswig-Holstein', '(', 'Germany', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]]], 7, True)
```

```
['Danish', 'Writing', 'system', 'Latin', '(', 'Danish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]]], 8, True)
```

```
['Danish', 'Pronunciation', 'dan', '\\xcb', '\\x80', 's', '\\xc9', '\\xa1',  
'\\xcc', '\\x8a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'dan',  
'\\xcb', '\\x80', 's', '\\xc9', '\\xa1', '\\xcc', '\\x8a']]]], 10, True)
```

```
['Danish', 'Official', 'language', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['Danish', 'Official', 'language', 'in', 'Faroe', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Danish', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Danish', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Danish', 'Official', 'language', 'in', 'Minority', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Danish', 'Official', 'language', 'in', ':', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Danish', 'Language', 'codes', 'ISO', '639-2', 'dan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

['Danish', 'Language', 'codes', 'ISO', '639-3', 'dan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Danish', 'Language', 'codes', 'ISO', '639-1', 'da']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dioula', 'Total', 'speakers', '2', ',', '700', ',', '000', '(', 'several', 'millions', 'more', 'including', 'second', 'language', 'speakers', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', ',', 'NUM', '\\\\(', 'several', 'millions', 'more', 'including', 'second', 'language', 'speakers', '\\\\)']], ['NUM', '2'], ['NUM', '700'], ['NUM', '000']], 17, True)

['Dioula', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Dioula', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dioula', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central', 'southern', 'Mali', 'and', 'abroad']], 7, True)

['Dioula', 'Language', 'codes', 'ISO', '639-2', 'dyu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dioula', 'Language', 'codes', 'ISO', '639-3', 'dyu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Etruscan', 'Spoken', 'in', 'Ancient', 'Etruria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Etruscan', 'Region', 'Italian', 'Peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Etruscan', 'Language', 'codes', 'ISO', '639-3', 'ett']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Estonian', 'Region', 'Northern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Estonian', 'Spoken', 'in', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Estonian', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Estonian', 'Writing', 'system', 'Historically', 'Runic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Estonian', 'Official', 'language', 'in', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Estonian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Estonian', 'Language', 'codes', 'ISO', '639-2', 'est']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Estonian', 'Language', 'codes', 'ISO', '639-3', 'est', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'est', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Estonian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Estonian', 'Language', 'codes', 'ISO', '639-3', 'vro', '\xe2', '\x80', '\x93', 'V', '\xc3', '\xb5', 'ro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'vro', '\xe2', '\x80', '\x93', 'V', '\xc3', '\xb5', 'ro']], ['NUM', '639-3']], 13, True)

['Estonian', 'Language', 'codes', 'ISO', '639-1', 'et']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['West', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['West', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['West', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['West', 'Region', 'West', 'Flanders']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'vls']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['West', 'Language', 'codes', 'Linguasphere', '52-ACB-ag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Greek', 'Total', 'speakers', 'c', '.', '13', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c',  
'.', 'NUM', 'million']], ['NUM', '13']], 7, True)

['Greek', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Greek', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Greek', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Greek', 'Spoken', 'in', ':', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':', 'Type: ANY', 'Type: ANY']]], 6, True)

['Greek', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Greek', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Greek', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Greek', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Greek', 'Pronunciation', 'elini', '\xcb', '\x88', 'ka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'elini', '\xcb', '\x88', 'ka']]], 6, True)

['Greek', 'Official', 'language', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Greek', 'Official', 'language', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Greek', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Greek', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',



'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Greek', 'Language', 'codes', 'ISO', '639-2', 'gre', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'gre', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Greek', 'Language', 'codes', 'ISO', '639-2', 'ell', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ell', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Greek', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Greek', 'Language', 'codes', 'ISO', '639-1', 'el']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Gallurese', 'Spoken', 'in', 'Gallura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gallurese', 'Spoken', 'in', 'northeastern', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Gallurese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gallurese', 'Language', 'codes', 'ISO', '639-3', 'sdn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gallurese', 'Language', 'codes', 'Linguasphere', '51-AAA-pd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Hebrew', 'Total', 'speakers', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hebrew', 'Total', 'speakers', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'United', 'States']]], 5, True)

['Hebrew', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hebrew', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hebrew', 'Pronunciation', 'Yemenite', ':', '\xca', '\x95', 'iv', '\xcb', '\x88', 'ri', '\xcb', '\x90', '\xce', '\xb8', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'Yemenite', ':', '\xca', '\x95', 'iv', '\xcb', '\x88', 'ri', '\xcb', '\x90', '\xce', '\xb8', ',']]], 15, True)

['Hebrew', 'Pronunciation', 'Ashkenazi', ':', '\xcb', '\x88', 'iv', '\xca', '\x80', 'is']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'Ashkenazi', ':', '\xcb', '\x88', 'iv', '\xca', '\x80', 'is']]], 10, True)

['Hebrew', 'Official', 'language', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Hebrew', 'Official', 'language', 'in', 'Judaism']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Hebrew', 'Language', 'codes', 'ISO', '639-2', 'heb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Hebrew', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Hebrew', 'Language', 'codes', 'Linguasphere', '12-AAB-a']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Hebrew', 'Language', 'codes', 'ISO', '639-1', 'he']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Hindi', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Hindi', 'Spoken', 'in', 'Nepal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Hindi', 'Spoken', 'in', 'Mauritius', 'and', 'significant', 'communities',  
'in', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Mauritius', 'and', 'significant', 'communities', 'in', 'USA']], 9, True)
```

```
['Hindi', 'Spoken', 'in', 'UK']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Hindi', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Hindi', 'Spoken', 'in', 'Malaysia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Hindi', 'Spoken', 'in', 'Singapore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Hindi', 'Spoken', 'in', 'United', 'Arab', 'Emirates']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'Arab', 'Emirates']], 6, True)
```

['Hindi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hindi', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hindi', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hindi', 'Official', 'language', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Hindi', 'Language', 'codes', 'ISO', '639-2', 'hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hindi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Hindi', 'Language', 'codes', 'ISO', '639-3', 'hin-hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Hindi', 'Language', 'codes', 'ISO', '639-1', 'hi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Hausa', 'Spoken', 'in', 'Benin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Hausa', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Spoken', 'in', 'Togo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Hausa', 'Total', 'speakers', '25', 'million', 'as', 'a', 'first', 'language', ',', '18', 'million', 'as', 'a', 'second', 'language']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'as', 'a', 'first', 'language', ',', 'NUM', 'million', 'as', 'a', 'second', 'language']], ['NUM', '25'], ['NUM', '18']], 16, True)

['Hausa', 'Writing', 'system', 'Latin', ',', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Hausa', 'Language', 'codes', 'ISO', '639-2', 'hau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hausa', 'Language', 'codes', 'ISO', '639-3', 'hau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hausa', 'Language', 'codes', 'ISO', '639-1', 'ha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Hawaiian', 'Official', 'language', 'in', 'recognized', 'as', 'minority', 'language', 'in', 'parts', 'of']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognized', 'as', 'minority', 'language', 'in', 'parts', 'of']], 11, True)

['Hawaiian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Hawaiian', 'Language', 'codes', 'ISO', '639-2', 'haw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hawaiian', 'Language', 'codes', 'ISO', '639-3', 'haw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Italian', 'Total', 'speakers', '62', 'million', 'or', '65', 'million', 'in', 'the', 'European', 'Union', ';', 'more', 'than', '20', 'million', 'as', 'second', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'or', 'NUM', 'million', 'in', 'the', 'European', 'Union', ';', 'more', 'than', 'NUM', 'million', 'as', 'second', 'language']], ['NUM', '62'], ['NUM', '65'], ['NUM', '20']], 20, True)

['Italian', 'Region', 'al', 'in', 'Slovenia', 'and', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'al', 'in', 'Slovenia', 'and', 'Croatia']]], 7, True)

['Italian', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'San', 'Marino']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Italian', 'Spoken', 'in', 'Vatican', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Italian', 'Spoken', 'in', 'Malta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Italian', 'Spoken', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Italian', 'Pronunciation', 'ita', '\xcb', '\x88', 'ljano']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ita', '\xcb', '\x88', 'ljano']]], 6, True)

['Italian', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Italian', 'Official', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Italian', 'Official', 'language', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['Italian', 'Official', 'language', 'in', 'San', 'Marino']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Italian', 'Official', 'language', 'in', 'Vatican', 'City']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Italian', 'Official', 'language', 'in', 'Sovereign', 'Military', 'Order',  
'of', 'Malta']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Sovereign', 'Military', 'Order', 'of', 'Malta']]], 9, True)

['Italian', 'Official', 'language', 'in', 'Croatia', '(', 'Istria',  
'County', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Croatia', '\\(', 'Istria', 'County', '\\)']]], 9, True)

['Italian', 'Official', 'language', 'in', 'Slovenia', '(', 'Slovenian',  
'Istria', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Slovenia', '\\(', 'Slovenian', 'Istria', '\\)']]], 9, True)

['Italian', 'Regulated', 'by', 'not', 'officially', 'by', 'Accademia',  
'della', 'Crusca']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'not',  
'officially', 'by', 'Accademia', 'della', 'Crusca']]], 9, True)

['Italian', 'Language', 'codes', 'ISO', '639-2', 'ita']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Italian', 'Language', 'codes', 'ISO', '639-3', 'ita']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Italian', 'Language', 'codes', 'Linguasphere', '51-AAA-q']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Italian', 'Language', 'codes', 'ISO', '639-1', 'it']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)



['Ido', 'Regulated', 'by', 'Uniono', 'por', 'la', 'Linguo', 'Internaciona', 'Ido']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Uniono', 'por', 'la', 'Linguo', 'Internaciona', 'Type: ANY']]], 9, True)

['Ido', 'Total', 'speakers', '(', 'but', 'see', 'community', 'section', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\(', 'but', 'see', 'community', 'section', '\\)']]], 9, True)

['Ido', 'Date', 'founded', '1907']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ido', 'Category', '(', 'purpose', ')', 'constructed', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Ido', 'Category', '(', 'purpose', ')', 'Ido']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Ido', 'Category', '(', 'sources', ')', 'based', 'on', 'Esperanto']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'based', 'on', 'Type: ANY']]], 8, True)

['Ido', 'Language', 'codes', 'ISO', '639-2', 'ido']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Ido', 'Language', 'codes', 'ISO', '639-3', 'ido']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ido', 'Language', 'codes', 'Linguasphere', '51-AAB-db']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Ido', 'Language', 'codes', 'ISO', '639-1', 'io']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'io']]], ['NUM', '639-1']], 6, True)

['Interlingua', 'Regulated', 'by', 'no', 'regulating', 'body']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'no', 'regulating', 'body']]], 6, True)

['Interlingua', 'Total', 'speakers', 'a', 'few', 'hundred']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a',  
'few', 'hundred']]], 6, True)

['Interlingua', 'Date', 'founded', '1951']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Interlingua', 'Category', '(', 'purpose', ')', 'international',  
'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'international', 'auxiliary', 'language']]], 8, True)

['Interlingua', 'Category', '(', 'purpose', ')', 'Interlingua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Interlingua', 'Writing', 'system', 'Latin', 'alphabet', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Interlingua', 'Writing', 'system', 'Italian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'Portuguese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'Spanish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'German']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'English']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Interlingua', 'Writing', 'system', 'Russian', 'and', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Interlingua', 'Pronunciation', '/', '\\xc9', '\\xaa', 'nt', '\\xc9', '\\x99',

'r', '\xcb', '\x88', 'l', '\xc9', '\xaa', '\xc5', '\x8b', '\xc9', '\xa1',  
'w', '\xc9', '\x99', '/', ';', 'IA', ':', 'inte', '\xc9', '\xbe', '\xcb',  
'\x88', 'li', '\xc5', '\x8b', '\xc9', '\xa1', 'wa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'\xc9', '\xaa', 'nt', '\xc9', '\x99', 'r', '\xcb', '\x88', 'l', '\xc9',  
'\xaa', '\xc5', '\x8b', '\xc9', '\xa1', 'w', '\xc9', '\x99', '/', ';',  
'IA', ':', 'inte', '\xc9', '\xbe', '\xcb', '\x88', 'li', '\xc5', '\x8b',  
'\xc9', '\xa1', 'wa']]], 36, True)

['Interlingua', 'Setting', 'and', 'usage', 'Scientific', 'registration',  
'of', 'international', 'vocabulary', ';', 'international', 'auxiliary',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Scientific', 'registration', 'of', 'international', 'vocabulary',  
';', 'international', 'auxiliary', 'language']]], 13, True)

['Interlingua', 'Language', 'codes', 'ISO', '639-2', 'ina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Interlingua', 'Language', 'codes', 'ISO', '639-3', 'ina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Interlingua', 'Language', 'codes', 'ISO', '639-1', 'ia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Japanese', 'Total', 'speakers', '130', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '130']], 5, True)

['Japanese', 'Spoken', 'in', 'Majority', ':', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Majority', ':', 'Japan']]], 6, True)

['Japanese', 'Writing', 'system', 'Chinese', 'characters']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Japanese', 'Writing', 'system', 'r', '\xc5', '\x8d', 'maji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'r',  
'\xc5', '\x8d', 'maji']]], 7, True)

['Japanese', 'Writing', 'system', 'Siddham', 'script', '(', 'occasionally',  
'in', 'Buddhist', 'temples', '.', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Siddham', 'script', '\\(', 'occasionally', 'in', 'Buddhist', 'temples', '.', '\\)']]], 12, True)

['Japanese', 'Pronunciation', 'n', '\\xca', '\\xb2', 'iho', '\\xc5', '\\x8b', '\\xc9', '\\xa1', 'o']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\\xca', '\\xb2', 'iho', '\\xc5', '\\x8b', '\\xc9', '\\xa1', 'o']]], 11, True)

['Japanese', 'Official', 'language', 'in', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Japanese', 'Official', 'language', 'in', 'Palau', '(', 'Angaur', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Palau', '\\(', 'Angaur', '\\)']]], 8, True)

['Japanese', 'Regulated', 'by', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Japanese', 'Language', 'codes', 'ISO', '639-2', 'jpn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Japanese', 'Language', 'codes', 'ISO', '639-3', 'jpn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Japanese', 'Language', 'codes', 'ISO', '639-1', 'ja']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Korean', 'Spoken', 'in', 'South', 'Korea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Korean', 'Spoken', 'in', 'North', 'Korea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Korean', 'Spoken', 'in', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Korean', 'Writing', 'system', 'Exclusive', 'use', 'of', 'Hangul', '(',

'Both', 'Koreas', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Exclusive', 'use', 'of', 'Hangul', '\\(', 'Both', 'Koreas', '\\)']]], 11,  
True)

['Korean', 'Writing', 'system', 'mix', 'of', 'Hangul', 'and', 'hanja', '(',  
'some', 'professional', 'scripts', 'in', 'S', '.', 'Korea', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'mix', 'of', 'Hangul', 'and', 'hanja', '\\(', 'some', 'professional',  
'scripts', 'in', 'S', '.', 'Korea', '\\)']]], 17, True)

['Korean', 'Writing', 'system', 'or', 'Cyrillic', 'alphabet', '(',  
'lesser', 'used', 'in', 'Goryeomal', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'or', 'Cyrillic', 'alphabet', '\\(', 'lesser', 'used', 'in', 'Goryeomal',  
'\\)']]], 12, True)

['Korean', 'Official', 'language', 'in', 'North', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Korean', 'Official', 'language', 'in', 'South', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Korean', 'Regulated', 'by', 'South', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Korean', 'Regulated', 'by', '\\xea', '\\xb5', '\\xad', '\\xeb', '\\xa6',  
'\\xbd', '\\xea', '\\xb5', '\\xad', '\\xec', '\\x96', '\\xb4', '\\xec', '\\x9b',  
'\\x90', '/', '\\xe5', '\\x9c', '\\x8b', '\\xe7', '\\xab', '\\x8b', '\\xe5',  
'\\x9c', '\\x8b', '\\xe8', '\\xaa', '\\x9e', '\\xe9', '\\x99', '\\xa2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'\\xea', '\\xb5', '\\xad', '\\xeb', '\\xa6', '\\xbd', '\\xea', '\\xb5', '\\xad',  
'\\xec', '\\x96', '\\xb4', '\\xec', '\\x9b', '\\x90', '/', '\\xe5', '\\x9c',  
'\\x8b', '\\xe7', '\\xab', '\\x8b', '\\xe5', '\\x9c', '\\x8b', '\\xe8', '\\xaa',  
'\\x9e', '\\xe9', '\\x99', '\\xa2']]], 34, True)

['Korean', 'Regulated', 'by', 'North', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Korean', 'Regulated', 'by', ':', 'Sahoe', 'Kwahagwon', '\\xc5', '\\x8e',  
'hak', 'Y', '\\xc5', '\\x8f', 'nguso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', ':',

```
'Sahoe', 'Kwahagwon', '\xc5', '\x8e', 'hak', 'Y', '\xc5', '\x8f',  
'nguso']]]], 13, True)
```

```
['Korean', 'Regulated', 'by', '\xec', '\x82', '\xac', '\xed', '\x9a',  
'\x8c', '\xea', '\xb3', '\xbc', '\xed', '\x95', '\x99', '\xec', '\x9b',  
'\x90', '\xec', '\x96', '\xb4', '\xed', '\x95', '\x99', '\xec', '\x97',  
'\xb0', '\xea', '\xb5', '\xac', '\xec', '\x86', '\x8c']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'\xec', '\x82', '\xac', '\xed', '\x9a', '\x8c', '\xea', '\xb3', '\xbc',  
'\xed', '\x95', '\x99', '\xec', '\x9b', '\x90', '\xec', '\x96', '\xb4',  
'\xed', '\x95', '\x99', '\xec', '\x97', '\xb0', '\xea', '\xb5', '\xac',  
'\xec', '\x86', '\x8c']]]], 33, True)
```

```
['Korean', 'Language', 'codes', 'ISO', '639-2', 'kor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Korean', 'Language', 'codes', 'ISO', '639-3', 'kor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Korean', 'Language', 'codes', 'Linguasphere', '45-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]]], 5, True)
```

```
['Korean', 'Language', 'codes', 'ISO', '639-1', 'ko']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Klingon', 'Total', 'speakers', 'Unknown', '.', 'Around', '12', 'fluent',  
'speakers', 'in', '1996', ',', 'according', 'to', 'Lawrence', 'Schoen',  
,', 'director', 'of', 'the', 'KLI']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Unknown', '.', 'Around', 'NUM', 'fluent', 'speakers', 'in', 'NUM', ',',  
'according', 'to', 'Lawrence', 'Schoen', ',', 'director', 'of', 'the',  
'KLI']], ['NUM', '12'], ['NUM', '1996']], 21, True)
```

```
['Klingon', 'Date', 'founded', '1984']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Klingon', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]]], 7, True)
```

```
['Klingon', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Klingon', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Klingon', 'Category', '(', 'purpose', ')', 'Klingon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Klingon', 'Writing', 'system', 'a', 'priori', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Klingon', 'Pronunciation', '/', '\\xcb', '\\x88', 't', '\\xcd', '\\xa1',  
'\\xc9', '\\xac', '\\xc9', '\\xaa', '\\xc5', '\\x8b', '\\xc9', '\\x91', 'n', 'xol',  
'/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'\\xcb', '\\x88', 't', '\\xcd', '\\xa1', '\\xc9', '\\xac', '\\xc9', '\\xaa',  
'\\xc5', '\\x8b', '\\xc9', '\\x91', 'n', 'xol', '/']]], 19, True)

['Klingon', 'Created', 'by', 'Marc', 'Okrand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Klingon', 'Language', 'codes', 'ISO', '639-2', 'tlh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Klingon', 'Language', 'codes', 'ISO', '639-3', 'tlh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kashubian', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kashubian', 'Region', 'Pomerania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kashubian', 'Official', 'language', 'in', 'In', 'official', 'use']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'In', 'official', 'use']]], 7, True)

['Kashubian', 'Official', 'language', 'in', 'as', 'a', 'regional',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'as', 'a', 'regional', 'language']]], 8, True)

['Kashubian', 'Official', 'language', 'in', 'in', 'some', 'communes', 'of',  
'Pomeranian', 'Voivodeship']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'in', 'some', 'communes', 'of', 'Pomeranian', 'Voivodeship']]], 10,  
True)

['Kashubian', 'Official', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Kashubian', 'Language', 'codes', 'ISO', '639-2', 'csb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kashubian', 'Language', 'codes', 'ISO', '639-3', 'csb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kashubian', 'Language', 'codes', 'Linguasphere', '53-AAA-cb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Loglan', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Loglan', 'Date', 'founded', '1955']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Loglan', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Loglan', 'Category', '(', 'purpose', ')', 'engineered', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Loglan', 'Category', '(', 'purpose', ')', 'logical', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)



```

['Loglan', 'Category', '(', 'purpose', ')', 'Loglan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY']]], 6, True)

['Loglan', 'Created', 'by', 'James', 'Cooke', 'Brown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'James',
'Cooke', 'Brown']]], 6, True)

['Loglan', 'Setting', 'and', 'usage', 'engineered', 'language', 'for',
'testing', 'the', 'Sapir', '\\e2', '\\x80', '\\x93', 'Whorf', 'Hypothesis',
'and', 'other', 'linguistic', 'research']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',
'usage', 'engineered', 'language', 'for', 'testing', 'the', 'Sapir',
'\\e2', '\\x80', '\\x93', 'Whorf', 'Hypothesis', 'and', 'other',
'linguistic', 'research']]], 19, True)

['Loglan', 'Language', 'codes', 'ISO', '639-3', '\\e2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\e2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Maltese', 'Spoken', 'in', 'Malta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Maltese', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Maltese', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Maltese', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Maltese', 'Spoken', 'in', 'United', 'Kingdom']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Maltese', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Maltese', 'Writing', 'system', 'Maltese-variant', 'of', 'the', 'Latin',

```

```

'Alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Maltese', 'Official', 'language', 'in', 'Malta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Maltese', 'Official', 'language', 'in', 'European', 'Union']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Maltese', 'Regulated', 'by', 'Il-Kunsill', 'Nazzjonali', 'tal-Ilsien',
'Malti']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Il-
Kunsill', 'Nazzjonali', 'tal-Ilsien', 'Malti']]], 7, True)

['Maltese', 'Language', 'codes', 'ISO', '639-2', 'mlt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Maltese', 'Language', 'codes', 'ISO', '639-3', 'mlt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maltese', 'Language', 'codes', 'Linguasphere', '12-AAC-c']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Maltese', 'Language', 'codes', 'ISO', '639-1', 'mt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Mandarin', 'Total', 'speakers', 'Native', ':', '845', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'Native', ':', 'NUM', 'million']], ['NUM', '845']], 7, True)

['Mandarin', 'Dialects', 'Northeastern']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mandarin', 'Dialects', 'Beijing']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mandarin', 'Dialects', 'Ji-Lu']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Jiao-Liao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Lower', 'Yangtze']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Dialects', 'Central', 'Plains']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Dialects', 'Lan-Yin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Dialects', 'Southwestern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mandarin', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Mandarin', 'Language', 'codes', 'ISO', '639-3', 'cmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Mandarin', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Manx', 'Total', 'speakers', 'Extinct', 'as', 'a', 'first', 'language',  
'in', '1974', ';', 'subsequently', 'revived', 'and', 'now', 'with',  
'about', 'a', 'hundred', 'competent', 'speakers', ',', 'including', 'a',  
'small', 'number', 'of', 'children', 'who', 'are', 'new', 'native',  
'speakers', ',', ']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Extinct', 'as', 'a', 'first', 'language', 'in', 'NUM', ';',  
'subsequently', 'revived', 'and', 'now', 'with', 'about', 'a', 'hundred',  
'competent', 'speakers', ',', 'including', 'a', 'small', 'number', 'of',  
'children', 'who', 'are', 'new', 'native', 'speakers', ',', ']], ['NUM',  
'1974']], 34, True)
```

```
['Manx', 'Pronunciation', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xa3',  
'ilk', ',', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xa3', 'il', '\xc9',  
'\xa1']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',  
'\x99', '\xcb', '\x88', '\xc9', '\xa3', 'ilk', ',', '\xc9', '\x99', '\xcb',  
'\x88', '\xc9', '\xa3', 'il', '\xc9', '\xa1']]], 19, True)
```

```
['Manx', 'Spoken', 'in', 'Isle', 'of', 'Man']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle',  
'of', 'Man']]], 6, True)
```

```
['Manx', 'Official', 'language', 'in', 'Isle', 'of', 'Man']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Isle', 'of', 'Man']]], 7, True)
```

```
['Manx', 'Language', 'codes', 'ISO', '639-2', 'glv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Manx', 'Language', 'codes', 'ISO', '639-3', 'glv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Manx', 'Language', 'codes', 'Linguasphere', '50-AAA-aj']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Manx', 'Language', 'codes', 'ISO', '639-1', 'gv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Marathi', 'Total', 'speakers', 'Total', '90', 'million', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Total', 'NUM', 'million', 'speakers']], ['NUM', '90']], 7, True)

['Marathi', 'Total', 'speakers', '70', 'million', 'native', ',', '20',  
'million', 'second', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', 'native', ',', 'NUM', 'million', 'second', 'language']],  
['NUM', '70'], ['NUM', '20']], 11, True)

['Marathi', 'Region', 'Maharashtra']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Marathi', 'Region', 'parts', 'of', 'Goa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Marathi', 'Region', 'Gujarat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Marathi', 'Region', 'Madhya', 'Pradesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Region', 'Karnataka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Marathi', 'Region', 'Andhra', 'Pradesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'India', ',', 'Mauritius', 'and', 'Israel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
,', 'Mauritius', 'and', 'Israel']]], 8, True)

['Marathi', 'Spoken', 'in', 'Marathi-speaking', 'populations', 'are',  
'found', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Marathi-  
speaking', 'populations', 'are', 'found', 'in', 'United', 'States']]], 10,  
True)

['Marathi', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Guyana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Trinidad', 'and', 'Tobago']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Marathi', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Marathi', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marathi', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Marathi', 'Writing', 'system', 'Devanagari', 'script', '(', 'standard', ')', ',', 'Modi', 'script', '(', 'shorthand', 'script', ')']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devanagari', 'script', '\\(', 'standard', '\\)', ',', 'Modi', 'script', '\\(', 'shorthand', 'script', '\\)']]], 15, True)

['Marathi', 'Pronunciation', 'm', '\\xc9', '\\x99', '\\xcb', '\\x88', '\\xc9', '\\xbe', 'a', '\\xca', '\\x88', '\\xca', '\\xb0', 'i']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm', '\\xc9', '\\x99', '\\xcb', '\\x88', '\\xc9', '\\xbe', 'a', '\\xca', '\\x88', '\\xca', '\\xb0', 'i']]], 15, True)

['Marathi', 'Official', 'language', 'in', 'India', '(', 'State', 'of', 'Maharashtra', ',', 'Union', 'territories', 'of', 'Daman-Diu', ')', 'and', 'Dadra', 'Nagar', 'Haveli']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\(', 'State', 'of', 'Maharashtra', ',', 'Union', 'territories', 'of', 'Daman-Diu', '\\)', 'and', 'Dadra', 'Nagar', 'Haveli']]], 19, True)

['Marathi', 'Language', 'codes', 'ISO', '639-2', 'mar']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Marathi', 'Language', 'codes', 'ISO', '639-3', 'mar']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marathi', 'Language', 'codes', 'ISO', '639-1', 'mr']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Norwegian', 'Total', 'speakers', '5', 'million', 'Norwegians']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'Norwegians']], ['NUM', '5']], 6, True)

['Norwegian', 'Writing', 'system', 'Latin', '(', 'Norwegian', 'variant', ')']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Norwegian', 'Pronunciation', 'n', '\\xc9', '\\x94', '\\xca', '\\x82', 'k']]:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\\xc9', '\\x94', '\\xca', '\\x82', 'k']]], 8, True)

['Norwegian', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Norwegian', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nor', '\xe2', '\x80',  
'\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nob', '\xe2', '\x80',  
'\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']]],  
['NUM', '639-2']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nno', '\xe2', '\x80',  
'\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']]], ['NUM', '639-2']],  
10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nor', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nob', '\xe2',  
'\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5',  
'l']]], ['NUM', '639-3']], 14, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nno', '\xe2', '\x80',  
'\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-3']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'no', '\xe2', '\x80', '\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-1']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-1']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-1']], 10, True)

['Northern', 'Dialects', 'Lobedu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Northern', 'Region', 'Gauteng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Northern', 'Region', 'Limpopo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Northern', 'Region', 'Mpumalanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Northern', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Northern', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-2', 'nso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'nso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Novial', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Novial', 'Date', 'founded', '1928']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Novial', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Novial', 'Category', '(', 'purpose', ')', 'international', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'international', 'auxiliary', 'language']], 8, True)

['Novial', 'Category', '(', 'purpose', ')', 'Novial']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)

['Novial', 'Created', 'by', 'Otto', 'Jespersen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Novial', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']], 7, True)

['Novial', 'Language', 'codes', 'ISO', '639-3', 'nov']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Novial', 'Language', 'codes', 'Linguasphere', '51-AAB-dc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Norn', 'Spoken', 'in', 'Shetland', ',', 'Orkney', 'and', 'Caithness']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Shetland', ',', 'Orkney', 'and', 'Caithness']]], 8, True)

['Norn', 'Language', 'codes', 'ISO', '639-3', 'nrn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norn', 'Language', 'codes', 'Linguasphere', '52-AAA-ac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Prussian', 'Spoken', 'in', 'Prussia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Prussian', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Prussian', 'Language', 'codes', 'ISO', '639-3', 'prg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Prussian', 'Language', 'codes', 'Linguasphere', '54-AAC-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Old', 'Spoken', 'in', 'Nordic', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'England', 'and', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Wales']]], 6, True)

['Old', 'Spoken', 'in', 'Isle', 'of', 'Man']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Isle', 'of', 'Man']]], 6, True)

```
['Old', 'Spoken', 'in', 'Normandy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Old', 'Spoken', 'in', 'Vinland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Old', 'Spoken', 'in', 'the', 'Volga', 'and', 'places', 'in-between']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'the',  
'Volga', 'and', 'places', 'in-between']]], 8, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'non']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'non']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'England', '(', 'except', 'the', 'extreme',  
'southwest', 'and', 'northwest', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'England', '\\\\(', 'except', 'the', 'extreme', 'southwest', 'and',  
'northwest', '\\\\)']]], 12, True)
```

```
['Old', 'Spoken', 'in', 'parts', 'of', 'modern', 'Scotland', 'south-east',  
'of', 'the', 'Forth']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',  
'of', 'modern', 'Scotland', 'south-east', 'of', 'the', 'Forth']]], 11,  
True)
```

```
['Old', 'Spoken', 'in', 'and', 'the', 'eastern', 'fringes', 'of', 'modern',  
'Wales']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',  
'the', 'eastern', 'fringes', 'of', 'modern', 'Wales']]], 10, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'ang']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'ang']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Occidental', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]]], 6, True)
```

```
['Occidental', 'Category', '(', 'purpose', ')', 'constructed', 'language',  
'based', 'on', 'Indo-European', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'constructed', 'language', 'based', 'on', 'Indo-  
European', 'languages']]]], 11, True)
```

```
['Occidental', 'Language', 'codes', 'ISO', '639-2', 'ile']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Occidental', 'Language', 'codes', 'ISO', '639-3', 'ile']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Occidental', 'Language', 'codes', 'ISO', '639-1', 'ie']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Polish', 'Total', 'speakers', '40', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']]], ['NUM', '40']], 5, True)
```

```
['Polish', 'Spoken', 'in', 'Poland', '.', 'Minorities', ':', 'Belarus']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Poland',  
'.', 'Minorities', ':', 'Belarus']]]], 8, True)
```

```
['Polish', 'Spoken', 'in', 'Ukraine']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Polish', 'Spoken', 'in', 'Lithuania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Polish', 'Spoken', 'in', 'Latvia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Polish', 'Spoken', 'in', 'United', 'Kingdom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)
```

['Polish', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Polish', 'Spoken', 'in', 'Czech', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Polish', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Polish', 'Writing', 'system', 'Latin', '(', 'Polish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Polish', 'Pronunciation', '\xcb', '\x88', 'p', '\xc9', '\x94', 'lski']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',

'\x88', 'p', '\xc9', '\x94', 'lski']]], 8, True)

['Polish', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Polish', 'Official', 'language', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Official', 'language', 'in', 'Minority', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Polish', 'Official', 'language', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Official', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Polish', 'Language', 'codes', 'ISO', '639-2', 'pol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Polish', 'Language', 'codes', 'ISO', '639-3', 'pol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Polish', 'Language', 'codes', 'ISO', '639-1', 'pl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Portuguese', 'Region', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Portuguese', 'Region', 'the', 'Americas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Portuguese', 'Region', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Portuguese', 'Region', 'Europe', 'and', 'Oceania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Portuguese', 'Spoken', 'in', 'See', 'geographic', 'distribution', 'of', 'Portuguese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'See', 'geographic', 'distribution', 'of', 'Type: ANY']], 8, True)

['Portuguese', 'Pronunciation', 'pu', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\xca', '\x83', '(', 'European', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pu', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\xca', '\x83', '\\(', 'European', '\\)', ',']], 17, True)

['Portuguese', 'Pronunciation', 'po', '\xcf', '\x87', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '(', 'j', ')', '\xca', '\x83', '(', 'BP-carioca', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xcf', '\x87', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\\(', 'j', '\\)', '\xca', '\x83', '\\(', 'BP-carioca', '\\)', ',']], 20, True)

['Portuguese', 'Pronunciation', 'po', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '(', 'j', ')', 's', '(', 'BP-paulistano', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\\(', 'j', '\\)', 's', '\\(', 'BP-paulistano', '\\)', ',']], 19, True)

['Portuguese', 'Pronunciation', 'po', '\xc9', '\xb9', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'ejs', '(', 'BP-caipira', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xc9', '\xb9', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'ejs', '\\(', 'BP-caipira', '\\)', ',']], 15, True)

['Portuguese', 'Pronunciation', 'po', '\xcf', '\x87', '(', 'h', ')', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '(', 'j', ')', 's', '(', 'BP-mineiro', ')', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'po', '\xcf', '\x87', '\\(', 'h', '\\)', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\\(', 'j', '\\)', 's', '\\(', 'BP-mineiro', '\\)', ',']], 22, True)

['Portuguese', 'Pronunciation', 'p', '\xc9', '\x94', 'htu', '\xcb', '\x88',



'\xc9', '\xa1', 'ejs', '(', 'BP-nordestino', ')', ', ']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p',  
'\xc9', '\x94', 'htu', '\xcb', '\x88', '\xc9', '\xa1', 'ejs', '\\(', 'BP-  
nordestino', '\\)'], ', ']]], 15, True)

['Portuguese', 'Official', 'language', 'in', '8', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)

['Portuguese', 'Official', 'language', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Cape', 'Verde']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)

['Portuguese', 'Official', 'language', 'in', 'East', 'Timor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)

['Portuguese', 'Official', 'language', 'in', 'Guinea-Bissau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'S', '\xc3', '\xa3', 'o',  
'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3',  
'\xad', 'ncipe']]]], 16, True)

['Portuguese', 'Official', 'language', 'in', '1', 'dependent', 'entity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'NUM', 'dependent', 'entity']], ['NUM', '1']], 7, True)

['Portuguese', 'Official', 'language', 'in', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Portuguese', 'Official', 'language', 'in', 'Numerous', 'international', 'organisations']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Numerous', 'international', 'organisations']]], 7, True)

['Portuguese', 'Language', 'codes', 'ISO', '639-2', 'por']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Portuguese', 'Language', 'codes', 'ISO', '639-3', 'por']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Portuguese', 'Language', 'codes', 'Linguasphere', '51-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Portuguese', 'Language', 'codes', 'ISO', '639-1', 'pt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Standard', 'Spoken', 'in', 'Republic', 'of', 'China', '(', 'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'China', '\\(', 'Taiwan', '\\)']]], 9, True)

['Standard', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Standard', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Standard', 'Official', 'language', 'in', 'Republic', 'of', 'China', '(', 'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'China', '\\(', 'Taiwan', '\\)']]], 10, True)

['Standard', 'Regulated', 'by', 'In', 'the', 'ROC', ':', 'National', 'Languages', 'Committee']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'In', 'the', 'ROC', ':', 'National', 'Languages', 'Committee']]], 10, True)

['Standard', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Punjabi', 'Total', 'speakers', '57', ',', '129', ',', '000', '(',  
'Encarta', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'Encarta', '\\)']], ['NUM', '57'],  
['NUM', '129'], ['NUM', '000']], 11, True)

['Punjabi', 'Region', 'Punjab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Punjabi', 'Spoken', 'in', 'Pakistan', 'and', 'India', '.', 'Minor',  
'populations', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Pakistan', 'and', 'India', '.', 'Minor', 'populations', 'in', 'United',  
'Kingdom']]], 12, True)

['Punjabi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Punjabi', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Punjabi', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Punjabi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Punjabi', 'Spoken', 'in', 'United', 'Arab', 'Emirates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'Arab', 'Emirates']]], 6, True)

['Punjabi', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Punjabi', 'Writing', 'system', 'Gurmukhi', 'in', 'Punjab', '(', 'India', ')', 'and', 'Sikh', 'diaspora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Gurmukhi', 'in', 'Punjab', '\\\\(', 'India', '\\\\)', 'and', 'Sikh', 'diaspora']]], 12, True)

['Punjabi', 'Writing', 'system', 'Shahmukhi', 'in', 'Punjab', '(', 'Pakistan', ')', 'and', 'Sarhad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Shahmukhi', 'in', 'Punjab', '\\\\(', 'Pakistan', '\\\\)', 'and', 'Sarhad']]], 11, True)

['Punjabi', 'Official', 'language', 'in', 'India', '(', 'Indian', 'Punjab', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\\\(', 'Indian', 'Punjab', '\\\\)']]], 9, True)

['Punjabi', 'Official', 'language', 'in', 'Pakistan', '(', 'Pakistani', 'Punjab', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Pakistan', '\\\\(', 'Pakistani', 'Punjab', '\\\\)']]], 9, True)

['Punjabi', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Punjabi', 'Language', 'codes', 'ISO', '639-2', 'pan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Punjabi', 'Language', 'codes', 'ISO', '639-3', 'pan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Punjabi', 'Language', 'codes', 'ISO', '639-1', 'pa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Romansh', 'Spoken', 'in', 'Switzerland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Romansh', 'Region', 'Graub', '\\xc3', '\\xbc', 'nden']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Graub', '\\xc3', '\\xbc', 'nden']]], 6, True)

```
['Romansh', 'Official', 'language', 'in', 'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Romansh', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Romansh', 'Language', 'codes', 'ISO', '639-2', 'roh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Romansh', 'Language', 'codes', 'ISO', '639-3', 'roh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Romansh', 'Language', 'codes', 'Linguasphere', '51-AAA-k']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Romansh', 'Language', 'codes', 'ISO', '639-1', 'rm']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Romanian,', 'Daco-Romanian', 'Total', 'speakers', 'First', 'language',  
, ':', '24', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Daco-Romanian', 'Total',  
'speakers', 'First', 'language', ':', 'NUM', 'million']]], ['NUM', '24']],  
9, True)
```

```
['Romanian,', 'Total', 'speakers', 'Second', 'language', ':', '4',  
'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Second', 'language', ':', 'NUM', 'million']]], ['NUM', '4']], 8, True)
```

```
['Romanian,', 'Spoken', 'in', 'By', 'a', 'majority']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'By',  
'a', 'majority']]], 6, True)
```

```
['Romanian,', 'Spoken', 'in', ':', 'Romania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',  
'Type: ANY']]], 5, True)
```

```
['Romanian,', 'Spoken', 'in', 'Republic', 'of', 'Moldova']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
```

'Republic', 'of', 'Moldova']]], 6, True)

['Romanian,', 'Spoken', 'in', 'Minority', 'speakers', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Minority', 'speakers', 'in']]], 6, True)

['Romanian,', 'Spoken', 'in', ':', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',  
'Type: ANY']]], 5, True)

['Romanian,', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Migrant', 'speakers', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Migrant', 'speakers', 'in']]], 6, True)

['Romanian,', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romanian,', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Romanian,', 'Pronunciation', 'ro', '\xcb', '\x88', 'm', '\xc9', '\xa8',  
'n', '\xc9', '\x99']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ro',  
'\xcb', '\x88', 'm', '\xc9', '\xa8', 'n', '\xc9', '\x99']]], 11, True)

['Romanian,', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Romanian,', 'Official', 'language', 'in', 'Moldova']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romanian,', 'Official', 'language', 'in', 'Vojvodina', '(', 'Serbia', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Vojvodina', '\\(', 'Serbia', '\\)']]], 8, True)

['Romanian,', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Romanian,', 'Official', 'language', 'in', 'Mount', 'Athos']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Romanian,', 'Official', 'language', 'in', 'Latin', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Romanian,', 'Regulated', 'by', 'Academia', 'Rom', '\\xc3', '\\xa2', 'n', '\\xc4', '\\x83']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Academia', 'Rom', '\\xc3', '\\xa2', 'n', '\\xc4', '\\x83']]], 10, True)

['Romanian,', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY']]], 6, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-2', 'rum', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rum', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-2', 'ron', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ron', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-3', 'ron']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Romanian,', 'Language', 'codes', 'ISO', '639-1', 'ro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Swedish', 'Writing', 'system', 'Latin', '(', 'Swedish', 'variant', ')']:



([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Swedish', 'Pronunciation', '\xcb', '\x88', 'sv', '\xc9', '\x9b', '\xcc', '\x82', 'nsk', '\xc3', '\xa2']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'sv', '\xc9', '\x9b', '\xcc', '\x82', 'nsk', '\xc3', '\xa2']]], 12, True)

['Swedish', 'Official', 'language', 'in', 'Sweden']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Swedish', 'Official', 'language', 'in', 'Finland']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Swedish', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swedish', 'Official', 'language', 'in', 'Nordic', 'Council']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swedish', 'Regulated', 'by', 'Research', 'Institute', 'for', 'the', 'Languages', 'of', 'Finland', '(', 'in', 'Finland', ')']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Research', 'Institute', 'for', 'the', 'Languages', 'of', 'Finland', '\\(', 'in', 'Finland', '\\)']]], 14, True)

['Swedish', 'Language', 'codes', 'ISO', '639-2', 'swe']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Swedish', 'Language', 'codes', 'ISO', '639-3', 'swe']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Swedish', 'Language', 'codes', 'ISO', '639-1', 'sv']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Spanish', 'Castilian', 'Total', 'speakers', 'First', 'language', '329', 'million', 'to', '400', 'million']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Castilian', 'Total',

```
'speakers', 'First', 'language', 'NUM', 'million', 'to', 'NUM',  
'million']], ['NUM', '329'], ['NUM', '400']], 11, True)
```

```
['Spanish,', 'Total', 'speakers', '500', 'million', 'as', 'first', 'or',  
'second', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', 'as', 'first', 'or', 'second', 'language']], ['NUM',  
'500']], 10, True)
```

```
['Spanish,', 'Spoken', 'in', '(', 'see', 'below', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\(',  
'see', 'below', '\\)']], 7, True)
```

```
['Spanish,', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)
```

```
['Spanish,', 'Pronunciation', 'espa', '\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol',  
,',', 'kaste', '\\xcb', '\\x88', '\\xca', '\\x8e', 'ano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'espa',  
'\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol', ',,', 'kaste', '\\xcb', '\\x88', '\\xca',  
'\\x8e', 'ano']], 15, True)
```

```
['Spanish,', 'Official', 'language', 'in', '21', 'countries']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'United', 'Nations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'European', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'Organization', 'of', 'Ibero-  
American', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Organization', 'of', 'Ibero-American', 'States']], 8, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'African', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['Spanish,', 'Official', 'language', 'in', 'Caricom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
```

'in', 'Type: ANY']]], 5, True)

['Spanish,', 'Official', 'language', 'in', 'World', 'Trade',  
'Organization']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'World', 'Trade', 'Organization']]], 7, True)

['Spanish,', 'Official', 'language', 'in', 'Andean', 'Community', 'of',  
'Nations']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Andean', 'Community', 'of', 'Nations']]], 8, True)

['Spanish,', 'Official', 'language', 'in', 'Mercosur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Spanish,', 'Official', 'language', 'in', 'Inter-American', 'Development',  
'Bank']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Inter-American', 'Development', 'Bank']]], 7, True)

['Spanish,', 'Official', 'language', 'in', 'Latin', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Spanish,', 'Official', 'language', 'in', 'Antarctic', 'Treaty']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Spanish,', 'Language', 'codes', 'ISO', '639-2', 'spa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Spanish,', 'Language', 'codes', 'ISO', '639-3', 'spa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Spanish,', 'Language', 'codes', 'Linguasphere', '51-AAA-b']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Spanish,', 'Language', 'codes', 'ISO', '639-1', 'es']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Logudorese', 'Spoken', 'in', 'Italy']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Logudorese', 'Region', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Logudorese', 'Language', 'codes', 'ISO', '639-3', 'src']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sardinian', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sardinian', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sardinian', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sardinian', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sardinian', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sardinian', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sardinian', 'Region', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Sardinian', 'Official', 'language', 'in', 'Sardinia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)
```

```
['Sardinian', 'Language', 'codes', 'ISO', '639-2', 'srd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'srd', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'srd', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', ':', 'sro', '\xe2', '\x80', '\x93', 'Campidanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'sro', '\xe2', '\x80', '\x93', 'Campidanese']], ['NUM', '639-3']], 11, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'sdn', '\xe2', '\x80', '\x93', 'Gallurese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sdn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'src', '\xe2', '\x80', '\x93', 'Logudorese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'src', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-3', 'sdc', '\xe2', '\x80', '\x93', 'Sassarese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sdc', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Sardinian', 'Language', 'codes', 'ISO', '639-1', 'sc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Swiss', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Swiss', 'Spoken', 'in', 'Switzerland', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Swiss', 'Pronunciation', '\xcb', '\x88', '\xca', '\x83', '\xca', '\x8b', 'its', '\xc9', '\x99', 'r', '\xcb', '\x8c', 'd', '\xcc', '\xa5', 'yt', '\xca', '\x83']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', '\xca', '\x83', '\xca', '\x8b', 'its', '\xc9', '\x99', 'r', '\xcb', '\x8c', 'd', '\xcc', '\xa5', 'yt', '\xca', '\x83']]], 20, True)

['Swiss', 'Language', 'codes', 'ISO', '639-2', 'gsw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Swiss', 'Language', 'codes', 'ISO', '639-3', 'gsw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sanskrit', 'Spoken', 'in', 'Greater', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Sanskrit', 'Writing', 'system', 'Devan', '\xc4', '\x81', 'gar', '\xc4', '\xab', '(', 'de', 'facto', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devan', '\xc4', '\x81', 'gar', '\xc4', '\xab', '\\(', 'de', 'facto', '\\)']], 13, True)

['Sanskrit', 'Writing', 'system', 'various', 'Br', '\xc4', '\x81', 'hm', '\xc4', '\xab', '\xe2', '\x80', '\x93', 'based', 'scripts']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'various', 'Br', '\xc4', '\x81', 'hm', '\xc4', '\xab', '\xe2', '\x80', '\x93', 'based', 'scripts']]], 15, True)

['Sanskrit', 'Writing', 'system', 'and', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Sanskrit', 'Pronunciation', 's', '\xc9', '\x99', '\xcc', '\x83', 'sk', '\xc9', '\xb9', '\xcc', '\xa9', 't', '\xcc', '\xaa', '\xc9', '\x99', 'm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's', '\xc9', '\x99', '\xcc', '\x83', 'sk', '\xc9', '\xb9', '\xcc', '\xa9', 't', '\xcc', '\xaa', '\xc9', '\x99', 'm']]], 18, True)

['Sanskrit', 'Official', 'language', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Sanskrit', 'Official', 'language', 'in', 'one', 'of', 'the', '22',  
'scheduled', 'languages', 'of', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'one', 'of', 'the', 'NUM', 'scheduled', 'languages', 'of', 'India']],  
['NUM', '22']], 12, True)

['Sanskrit', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Sanskrit', 'Language', 'codes', 'ISO', '639-2', 'san']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Sanskrit', 'Language', 'codes', 'ISO', '639-3', 'san']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sanskrit', 'Language', 'codes', 'ISO', '639-1', 'sa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Serbo-Croatian', 'Dialects', '\xc4', '\x8c', 'akavian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xc4',  
'\x8c', 'akavian']]], 5, True)

['Serbo-Croatian', 'Dialects', 'Kajkavian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Serbo-Croatian', 'Dialects', '\xc5', '\xa0', 'tokavian', '(', 'standard',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xc5',  
'\xa0', 'tokavian', '\\(', 'standard', '\\)']]], 8, True)

['Serbo-Croatian', 'Dialects', 'Torlakian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Serbo-Croatian', 'Region', 'Serb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Serbo-Croatian', 'Region', 'Croat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

```

['Serbo-Croatian', 'Region', 'Bosniak']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Serbo-Croatian', 'Region', 'Montenegrin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Serbo-Croatian', 'Spoken', 'in', 'Serbia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Serbo-Croatian', 'Spoken', 'in', 'Croatia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Serbo-Croatian', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Serbo-Croatian', 'Spoken', 'in', 'Montenegro']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Serbo-Croatian', 'Spoken', 'in', 'Kosovoa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Serbo-Croatian', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Serbo-Croatian', 'Writing', 'system', 'Cyrillic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Serbo-Croatian', 'Standard', 'forms', 'Serbian']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Serbo-Croatian', 'Standard', 'forms', 'Croatian']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Serbo-Croatian', 'Standard', 'forms', 'Bosnian']:

```



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Standard', 'forms', 'Montenegrin']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-2', 'scr', ',', 'scc', '(', 'deprecated', ')']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'scr', ',', 'scc', '\\(', 'deprecated', '\\)']]], ['NUM', '639-2']], 11, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'hbs', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hbs', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', ':', 'srp', '\\xe2', '\\x80', '\\x93', 'Serbian']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'srp', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM', '639-3']], 11, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'hrv', '\\xe2', '\\x80', '\\x93', 'Croatian']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hrv', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-3', 'bos', '\\xe2', '\\x80', '\\x93', 'Bosnian']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bos', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'Linguasphere', '53-AAA-g']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Serbo-Croatian', 'Language', 'codes', 'ISO', '639-1', 'sh', '(',  
'deprecated', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sh', '\\\\(', 'deprecated', '\\\\)']], ['NUM', '639-1']], 9,  
True)
```

```
['Slovene/Slovenian', 'Dialects', 'Prekmurje', 'dialect']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Slovene/Slovenian', 'Dialects', 'Resian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Slovene/Slovenian', 'Dialects', 'approx', '.', '32', 'unstandardised',  
'dialects']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'approx',  
'.', 'NUM', 'unstandardised', 'dialects']], ['NUM', '32']], 7, True)
```

```
['Slovene/Slovenian', 'Spoken', 'in', 'Slovenia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Slovene/Slovenian', 'Spoken', 'in', 'Italy', '(', 'in', 'Friuli',  
'Venezia', 'Giulia', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy',  
'\\\\(', 'in', 'Friuli', 'Venezia', 'Giulia', '\\\\)']], 10, True)
```

```
['Slovene/Slovenian', 'Spoken', 'in', 'Austria', '(', 'in', 'Carinthia',  
'and', 'Styria', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Austria', '\\\\(', 'in', 'Carinthia', 'and', 'Styria', '\\\\)']], 10, True)
```

```
['Slovene/Slovenian', 'Spoken', 'in', 'Hungary', '(', 'in', 'Vas', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Hungary', '\\\\(', 'in', 'Vas', '\\\\)']], 8, True)
```

```
['Slovene/Slovenian', 'Spoken', 'in', 'Croatia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Slovene/Slovenian', 'Spoken', 'in', 'and', 'emigrant', 'groups', 'in',  
'various', 'countries']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',  
'emigrant', 'groups', 'in', 'various', 'countries']]], 9, True)
```

['Slovene/Slovenian', 'Pronunciation', 'slo', '\xcb', '\x88', '\xca', '\x8b', 'enski', '\xcb', '\x88', 'j', '\xc9', '\x9b', 'zik', ',', 'slo', '\xcb', '\x88', '\xca', '\x8b', 'en', '\xca', '\x83', 't', '\xca', '\x83', 'ina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'slo', '\xcb', '\x88', '\xca', '\x8b', 'enski', '\xcb', '\x88', 'j', '\xc9', '\x9b', 'zik', ',', 'slo', '\xcb', '\x88', '\xca', '\x8b', 'en', '\xca', '\x83', 't', '\xca', '\x83', 'ina']]], 27, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'Slovenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Slovene/Slovenian', 'Official', 'language', 'in', ':', 'Austria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'Hungary']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Slovene/Slovenian', 'Official', 'language', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Slovene/Slovenian', 'Regulated', 'by', 'Slovenian', 'Academy', 'of', 'Sciences', 'and', 'Arts']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Slovenian', 'Academy', 'of', 'Sciences', 'and', 'Arts']]], 9, True)

['Slovene/Slovenian', 'Language', 'codes', 'ISO', '639-2', 'slv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Slovene/Slovenian', 'Language', 'codes', 'ISO', '639-3', 'slv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Slovene/Slovenian', 'Language', 'codes', 'Linguasphere', 'Slovene-speaking', 'areas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Slovene-speaking', 'areas']]], 6, True)

['Slovene/Slovenian', 'Language', 'codes', 'ISO', '639-1', 'sl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Slovak', 'Total', 'speakers', 'over', '7', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'over', 'NUM', 'million']], ['NUM', '7']], 6, True)

['Slovak', 'Region', 'Central', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Slovak', 'Spoken', 'in', 'Slovakia', 'and', 'as', 'a', 'minority',  
'language', 'also', 'in', 'the']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Slovakia', 'and', 'as', 'a', 'minority', 'language', 'also', 'in',  
'the']], 12, True)

['Slovak', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Slovak', 'Spoken', 'in', 'Canada', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', '']], 5, True)

['Slovak', 'Spoken', 'in', 'Serbia', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', '']], 5, True)

['Slovak', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Slovak', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Slovak', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Slovak', 'Official', 'language', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Slovak', 'Official', 'language', 'in', 'Vojvodina', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Vojvodina', 'in', 'Serbia']]], 7, True)

['Slovak', 'Recognised', 'minority', 'language', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Slovak', 'Language', 'codes', 'ISO', '639-2', 'slo', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'slo', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Slovak', 'Language', 'codes', 'ISO', '639-2', 'slk', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'slk', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Slovak', 'Language', 'codes', 'ISO', '639-3', 'slk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Slovak', 'Language', 'codes', 'ISO', '639-1', 'sk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Scottish', 'Gaelic', 'Total', 'speakers', '58', ',', '552', 'in',  
'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Gaelic', 'Total',  
'speakers', 'NUM', ',', 'NUM', 'in', 'Scotland']], ['NUM', '58'], ['NUM',  
'552']], 9, True)

['Scottish', 'Region', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Scottish', 'Region', 'Cape', 'Breton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Scottish', 'Region', 'Nova', 'Scotia', 'and', 'Glengarry', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nova',  
'Scotia', 'and', 'Glengarry', 'County']]], 7, True)

['Scottish', 'Region', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Scottish', 'Spoken', 'in', 'United', 'Kingdom']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY' ] ]], 5, True)

['Scottish', 'Spoken', 'in', 'Canada']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 4, True)

['Scottish', 'Spoken', 'in', 'United', 'States']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY' ] ]], 5, True)

['Scottish', 'Spoken', 'in', 'Australia']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 4, True)

['Scottish', 'Spoken', 'in', 'New', 'Zealand']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY' ] ]], 5, True)

['Scottish', 'Writing', 'system', 'Latin', '(', 'Gaelic', 'variant', ')']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Writing', 'system', 'Type: ANY', '\\(', 'Gaelic', 'variant', '\\)' ] ]], 8, True)

['Scottish', 'Pronunciation', '\\xcb', '\\x88', 'ka', '\\xcb', '\\x90', 'lik', '\\xca', '\\xb2']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Pronunciation', '\\xcb', '\\x88', 'ka', '\\xcb', '\\x90', 'lik', '\\xca', '\\xb2' ] ]], 10, True)

['Scottish', 'Official', 'language', 'in', 'Scotland']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Official', 'language', 'in', 'Type: ANY' ] ]], 5, True)

['Scottish', 'Regulated', 'by']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ]], 3, True)

['Scottish', 'Language', 'codes', 'ISO', '639-2', 'gla']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-2' ]], 6, True)

['Scottish', 'Language', 'codes', 'ISO', '639-3', 'gla']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ]], 6, True)

```
['Scottish', 'Language', 'codes', 'Linguasphere', '50-AAA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Scottish', 'Language', 'codes', 'ISO', '639-1', 'gd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Swahili', 'Spoken', 'in', 'Burundi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'DR', 'Congo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Swahili', 'Spoken', 'in', 'Kenya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Mozambique']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Oman']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Mauritius']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Rwanda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Uganda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Swahili', 'Spoken', 'in', 'Malawi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Swahili', 'Total', 'speakers', 'First', 'language', ':', '5', '\xe2',  
\x80', \x93', '10', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'First', 'language', ':', 'NUM', '\xe2', '\x80', '\x93', 'NUM',  
'million']], ['NUM', '5'], ['NUM', '10']], 12, True)

['Swahili', 'Writing', 'system', 'Latin', ',', 'Arabic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swahili', 'Official', 'language', 'in', 'African', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swahili', 'Official', 'language', 'in', 'Kenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Swahili', 'Official', 'language', 'in', 'Tanzania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Swahili', 'Official', 'language', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Swahili', 'Regulated', 'by', 'Baraza', 'la', 'Kiswahili', 'la', 'Taifa',  
'(', 'Tanzania', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Baraza', 'la', 'Kiswahili', 'la', 'Taifa', '\\(', 'Tanzania', '\\)']]],  
11, True)

['Swahili', 'Language', 'codes', 'ISO', '639-2', 'swa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Swahili', 'Language', 'codes', 'ISO', '639-3', 'swa', '\xe2', '\x80',  
\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'swa', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Swahili', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Swahili', 'Language', 'codes', 'ISO', '639-3', ':', 'swc', '\xe2', '\x80', '\x93', 'Congo', 'Swahili']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'swc', '\xe2', '\x80', '\x93', 'Congo', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Swahili', 'Language', 'codes', 'ISO', '639-3', 'swh', '\xe2', '\x80', '\x93', 'Coastal', 'Swahili']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'swh', '\xe2', '\x80', '\x93', 'Coastal', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Swahili', 'Language', 'codes', 'Linguasphere', '99-AUS-m']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Swahili', 'Language', 'codes', 'ISO', '639-1', 'sw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tamil', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tamil', 'Spoken', 'in', 'Sri', 'Lanka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tamil', 'Spoken', 'in', 'Singapore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tamil', 'Spoken', 'in', 'where', 'it', 'has', 'official', 'status', ';', 'with', 'significant', 'minorities', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'where', 'it', 'has', 'official', 'status', ';', 'with', 'significant', 'minorities', 'in', 'Canada']], 14, True)

['Tamil', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tamil', 'Spoken', 'in', 'United', 'Kingdom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Tamil', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tamil', 'Spoken', 'in', 'Mauritius', 'and', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Tamil', 'Spoken', 'in', 'and', 'emigrant', 'communities', 'around', 'the', 'world']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'emigrant', 'communities', 'around', 'the', 'world']]], 9, True)

['Tamil', 'Writing', 'system', 'Tamil', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tamil', 'Pronunciation', 't', '\xcc', '\xaa', '\xc9', '\x90', 'm', '\xc9', '\xa8', '\xc9', '\xbb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't', '\xcc', '\xaa', '\xc9', '\x90', 'm', '\xc9', '\xa8', '\xc9', '\xbb']]], 12, True)

['Tamil', 'Official', 'language', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Tamil', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tamil', 'Language', 'codes', 'ISO', '639-2', 'tam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tamil', 'Language', 'codes', 'ISO', '639-3', 'tam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tamil', 'Language', 'codes', 'ISO', '639-1', 'ta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Setswana', 'Spoken', 'in', 'Botswana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Setswana', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Setswana', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Setswana', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Setswana', 'Region', 'Southern', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Setswana', 'Official', 'language', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Setswana', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Setswana', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Setswana', 'Language', 'codes', 'ISO', '639-2', 'tsn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Setswana', 'Language', 'codes', 'ISO', '639-3', 'tsn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Setswana', 'Language', 'codes', 'ISO', '639-1', 'tn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tocharian', 'Writing', 'system', 'Tocharian', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tocharian', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vietnamese', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Vietnamese', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Vietnamese', 'Spoken', 'in', 'Vietnamese', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Vietnamese', 'Writing', 'system', 'Vietnamese', 'variant', 'of', 'Latin',  
'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Vietnamese', 'Official', 'language', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Vietnamese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-2', 'vie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-3', 'vie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vietnamese', 'Language', 'codes', 'Linguasphere', '46-EBA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Vietnamese', 'Language', 'codes', 'Linguasphere', 'Extent', 'of',  
'Vietnamese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Extent', 'of', 'Type: ANY']], 7, True)

```

['Vietnamese', 'Language', 'codes', 'ISO', '639-1', 'vi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Volap\xc3\xbck', 'Total', 'speakers', '20']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Volap\xc3\xbck', 'Writing', 'system', 'German']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Volap\xc3\xbck', 'Writing', 'system', 'and', 'French']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Volap\xc3\xbck', 'Category', '(', 'purpose', ')', 'constructed',
'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Volap\xc3\xbck', 'Category', '(', 'purpose', ')', 'Volap', '\xc3',
'\xbc', 'k']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Volap', '\xc3', '\xbc', 'k']]], 9, True)

['Volap\xc3\xbck', 'Created', 'by', 'Johann', 'Martin', 'Schleyer']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',
'Johann', 'Martin', 'Schleyer']]], 6, True)

['Volap\xc3\xbck', 'Language', 'codes', 'ISO', '639-2', 'vol']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Volap\xc3\xbck', 'Language', 'codes', 'ISO', '639-3', 'vol']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Volap\xc3\xbck', 'Language', 'codes', 'ISO', '639-1', 'vo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Yiddish', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

```

['Yiddish', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Yiddish', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Belarus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yiddish', 'Spoken', 'in', 'Moldova']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Lithuania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Belgium']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yiddish', 'Spoken', 'in', 'France', 'and', 'elsewhere']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Yiddish', 'Writing', 'system', 'uses', 'a', 'Hebrew-based', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'uses', 'a', 'Hebrew-based', 'alphabet']]], 7, True)

['Yiddish', 'Pronunciation', '\xcb', '\x88', 'j', '\xc9', '\xaa', 'd', '\xc9', '\xaa', '\xca', '\x83', ',', '\xcb', '\x88', 'jid', '\xc9', '\xaa', '\xca', '\x83']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'j', '\xc9', '\xaa', 'd', '\xc9', '\xaa', '\xca', '\x83', ',', '\xcb', '\x88', 'jid', '\xc9', '\xaa', '\xca', '\x83']]], 20, True)

['Yiddish', 'Official', 'language', 'in', 'Sweden']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Yiddish', 'Official', 'language', 'in', 'Recognized', 'as', 'a', 'minority', 'language', 'in', 'the', 'Russian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Recognized', 'as', 'a', 'minority', 'language', 'in', 'the', 'Type: ANY']]], 12, True)

['Yiddish', 'Regulated', 'by', 'no', 'formal', 'bodies', ';']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'no',

'formal', 'bodies', ';']]], 7, True)

['Yiddish', 'Regulated', 'by', 'YIVO', 'de', 'facto']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'YIVO', 'de', 'facto']]], 6, True)

['Yiddish', 'Language', 'codes', 'ISO', '639-2', 'yid']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Yiddish', 'Language', 'codes', 'ISO', '639-3', 'yid', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'yid', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Yiddish', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Yiddish', 'Language', 'codes', 'ISO', '639-1', 'yi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Indonesian', 'Total', 'speakers', '23', 'million', 'native', ';',  
'approx', '.', '165', 'million', 'total']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', 'native', ';', 'approx', '.', 'NUM', 'million',  
'total']], ['NUM', '23'], ['NUM', '165']], 12, True)

['Indonesian', 'Region', 'Southeast', 'Asia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Indonesian', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Indonesian', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Indonesian', 'Official', 'language', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)



```
['Indonesian', 'Regulated', 'by', 'Pusat', 'Bahasa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Indonesian', 'Language', 'codes', 'ISO', '639-2', 'ind']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Indonesian', 'Language', 'codes', 'ISO', '639-3', 'ind']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Indonesian', 'Language', 'codes', 'ISO', '639-1', 'id']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Malayalam', 'Total', 'speakers', ':', '\xe2', '\x80', '\xa2', '773', ',',  
'624', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', ':',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'UAE']]], ['NUM', '773'],  
['NUM', '624']], 12, True)
```

```
['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '134', ',',  
'728', 'in', 'Kuwait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'134'], ['NUM', '728']], 11, True)
```

```
['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '134', ',',  
'019', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'134'], ['NUM', '019']], 11, True)
```

```
['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '105', ',',  
'655', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'105'], ['NUM', '655']], 11, True)
```

```
['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '94', ',',  
'310', 'in', 'Qatar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM',  
'94'], ['NUM', '310']], 11, True)
```

```

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '58', ',', '146', 'in', 'Bahrain']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '58'], ['NUM', '146']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '26', ',', '237', 'in', 'UK']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '26'], ['NUM', '237']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '15', ',', '600', 'in', 'other', 'Europe']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'other', 'Europe']], ['NUM', '15'], ['NUM', '600']], 12, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '11', ',', '346', 'in', 'Canada']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '11'], ['NUM', '346']], 11, True)

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '10', ',', '636', 'in', 'Malaysia']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\xa2', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '10'], ['NUM', '636']], 11, True)

['Malayalam', 'Spoken', 'in', 'India']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Malayalam', 'Region', 'Kerala']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Malayalam', 'Region', 'Lakshadweep']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Malayalam', 'Region', 'Karnataka']:
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

```

```
['Malayalam', 'Region', 'Tamil', 'Nadu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Malayalam', 'Region', 'Mah', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Malayalam', 'Region', 'Andaman', 'and', 'Nicobar', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Andaman',  
'and', 'Nicobar', 'Islands']]], 6, True)
```

```
['Malayalam', 'Official', 'language', 'in', 'India', '(', 'Kerala', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'India', '\\(', 'Type: ANY', '\\)']]], 8, True)
```

```
['Malayalam', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Malayalam', 'Language', 'codes', 'ISO', '639-2', 'mal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Malayalam', 'Language', 'codes', 'ISO', '639-3', 'mal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Malayalam', 'Language', 'codes', 'ISO', '639-1', 'ml']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Language', 'Spoken', 'in', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Language', 'Spoken', 'in', 'Kerala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Language', 'Spoken', 'in', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Language', 'Spoken', 'in', 'Andhra', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'Type: ANY']]], 5, True)

['Language', 'Spoken', 'in', 'Goa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Language', 'Spoken', 'in', 'Tamil', 'Nadu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Language', 'Spoken', 'in', 'USA', 'and', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Language', 'Official', 'language', 'in', 'India', '(', 'Karnataka', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Language', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Karnataka']]], 10, True)

['Language', 'Language', 'codes', 'ISO', '639-2', 'kan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Language', 'Language', 'codes', 'ISO', '639-3', 'kan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Language', 'Language', 'codes', 'ISO', '639-1', 'kn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Telugu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Telugu', 'Writing', 'system', 'Telugu', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Telugu', 'Official', 'language', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['Telugu', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Telugu', 'Language', 'codes', 'ISO', '639-2', 'tel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Telugu', 'Language', 'codes', 'ISO', '639-3', 'tel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Telugu', 'Language', 'codes', 'ISO', '639-1', 'te']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Dari', 'Dialects', 'Kaboli']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dari', 'Dialects', 'Mazari']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dari', 'Dialects', 'Herati']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dari', 'Dialects', 'Badakhshi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dari', 'Dialects', 'Panjshiri']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dari', 'Dialects', 'Laghmani']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dari', 'Dialects', 'Sistani']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

```

['Dari', 'Dialects', 'Aimaqi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Dialects', 'Hazaragi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dari', 'Region', 'Central', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Region', 'Middle', 'East']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Region', 'South', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dari', 'Official', 'language', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Dari', 'Regulated', 'by', 'Academy', 'of', 'Sciences', 'of', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Academy', 'of', 'Sciences', 'of', 'Afghanistan']]], 8, True)

['Dari', 'Language', 'codes', 'ISO', '639-2', 'per', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'per', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Dari', 'Language', 'codes', 'ISO', '639-2', 'fas', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fas', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Dari', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Dari', 'Language', 'codes', 'ISO', '639-3', 'aiq', '\\xe2', '\\x80',

```

```

'\x93', 'Aimaq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'aiq', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Dari', 'Language', 'codes', 'ISO', '639-3', 'haz', '\xe2', '\x80',
'\x93', 'Hazaragi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'haz', '\xe2', '\x80', '\x93', 'Hazaragi']], ['NUM',
'639-3']], 10, True)

['Dari', 'Language', 'codes', 'ISO', '639-1', 'fa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Kurdish', 'Total', 'speakers', '16', 'million', 'to', '35', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'to', 'NUM', 'million']], ['NUM', '16'], ['NUM', '35']],
8, True)

['Kurdish', 'Region', 'West', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kurdish', 'Spoken', 'in', 'Turkey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kurdish', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kurdish', 'Spoken', 'in', 'Iraq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kurdish', 'Spoken', 'in', 'Syria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kurdish', 'Spoken', 'in', 'Lebanon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kurdish', 'Spoken', 'in', 'Georgia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Turkmenistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurdish', 'Spoken', 'in', '(', 'see', 'article', 'for', 'full', 'list', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\(', 'see', 'article', 'for', 'full', 'list', '\\)']]], 10, True)

['Kurdish', 'Official', 'language', 'in', 'Iran', ':', 'constitutional', 'status', 'as', 'a', 'regional', 'language', ',', 'but', 'never', 'implemented']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Iran', ':', 'constitutional', 'status', 'as', 'a', 'regional', 'language', ',', 'but', 'never', 'implemented']]], 16, True)

['Kurdish', 'Official', 'language', 'in', 'Armenia', ':', 'minority', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Armenia', ':', 'minority', 'language']]], 8, True)

['Kurdish', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-2', 'kur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'kur', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'kur', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', ':', 'ckb', '\xe2', '\x80', '\x93', 'Sorani']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ckb', '\xe2', '\x80', '\x93', 'Sorani']], ['NUM', '639-3']], 11, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'kmr', '\xe2', '\x80', '\x93', 'Kurmanji']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kmr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'lki', '\xe2', '\x80', '\x93', 'Laki']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lki', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Kurdish', 'Language', 'codes', 'ISO', '639-1', 'ku']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sudovian', 'Spoken', 'in', 'Yotvingia', 'and', 'S', '\xc5', '\xab', 'duva']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Yotvingia', 'and', 'S', '\xc5', '\xab', 'duva']], 9, True)

['Sudovian', 'Region', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sudovian', 'Language', 'codes', 'ISO', '639-3', 'xsv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Region', 'Eastern', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Writing', 'system', 'Glagolitic', 'alphabet', ';', 'Cyrillic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'chu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'chu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Language', 'codes', 'Linguasphere', '53-AAA-a']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-1', 'cu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Latvian', 'Total', 'speakers', '110', ',', '000', '(', 'Abroad', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'Abroad', '\\\\)']], ['NUM', '110'], ['NUM', '000']], 9, True)

['Latvian', 'Region', 'Baltic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Latvian', 'Spoken', 'in', 'Latvia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Latvian', 'Writing', 'system', 'Latvian', 'variant', 'of', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Latvian', 'Official', 'language', 'in', 'Latvia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Latvian', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Latvian', 'Language', 'codes', 'ISO', '639-2', 'lav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Latvian', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Latvian', 'Language', 'codes', 'ISO', '639-3', ':', 'lav', '\xe2',  
'\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'lav', '\xe2', '\x80', '\x93', 'Macrolanguage']],  
['NUM', '639-3']], 11, True)

['Latvian', 'Language', 'codes', 'Linguasphere', '54-AAB-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Latvian', 'Language', 'codes', 'ISO', '639-1', 'lv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Mon', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Mon', 'Pronunciation', 'p', '\xca', '\xb0', 'esa', 'm', '\xc9', '\x91',  
'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p',  
'\xca', '\xb0', 'esa', 'm', '\xc9', '\x91', 'n']], 10, True)

['Mon', 'Language', 'codes', 'ISO', '639-3', 'mnw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Michif', 'Total', 'speakers', '500', '\xe2', '\x80', '\x93', '1', ',',  
'000', 'endangered']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\xe2', '\x80', '\x93', 'NUM', ',', 'NUM', 'endangered']], ['NUM',  
'500'], ['NUM', '1'], ['NUM', '000']], 11, True)

['Michif', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Michif', 'Region', 'M', '\xc3', '\xa9', 'tis', 'communities', 'in', 'the', 'Prairies', ';', 'mostly', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'M', '\xc3', '\xa9', 'tis', 'communities', 'in', 'the', 'Prairies', ';', 'mostly', 'Manitoba']]], 13, True)

['Michif', 'Region', 'Saskatchewan', 'and', 'northwestern', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Saskatchewan', 'and', 'northwestern', 'Ontario']]], 6, True)

['Michif', 'Official', 'language', 'in', 'M', '\xc3', '\xa9', 'tis', 'Nation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'M', '\xc3', '\xa9', 'tis', 'Nation']]], 9, True)

['Michif', 'Regulated', 'by', 'M', '\xc3', '\xa9', 'tis', 'National', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'M', '\xc3', '\xa9', 'tis', 'National', 'Council']]], 9, True)

['Michif', 'Language', 'codes', 'ISO', '639-3', 'crg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['British', 'Signed', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'United', 'Kingdom']]], 5, True)

['British', 'Total', 'signers', 'About', '125', ',', '000', 'first-language', 'signers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'About', 'NUM', ',', 'NUM', 'first-language', 'signers']], ['NUM', '125'], ['NUM', '000']], 9, True)

['Coptic', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coptic', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coptic', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Coptic', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Coptic', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Coptic', 'Writing', 'system', 'Coptic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Coptic', 'Language', 'codes', 'ISO', '639-2', 'cop']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Coptic', 'Language', 'codes', 'ISO', '639-3', 'cop']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Thai', 'Total', 'speakers', 'over', '60', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'over', 'NUM', 'million']], ['NUM', '60']], 6, True)
```

```
['Thai', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Thai', 'Spoken', 'in', 'Northern', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Thai', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Thai', 'Spoken', 'in', 'Southern', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Thai', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Thai', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Thai', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Thai', 'Spoken', 'in', 'France']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Thai', 'Spoken', 'in', 'England']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Thai', 'Writing', 'system', 'Thai', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Thai', 'Pronunciation', 'p', '\xca', '\xb0', '\xc4', '\x81', '\xcb', '\x90', 's', '\xc7', '\x8e', '\xcb', '\x90', 't', '\xca', '\xb0', '\xc4', '\x81', 'j']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\xca', '\xb0', '\xc4', '\x81', '\xcb', '\x90', 's', '\xc7', '\x8e', '\xcb', '\x90', 't', '\xca', '\xb0', '\xc4', '\x81', 'j']]], 20, True)

['Thai', 'Official', 'language', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Thai', 'Regulated', 'by', 'The', 'Royal', 'Institute']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'The', 'Royal', 'Institute']]], 6, True)

['Thai', 'Language', 'codes', 'ISO', '639-2', 'tha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Thai', 'Language', 'codes', 'ISO', '639-3', 'tha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thai', 'Language', 'codes', 'Linguasphere', '47-AAA-b']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Thai', 'Language', 'codes', 'ISO', '639-1', 'th']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Polari', 'Spoken', 'in', 'United', 'Kingdom', 'and', 'Republic', 'of', 'Ireland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'Kingdom', 'and', 'Republic', 'of', 'Ireland']]], 9, True)

['Polari', 'Total', 'speakers', 'no', 'estimate', 'available']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'no', 'estimate', 'available']]], 6, True)

['Polari', 'Language', 'codes', 'ISO', '639-3', 'pld']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Duala', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Duala', 'Language', 'codes', 'ISO', '639-3', 'dua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Akkadian', 'Spoken', 'in', 'Assyria', 'and', 'Babylonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Akkadian', 'Region', 'Mesopotamia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Akkadian', 'Writing', 'system', 'Sumero-Akkadian', 'cuneiform']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Akkadian', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Akkadian', 'Language', 'codes', 'ISO', '639-2', 'akk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Akkadian', 'Language', 'codes', 'ISO', '639-3', 'akk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Middle', 'Spoken', 'in', 'to', 'some', 'extent', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'to', 'some', 'extent', 'in', 'Ireland']], 8, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'enm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'enm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['M\\xc4\\x81ori', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['M\\xc4\\x81ori', 'Region', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['M\\xc4\\x81ori', 'Official', 'language', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-2', 'mao', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mao', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-2', 'mri', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mri', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-3', 'mri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['M\\xc4\\x81ori', 'Language', 'codes', 'ISO', '639-1', 'mi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Cree', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cree', 'Official', 'language', 'in', 'Northwest', 'Territories', '(', 'Canada', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Northwest', 'Territories', '\\(', 'Canada', '\\)']], 9, True)

['Cree', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Cree', 'Language', 'codes', 'ISO', '639-2', 'cre']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'cre', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cre', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', ':', 'nsk', '\\xe2', '\\x80', '\\x93', 'Naskapi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'nsk', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'moe', '\\xe2', '\\x80', '\\x93', 'Montagnais']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'moe', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Cree', 'Language', 'codes', 'ISO', '639-3', 'atj', '\\xe2', '\\x80', '\\x93', 'Atikamekw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'atj', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Cree', 'Language', 'codes', 'ISO', '639-1', 'cr']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tetum', 'Spoken', 'in', 'East', 'Timor']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tetum', 'Spoken', 'in', 'Indonesia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetum', 'Spoken', 'in', 'Australia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetum', 'Spoken', 'in', 'Portugal']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetum', 'Region', 'Southeast', 'Asia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetum', 'Official', 'language', 'in', 'East', 'Timor']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Tetum', 'Regulated', 'by', 'National', 'Institute', 'of', 'Linguistics']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'National', 'Institute', 'of', 'Linguistics']], 7, True)

['Tetum', 'Language', 'codes', 'ISO', '639-2', 'tet']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tetum', 'Language', 'codes', 'ISO', '639-3', 'tet']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hakka', 'Total', 'speakers', '34', 'million']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '34']], 5, True)

['Hakka', 'Spoken', 'in', 'Mainland', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hakka', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Mauritius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hakka', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hakka', 'Writing', 'system', 'hanzi', ',', 'romanization']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Hakka', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Hakka', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Hakka', 'Language', 'codes', 'ISO', '639-3', 'hak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyabwa', 'Spoken', 'in', 'Ivory', 'Coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Nyabwa', 'Language', 'codes', 'ISO', '639-3', 'nwb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quebec', 'Spoken', 'in', 'Quebec', '(', 'mainly', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Quebec', 'Spoken', 'in', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quebec', 'Spoken', 'in', 'Western', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Quebec', 'Spoken', 'in', 'New', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Quebec', 'Regulated', 'by', 'Office', 'qu', '\\xc3', '\\xa9', 'b', '\\xc3',  
'\\xa9', 'cois', 'de', 'la', 'langue', 'fran', '\\xc3', '\\xa7', 'aise']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Office', 'qu', '\\xc3', '\\xa9', 'b', '\\xc3', '\\xa9', 'cois', 'de', 'la',  
'langue', 'fran', '\\xc3', '\\xa7', 'aise']], 18, True)

['Quebec', 'Language', 'codes', 'ISO', '639-2', 'fre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Quebec', 'Language', 'codes', 'ISO', '639-3', 'fra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quebec', 'Language', 'codes', 'ISO', '639-1', 'fr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dida', 'Spoken', 'in', 'Ivory', 'Coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Dida', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dida', 'Language', 'codes', 'ISO', '639-3', ':', 'gud', '\xe2', '\x80',  
'\x93', 'Yocobou', '\xc3', '\xa9', 'Dida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'gud', '\xe2', '\x80', '\x93', 'Yocobou', '\xc3',  
'\xa9', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Taiwanese', 'Region', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Taiwanese', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Taiwanese', 'Official', 'language', 'in', 'None', ',', 'de', 'facto',  
'status', 'in', 'Taiwan', 'since', 'it', 'is', 'one', 'of', 'the',  
'statutory', 'languages', 'for', 'public', 'transport', 'announcements',  
'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'None', ',', 'de', 'facto', 'status', 'in', 'Taiwan', 'since', 'it',  
'is', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public',  
'transport', 'announcements', 'in', 'Taiwan']]], 25, True)

['Taiwanese', 'Regulated', 'by', 'National', 'Languages', 'Committee', '(',  
'Ministry', 'of', 'Education', ',', 'ROC', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'National', 'Languages', 'Committee', '\\\\(', 'Ministry', 'of', 'Education',  
,',', 'ROC', '\\\)']]], 13, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\)']]], ['NUM', '639-2']], 9, True)

```
['Taiwanese', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Taiwanese', 'Language', 'codes', 'ISO', '639-3', 'nan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taiwanese', 'Language', 'codes', 'ISO', '639-1', 'zh']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Syriac', 'Spoken', 'in', 'Mesopotamia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Syriac', 'Spoken', 'in', 'Assyria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Syriac', 'Spoken', 'in', 'Roman', 'Syria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Syriac', 'Writing', 'system', 'Syriac', 'abjad']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Syriac', 'Pronunciation', '/', 'l', '\\xc9', '\\x9b', '\\xca', '\\x83',
'\\xcb', '\\x90', '\\xc9', '\\x91', 'n', '\\xc9', '\\x91', '\\xca', '\\x94',
'surj', '\\xc9', '\\x91', 'j', '\\xc9', '\\x91', '\\xca', '\\x94', '/']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',
'l', '\\xc9', '\\x9b', '\\xca', '\\x83', '\\xcb', '\\x90', '\\xc9', '\\x91', 'n',
'\\xc9', '\\x91', '\\xca', '\\x94', 'surj', '\\xc9', '\\x91', 'j', '\\xc9',
'\\x91', '\\xca', '\\x94', '/']], 26, True)
```

```
['Syriac', 'Language', 'codes', 'ISO', '639-2', 'syc', '(', 'classical',
')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'syc', '\\(', 'classical', '\\)']], ['NUM', '639-2']], 9,
True)
```

```
['Balochi', 'Spoken', 'in', 'Pakistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Iran']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Afghanistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Turkmenistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'UAE']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Spoken', 'in', 'Oman']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balochi', 'Language', 'codes', 'ISO', '639-2', 'bal']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bal', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'bal', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Balochi', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Zulu', 'Spoken', 'in', 'South', 'Africa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Zulu', 'Spoken', 'in', 'Zimbabwe']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Lesotho']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Spoken', 'in', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Region', 'KwaZulu-Natal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Zulu', 'Region', 'eastern', 'Gauteng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Region', 'eastern', 'Free', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Zulu', 'Region', 'southern', 'Mpumalanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zulu', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Zulu', 'Language', 'codes', 'ISO', '639-2', 'zul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Zulu', 'Language', 'codes', 'ISO', '639-3', 'zul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zulu', 'Language', 'codes', 'ISO', '639-1', 'zu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Nepali', 'Total', 'speakers', 'native', '\xe2', '\x80', '\x93', '14',  
'million']:
```



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'native', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '14']], 9,  
True)
```

```
['Nepali', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Nepali', 'Spoken', 'in', 'Myanmar', '(', 'Burma', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)
```

```
['Nepali', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Nepali', 'Official', 'language', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['Nepali', 'Official', 'language', 'in', 'Sikkim', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', '\\(', 'India', '\\)']], 8, True)
```

```
['Nepali', 'Language', 'codes', 'ISO', '639-2', 'nep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Nepali', 'Language', 'codes', 'ISO', '639-3', 'nep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nepali', 'Language', 'codes', 'ISO', '639-1', 'ne']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Scots', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Scots', 'Spoken', 'in', 'Northern', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Scots', 'Spoken', 'in', 'Republic', 'of', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Republic', 'of', 'Ireland']], 6, True)

['Scots', 'Region', 'Northern', 'Isles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Scots', 'Region', 'Caithness']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Scots', 'Region', 'Ulster', ':', 'Counties', 'Down']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ulster', ':', 'Counties', 'Down']], 6, True)

['Scots', 'Region', 'Antrim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Scots', 'Region', 'Londonderry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Scots', 'Region', 'Donegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Scots', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Scots', 'Regulated', 'by', '\xe2', '\x80', '\x94', 'Ireland', ':', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', '\xe2', '\x80', '\x94', 'Ireland', ':', 'None']], 9, True)

['Scots', 'Regulated', 'by', 'although', 'the', 'cross-border', 'Ulster-Scots', 'Agency']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'although', 'the', 'cross-border', 'Ulster-Scots', 'Agency']]], 8, True)

['Scots', 'Regulated', 'by', 'established', 'by', 'the', 'Implementation', 'Agreement', 'following', 'the', 'Good', 'Friday', 'Agreement', 'promotes', 'usage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'established', 'by', 'the', 'Implementation', 'Agreement', 'following', 'the', 'Good', 'Friday', 'Agreement', 'promotes', 'usage']]], 15, True)

['Scots', 'Language', 'codes', 'ISO', '639-2', 'sco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Scots', 'Language', 'codes', 'ISO', '639-3', 'sco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sumerian', 'Spoken', 'in', 'Sumer', 'and', 'Akkad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Sumerian', 'Region', 'Iraq', '(', 'Mesopotamia', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Iraq', '\\\\(', 'Mesopotamia', '\\\\)']]], 6, True)

['Sumerian', 'Language', 'codes', 'ISO', '639-2', 'sux']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sumerian', 'Language', 'codes', 'ISO', '639-3', 'sux']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hindi-Urdu', 'Dialects', 'Dakhni']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hindi-Urdu', 'Region', 'South', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hindi-Urdu', 'Writing', 'system', 'Devanagari', 'script', ',']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Hindi-Urdu', 'Writing', 'system', 'Perso-Arabic', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-2', 'hin', ',', 'ur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hi', ',', 'ur']], ['NUM', '639-2']], 8, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-3', ':', 'hin', '\xe2', '\x80', '\x93', 'Hindi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'hin', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-3', 'urd', '\xe2', '\x80', '\x93', 'Urdu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'urd', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Hindi-Urdu', 'Language', 'codes', 'ISO', '639-1', 'hi', ',', 'ur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hi', ',', 'ur']], ['NUM', '639-1']], 8, True)

["Ge'ez,", 'Ethiopic', 'Spoken', 'in', 'Ethiopia', ',', 'Eritrea', 'and', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Ethiopic', 'Spoken', 'in', 'Ethiopia', ',', 'Eritrea', 'and', 'Israel']]], 9, True)

["Ge'ez,", 'Writing', 'system', "Ge'ez", 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

["Ge'ez,", 'Pronunciation', '\xc9', '\xa1', '\xc9', '\xa8', '\xca', '\x95', '\xc9', '\xa8', 'z']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9', '\xa1', '\xc9', '\xa8', '\xca', '\x95', '\xc9', '\xa8', 'z']]], 11, True)

["Ge'ez,", 'Official', 'language', 'in', 'Liturgical', 'language', 'of',

'the', 'Ethiopian', 'Orthodox', 'Tewahedo', 'Church']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Liturgical', 'language', 'of', 'the', 'Ethiopian', 'Orthodox',  
'Tewahedo', 'Type: ANY']]], 12, True)

["Ge'ez,", 'Official', 'language', 'in', 'Eritrean', 'Orthodox',  
'Tewahedo', 'Church']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Eritrean', 'Orthodox', 'Tewahedo', 'Type: ANY']]], 8, True)

["Ge'ez,", 'Official', 'language', 'in', 'Ethiopic', 'Catholic', 'Church']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Ethiopic', 'Catholic', 'Type: ANY']]], 7, True)

["Ge'ez,", 'Official', 'language', 'in', 'and', 'Beta', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'and', 'Beta', 'Israel']]], 7, True)

["Ge'ez,", 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

["Ge'ez,", 'Language', 'codes', 'ISO', '639-2', 'gez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

["Ge'ez,", 'Language', 'codes', 'ISO', '639-3', 'gez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ad\xc3\xbbnaic', 'Category', '(', 'sources', ')', 'a', 'priori',  
'language', ',', 'but', 'related', 'to', 'other', 'languages', 'of',  
'Arda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'a', 'priori', 'language', ',', 'but', 'related', 'to',  
'other', 'languages', 'of', 'Arda']]], 16, True)

['Ad\xc3\xbbnaic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Ad\xc3\xbbnaic', 'Setting', 'and', 'usage', 'Fantasy', 'world', 'of',  
'Middle-earth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Fantasy', 'world', 'of', 'Middle-earth']]], 8, True)

['Ad\\xc3\\xbbnaic', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.',  
'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',  
'.', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Ad\\xc3\\xbbnaic', 'Category', '(', 'purpose', ')', 'constructed',  
'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Ad\\xc3\\xbbnaic', 'Category', '(', 'purpose', ')', 'artistic',  
'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Ad\\xc3\\xbbnaic', 'Category', '(', 'purpose', ')', 'fictional',  
'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Ad\\xc3\\xbbnaic', 'Category', '(', 'purpose', ')', 'Ad', '\\xc3', '\\xbb',  
'naic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Ad', '\\xc3', '\\xbb', 'naic']]], 9, True)

['Ad\\xc3\\xbbnaic', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ad\\xc3\\xbbnaic', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',  
'\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Ainu', 'Region', 'Hokkaid', '\\xc5', '\\x8d']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ainu', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ainu', 'Writing', 'system', 'Katakana', ',', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', ',', 'Type: ANY', 'alphabet']]], 7, True)

['Ainu', 'Pronunciation', '\xcb', '\x88', 'ainu', 'i', '\xcb', '\x88', 'tak']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'ainu', 'i', '\xcb', '\x88', 'tak']], 9, True)

['Ainu', 'Language', 'codes', 'ISO', '639-3', 'ain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kazakh', 'Region', 'Central', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Mongolia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Tajikistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Turkmenistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kazakh', 'Spoken', 'in', 'Uzbekistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kazakh', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kazakh', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kazakh', 'Writing', 'system', 'Cyrillic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kazakh', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kazakh', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kazakh', 'Pronunciation', 'q', '\xc9', '\x91', 'z', '\xc9', '\x91', 'q', '\xcb', '\x88', 'tl', '\xc9', '\x99']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'q', '\xc9', '\x91', 'z', '\xc9', '\x91', 'q', '\xcb', '\x88', 'tl', '\xc9', '\x99']]], 14, True)

['Kazakh', 'Official', 'language', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Kazakh', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kazakh', 'Language', 'codes', 'ISO', '639-2', 'kaz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kazakh', 'Language', 'codes', 'ISO', '639-3', 'kaz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kazakh', 'Language', 'codes', 'ISO', '639-3', 'Area', 'with',



'significant', 'Kazakh-speaking', 'population']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Area', 'with', 'significant', 'Kazakh-speaking', 'population' ] ], [ 'NUM', '639-3' ] ], 10, True)

['Kazakh', 'Language', 'codes', 'ISO', '639-1', 'kk']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-1' ] ], 6, True)

['Uzbek', 'Spoken', 'in', 'Uzbekistan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'Kyrgyzstan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'Afghanistan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'Kazakhstan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'Turkmenistan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'Tajikistan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'Russia']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Spoken', 'in', 'China']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Uzbek', 'Official', 'language', 'in', 'Uzbekistan']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Official', 'language', 'in', 'Type: ANY' ] ] ], 5, True)

['Uzbek', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-2', 'uzb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-3', 'uzb', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'uzb', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Uzbek', 'Language', 'codes', 'ISO', '639-1', 'uz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tok', 'Total', 'speakers', '5-6', 'million', ';', 'approx', '.', '1', 'million', 'native', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', ';', 'approx', '.', 'NUM', 'million', 'native', 'speakers']], ['NUM', '5-6'], ['NUM', '1']], 12, True)

['Tok', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tok', 'Language', 'codes', 'ISO', '639-2', 'tpi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tok', 'Language', 'codes', 'ISO', '639-3', 'tpi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tok', 'Language', 'codes', 'Linguasphere', '52-ABB-cc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Serbian', 'Pronunciation', 'sr', '\xcc', '\xa9', '\xcc', '\x82', 'pski', '\xcb', '\x90']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sr',

'\xcc', '\xa9', '\xcc', '\x82', 'pski', '\xcb', '\x90']]], 10, True)

['Serbian', 'Official', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Serbian', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Bosnia', 'and', 'Herzegovina']]], 7, True)

['Serbian', 'Official', 'language', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Greece', '(',  
'in', 'Mount', 'Athos', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Greece', '\\(', 'in', 'Mount', 'Athos', '\\)']]], 11,  
True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Recognised', 'minority', 'language', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Serbian', 'Language', 'codes', 'ISO', '639-2', 'srp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Serbian', 'Language', 'codes', 'ISO', '639-3', 'srp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Serbian', 'Language', 'codes', 'ISO', '639-1', 'sr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['(Paraguayan)', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(Paraguayan)', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(Paraguayan)', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(Paraguayan)', 'Pronunciation', '\xca', '\x94', '\xc3', '\xa3', '\xca', '\x8b', '\xcc', '\x83', '\xc3', '\xa3', '\xc9', '\xb2', '\xe1', '\xba', '\xbd', '\xcb', '\x88', '\xca', '\x94', '\xe1', '\xba', '\xbd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca', '\x94', '\xc3', '\xa3', '\xca', '\x8b', '\xcc', '\x83', '\xc3', '\xa3', '\xc9', '\xb2', '\xe1', '\xba', '\xbd', '\xcb', '\x88', '\xca', '\x94', '\xe1', '\xba', '\xbd']], 24, True)

['(Paraguayan)', 'Official', 'language', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['(Paraguayan)', 'Official', 'language', 'in', 'Corrientes', '(' , 'Argentina', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Corrientes', '\\(', 'Argentina', '\\)']], 8, True)

['(Paraguayan)', 'Official', 'language', 'in', 'Brazil', '(' , 'regional', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Brazil', '\\(', 'regional', '\\)']], 8, True)

['(Paraguayan)', 'Official', 'language', 'in', 'Mercosur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['(Paraguayan)', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['(Paraguayan)', 'Language', 'codes', 'ISO', '639-2', 'grn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['(Paraguayan)', 'Language', 'codes', 'ISO', '639-3', 'gug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['(Paraguayan)', 'Language', 'codes', 'ISO', '639-1', 'gn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bislama', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bislama', 'Total', 'speakers', '6', ',', '200', '(', 'first', 'language',  
)', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\\\(', 'first', 'language', '\\\\)', ',', '']], ['NUM', '6'],  
'NUM', '200']], 11, True)

['Bislama', 'Total', 'speakers', '200', ',', '000', '(', 'additional',  
'language', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\\\(', 'additional', 'language', '\\\\)', '']], ['NUM',  
'200'], ['NUM', '000']], 10, True)

['Bislama', 'Official', 'language', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Bislama', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Bislama', 'Language', 'codes', 'ISO', '639-2', 'bis']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bislama', 'Language', 'codes', 'ISO', '639-3', 'bis']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bislama', 'Language', 'codes', 'Linguasphere', '52-ABB-ce']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

```

['Bislama', 'Language', 'codes', 'ISO', '639-1', 'bi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Lithuanian', 'Total', 'speakers', '500', ',', '000', '(', 'Abroad', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Abroad', '\\)']], ['NUM', '500'], ['NUM',
'000']], 9, True)

['Lithuanian', 'Total', 'speakers', '4', 'million', '(', 'Worldwide', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', '\\(', 'Worldwide', '\\)']], ['NUM', '4']], 8, True)

['Lithuanian', 'Dialects', 'Samogitian']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Lithuanian', 'Region', 'Europe']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Lithuanian', 'Spoken', 'in', 'Lithuania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Lithuanian', 'Writing', 'system', 'Latin', '(', 'Lithuanian', 'variant',
')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Lithuanian', 'Official', 'language', 'in', 'Lithuania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']], 5, True)

['Lithuanian', 'Official', 'language', 'in', 'European', 'Union']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Lithuanian', 'Official', 'language', 'in', 'Gmina', 'Pu', '\\xc5', '\\x84',
'sk', 'in', 'the', 'Podlaskie', 'Voivodeship', 'of', 'Poland']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Gmina', 'Pu', '\\xc5', '\\x84', 'sk', 'in', 'the', 'Podlaskie',
'Voivodeship', 'of', 'Poland']], 15, True)

['Lithuanian', 'Language', 'codes', 'ISO', '639-2', 'lit']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lithuanian', 'Language', 'codes', 'ISO', '639-3', 'lit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lithuanian', 'Language', 'codes', 'Linguasphere', '54-AAA-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Lithuanian', 'Language', 'codes', 'ISO', '639-1', 'lt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Alsatian', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Alsatian', 'Region', 'Alsace']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Alsatian', 'Language', 'codes', 'ISO', '639-3', 'gsw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kyrgyz', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Xinjiang', '(', 'China', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)

['Kyrgyz', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kyrgyz', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kyrgyz', 'Official', 'language', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Kyrgyz', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Kyrgyz', 'Language', 'codes', 'ISO', '639-2', 'kir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kyrgyz', 'Language', 'codes', 'ISO', '639-3', 'kir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kyrgyz', 'Language', 'codes', 'ISO', '639-1', 'ky']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ganda', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ganda', 'Region', 'Mainly', 'Buganda', 'region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ganda', 'Language', 'codes', 'ISO', '639-2', 'lug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ganda', 'Language', 'codes', 'ISO', '639-3', 'lug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ganda', 'Language', 'codes', 'ISO', '639-1', 'lg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tigrinya', 'Region', 'Eritrea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Tigrinya', 'Region', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Tigrinya', 'Region', 'especially', 'in', 'Tigray']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Tigrinya', 'Spoken', 'in', 'Eritrea', ',', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Tigrinya', 'Writing', 'system', 'Ge'ez', 'alphabet', 'abugida']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

```
['Tigrinya', 'Pronunciation', '/', 't', '\xc9', '\xa8', '\xc9', '\xa1',  
'r', '\xc9', '\xa8', '\xc9', '\xb2', 'a', '/']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
't', '\xc9', '\xa8', '\xc9', '\xa1', 'r', '\xc9', '\xa8', '\xc9', '\xb2',  
'a', '/']]], 15, True)
```

```
['Tigrinya', 'Official', 'language', 'in', 'Eritrea', '(', 'working',  
'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Eritrea', '\\(', 'working', 'language', '\\)']]], 9, True)
```

```
['Tigrinya', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Tigrinya', 'Language', 'codes', 'ISO', '639-2', 'tir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Tigrinya', 'Language', 'codes', 'ISO', '639-3', 'tir']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tigrinya', 'Language', 'codes', 'ISO', '639-1', 'ti']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Amharic', 'Total', 'speakers', '17', ',', '528', ',', '500', ';', '14',  
, ',', '743', ',', '556', 'monolingual', '(', 'Ethnologue', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', ';', 'NUM', ',', 'NUM', ',', 'NUM',  
'monolingual', '\\(', 'Ethnologue', '\\)']]], ['NUM', '17'], ['NUM', '528'],  
['NUM', '500'], ['NUM', '14'], ['NUM', '743'], ['NUM', '556']], 18, True)
```

```
['Amharic', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Amharic', 'Spoken', 'in', 'Eritrea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Amharic', 'Writing', 'system', 'Ge'ez', 'alphabet', 'abugida']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Amharic', 'Pronunciation', 'amar', '\xc9', '\xa8', '\xc9', '\xb2',  
'\xc9', '\xb2', 'a']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'amar',  
'\xc9', '\xa8', '\xc9', '\xb2', '\xc9', '\xb2', 'a']]]], 10, True)
```

```
['Amharic', 'Official', 'language', 'in', 'Ethiopia', 'and', 'the',  
'following', 'specific', 'regions', ':', 'Addis', 'Ababa', 'City',  
'Council']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Ethiopia', 'and', 'the', 'following', 'specific', 'regions', ':',  
'Addis', 'Ababa', 'City', 'Council']]]], 15, True)
```

```
['Amharic', 'Official', 'language', 'in', 'Amhara', 'Region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Amharic', 'Official', 'language', 'in', 'Benishangul-Gumuz', 'Region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Amharic', 'Official', 'language', 'in', 'Dire', 'Dawa', 'Administrative',  
'council']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Dire', 'Dawa', 'Administrative', 'council']]]], 8, True)
```

```
['Amharic', 'Official', 'language', 'in', 'Gambela', 'Region']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['Amharic', 'Official', 'language', 'in', 'SNNPR']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['Amharic', 'Regulated', 'by', 'no', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]]], 6, True)
```

```
['Amharic', 'Language', 'codes', 'ISO', '639-2', 'amh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Amharic', 'Language', 'codes', 'ISO', '639-3', 'amh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Amharic', 'Language', 'codes', 'Linguasphere', '12-ACB-a']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Amharic', 'Language', 'codes', 'ISO', '639-1', 'am']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Wolof', 'Spoken', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wolof', 'Spoken', 'in', 'Gambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wolof', 'Spoken', 'in', 'Mauritania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wolof', 'Region', 'West', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wolof', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Wolof', 'Regulated', 'by', 'CLAD', '(', 'Centre', 'de', 'linguistique', 'appliqu', '\xc3', '\xa9', 'e', 'de', 'Dakar', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'CLAD', '\\(', 'Centre', 'de', 'linguistique', 'appliqu', '\xc3', '\xa9', 'e', 'de', 'Dakar', '\\)']]], 15, True)

['Wolof', 'Language', 'codes', 'ISO', '639-2', 'wol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Wolof', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wolof', 'Language', 'codes', 'ISO', '639-3', ':', 'wol', '\xe2', '\x80', '\x93', 'Wolof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'wol', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Wolof', 'Language', 'codes', 'ISO', '639-3', 'wof', '\xe2', '\x80', '\x93', 'Gambian', 'Wolof']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wof', '\xe2', '\x80', '\x93', 'Gambian', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Wolof', 'Language', 'codes', 'ISO', '639-1', 'wo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Xhosa', 'Spoken', 'in', 'South', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Xhosa', 'Spoken', 'in', 'Lesotho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xhosa', 'Official', 'language', 'in', 'South', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Xhosa', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Xhosa', 'Language', 'codes', 'ISO', '639-2', 'xho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Xhosa', 'Language', 'codes', 'ISO', '639-3', 'xho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xhosa', 'Language', 'codes', 'ISO', '639-1', 'xh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Istro-Romanian', 'Total', 'speakers', '1000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Istro-Romanian', 'Spoken', 'in', 'Croatia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Istro-Romanian', 'Region', 'Istria']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Istro-Romanian', 'Language', 'codes', 'ISO', '639-3', 'ruo']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Istro-Romanian', 'Language', 'codes', 'Linguasphere', '51-AAD-a']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Tuvaluan', 'Spoken', 'in', 'Tuvalu']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvaluan', 'Spoken', 'in', 'Fiji']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvaluan', 'Spoken', 'in', 'Kiribati']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvaluan', 'Spoken', 'in', 'Nauru']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvaluan', 'Spoken', 'in', 'New', 'Zealand']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tuvaluan', 'Total', 'speakers', '10', ',', '670', 'in', 'Tuvalu', '.', '13', ',', '051', 'total']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Tuvalu', '.', 'NUM', ',', 'NUM', 'total']], ['NUM', '10'], ['NUM', '670'], ['NUM', '13'], ['NUM', '051']], 13, True)

['Tuvaluan', 'Official', 'language', 'in', 'Tuvalu']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Tuvaluan', 'Regulated', 'by', 'No', 'official', 'regulation']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Tuvaluan', 'Language', 'codes', 'ISO', '639-2', 'tvl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tuvaluan', 'Language', 'codes', 'ISO', '639-3', 'tvl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Region', 'mare', 'nostrum', '(', 'Mediterranean', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'mare',  
'nostrum', '\\(', 'Mediterranean', '\\)']], 7, True)

['Classical', 'Writing', 'system', 'Latin', 'alphabet', 'Literature']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Classical', 'Writing', 'system', 'documents']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Classical', 'Writing', 'system', 'inscriptions']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Classical', 'Pronunciation', 'la', '\\xcb', '\\x88', 'ti', '\\xcb', '\\x90',  
'nita', '\\xcb', '\\x90', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'la',  
'\\xcb', '\\x88', 'ti', '\\xcb', '\\x90', 'nita', '\\xcb', '\\x90', 's']], 12,  
True)

['Classical', 'Official', 'language', 'in', 'Roman', 'republic', ',',  
'Roman', 'empire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Roman', 'republic', ',', 'Roman', 'empire']], 9, True)

['Classical', 'Regulated', 'by', 'Schools', 'of', 'grammar', 'and',  
'rhetoric']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Schools', 'of', 'grammar', 'and', 'rhetoric']], 8, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'lat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',

```

'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'lat-cla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Language', 'codes', 'ISO', '639-1', 'la']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Pashto', 'Total', 'speakers', 'estimated', '50', 'to', '60', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'estimated', 'NUM', 'to', 'NUM', 'million']], ['NUM', '50'], ['NUM',
'60']], 8, True)

['Pashto', 'Region', 'South-Central', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pashto', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pashto', 'Spoken', 'in', 'Pakistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pashto', 'Spoken', 'in', 'Iran', '(', 'minor', ')', 'and', 'by', 'the',
'Pashtun', 'diaspora', 'around', 'the', 'world']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iran',
'\\(', 'minor', '\\)', 'and', 'by', 'the', 'Pashtun', 'diaspora', 'around',
'the', 'world']]], 15, True)

['Pashto', 'Writing', 'system', 'Pashto', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Pashto', 'Pronunciation', 'pa', '\xca', '\x82', '\xcb', '\x88', 'to']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pa',
'\xca', '\x82', '\xcb', '\x88', 'to']]], 8, True)

['Pashto', 'Pronunciation', 'pa', '\xc3', '\xa7', '\xcb', '\x88', 'to']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pa',

```



'\xc3', '\xa7', '\xcb', '\x88', 'to']]], 8, True)

['Pashto', 'Pronunciation', 'pax', '\xcb', '\x88', 'to']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'pax',  
'\xcb', '\x88', 'to']]], 6, True)

['Pashto', 'Official', 'language', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Pashto', 'Official', 'language', 'in', 'Pakistan', '(', 'provincial',  
'language', 'in', 'Khyber', 'Pakhtunkhwa', '.', 'and', 'FATA', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Pakistan', '\\(', 'provincial', 'language', 'in', 'Khyber',  
'Pakhtunkhwa', '.', 'and', 'FATA', '\\)']]], 15, True)

['Pashto', 'Regulated', 'by', 'Academy', 'of', 'Sciences', 'of',  
'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Academy', 'of', 'Sciences', 'of', 'Afghanistan']]], 8, True)

['Pashto', 'Language', 'codes', 'ISO', '639-2', 'pus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pashto', 'Language', 'codes', 'ISO', '639-3', 'pus', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'pus', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Pashto', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Pashto', 'Language', 'codes', 'Linguasphere', '58-ABD-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Pashto', 'Language', 'codes', 'ISO', '639-1', 'ps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Classical', 'Spoken', 'in', 'mainland', 'China', ';', 'Taiwan', ';',  
'Japan', ';', 'Korea', 'and', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',

'mainland', 'China', ';', 'Taiwan', ';', 'Japan', ';', 'Korea', 'and',  
'Vietnam']]], 13, True)

['Classical', 'Total', 'speakers', 'Not', 'a', 'spoken', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Not', 'a', 'spoken', 'language']]], 7, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Classical', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'lzh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Malay', 'Spoken', 'in', 'Brunei']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Malay', 'Spoken', 'in', 'Singapore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Malay', 'Spoken', 'in', 'Philippines', '(', 'mostly', 'in', 'Tawi',  
'Tawi', 'and', 'Sulu', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Philippines', '\\\\(', 'mostly', 'in', 'Tawi', 'Tawi', 'and', 'Sulu', '\\  
\\)']]], 12, True)

['Malay', 'Spoken', 'in', 'Thailand']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Malay', 'Spoken', 'in', 'Myanmar']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Malay', 'Spoken', 'in', 'Sri', 'Lanka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Malay', 'Spoken', 'in', 'Cocos', '(', 'Keeling', ')', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Cocos', '\\\\(', 'Keeling', '\\\\)', 'Islands']]], 8, True)

['Malay', 'Spoken', 'in', 'Christmas', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Malay', 'Dialects', 'Malay', 'dialects']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Malay', 'Standard', 'forms', 'Malaysian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Malay', 'Standard', 'forms', 'Indonesian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Malay', 'Official', 'language', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Malay', 'Official', 'language', 'in', 'Brunei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Malay', 'Official', 'language', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Malay', 'Language', 'codes', 'ISO', '639-2', 'may', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'may', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Malay', 'Language', 'codes', 'ISO', '639-2', 'msa', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'msa', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'msa', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'msa', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', ':', 'zsm', '\xe2', '\x80', '\x93', 'Malaysian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'zsm', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'ind', '\xe2', '\x80', '\x93', 'Indonesian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ind', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Malay', 'Language', 'codes', 'ISO', '639-3', 'coa', '\xe2', '\x80', '\x93', 'Cocos', 'Islands', 'Malay']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'coa', '\xe2', '\x80', '\x93', 'Cocos', 'Islands', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Malay', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Valencian', 'Region', 'of', 'Murcia', '(', 'Carche', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'of', 'Murcia', '\\(', 'Carche', '\\)']], 7, True)

['Valencian', 'Region', 'See', 'also', 'geographic', 'distribution', 'of', 'Catalan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'See', 'also', 'geographic', 'distribution', 'of', 'Type: ANY']], 8, True)

['Valencian', 'Spoken', 'in', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Valencian', 'Spoken', 'in', 'Valencian', 'Community']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

```
['Valencian', 'Spoken', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3,  
True)
```

```
['Valencian', 'Writing', 'system', 'Latin', '(', 'Catalan', 'variant',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Valencian', 'Pronunciation', 'valensi', '\xcb', '\x88', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation',  
'valensi', '\xcb', '\x88', 'a']]], 6, True)
```

```
['Valencian', 'Official', 'language', 'in', 'In', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Valencian', 'Regulated', 'by', 'Acad', '\xc3', '\xa8', 'mia',  
'Valenciana', 'de', 'la', 'Llengua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Acad', '\xc3', '\xa8', 'mia', 'Valenciana', 'de', 'la', 'Llengua']]], 11,  
True)
```

```
['Valencian', 'Language', 'codes', 'ISO', '639-2', 'cat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Valencian', 'Language', 'codes', 'ISO', '639-3', 'cat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Valencian', 'Language', 'codes', 'Linguasphere', '51-AAA-e']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Valencian', 'Language', 'codes', 'ISO', '639-1', 'ca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Tutchone', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tutchone', 'Region', 'Yukon', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tutchone', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tutchone', 'Language', 'codes', 'ISO', '639-3', ':', 'tce', '\xe2',  
\x80', '\x93', 'Southern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'tce', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Tutchone', 'Language', 'codes', 'ISO', '639-3', 'ttm', '\xe2', '\x80',  
\x93', 'Northern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ttm', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Lao', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lao', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lao', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lao', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lao', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lao', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lao', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Lao', 'Spoken', 'in', 'Argentina', '(', 'Misiones', 'Province', ')']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Argentina', '\\(', 'Misiones', 'Province', '\\)']]], 8, True)

['Lao', 'Writing', 'system', 'Lao', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lao', 'Pronunciation', 'p', '\\xca', '\\xb0', '\\xc3', '\\xa1', '\\xcb', '\\x90', 's', '\\xc7', '\\x8e', '\\xcb', '\\x90', 'l', '\\xc3', '\\xa1', '\\xcb', '\\x90', 'w']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\\xca', '\\xb0', '\\xc3', '\\xa1', '\\xcb', '\\x90', 's', '\\xc7', '\\x8e', '\\xcb', '\\x90', 'l', '\\xc3', '\\xa1', '\\xcb', '\\x90', 'w']]], 20, True)

['Lao', 'Official', 'language', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Lao', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Lao', 'Language', 'codes', 'ISO', '639-2', 'lao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Lao', 'Language', 'codes', 'ISO', '639-3', 'lao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lao', 'Language', 'codes', 'ISO', '639-1', 'lo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Gujarati', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gujarati', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gujarati', 'Spoken', 'in', 'Mauritius']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Portugal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Spoken', 'in', 'Panama']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gujarati', 'Writing', 'system', 'Gujarati', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gujarati', 'Writing', 'system', 'mainly', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Gujarati', 'Pronunciation', '/', '\xc9', '\xa1', 'ud', '\xca', '\x92', '(', '\xc9', '\x99', ')', '\xcb', '\x88', '\xc9', '\xbe', 'at', '\xcc', '\xaa', 'i', '/']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', '\xc9', '\xa1', 'ud', '\xca', '\x92', '\\(', '\xc9', '\x99', '\\)', '\xcb', '\x88', '\xc9', '\xbe', 'at', '\xcc', '\xaa', 'i', '/']]], 21, True)

['Gujarati', 'Official', 'language', 'in', 'Gujarat', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\(', 'India', '\\)']]], 8, True)

['Gujarati', 'Official', 'language', 'in', 'Dadra', 'and', 'Nagar', 'Haveli', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Dadra', 'and', 'Nagar', 'Haveli', '\\(', 'India', '\\)']]], 11, True)

['Gujarati', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Gujarati', 'Language', 'codes', 'ISO', '639-2', 'guj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gujarati', 'Language', 'codes', 'ISO', '639-3', 'guj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gujarati', 'Language', 'codes', 'ISO', '639-1', 'gu']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-1' ] ], 6, True)

['Southern', 'Spoken', 'in', 'South', 'Africa']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY' ] ] ], 5, True)

['Southern', 'Region', 'Mpumalanga']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 3, True)

['Southern', 'Region', 'Limpopo']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 3, True)

['Southern', 'Region', 'Gauteng']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 3, True)

['Southern', 'Region', 'North', 'West']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Southern', 'Official', 'language', 'in', 'South', 'Africa']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY' ] ] ], 6, True)

['Southern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY' ] ] ], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'nbl']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-2' ] ], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'nbl']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-1', 'nr']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-1' ] ], 6, True)

['Dalmatian', 'Spoken', 'in', 'Croatia', ',', 'Montenegro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Dalmatian', 'Region', 'Adriatic', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dalmatian', 'Language', 'codes', 'ISO', '639-3', 'dln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Dalmatian', 'Language', 'codes', 'Linguasphere', '51-AAA-t']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Vandalic', 'Language', 'codes', 'ISO', '639-3', 'xvn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Romani', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Macedonia']]], 8, True)

['Romani', 'Official', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Slovenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Romani', 'Official', 'language', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Romani', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Romani', 'Language', 'codes', 'ISO', '639-2', 'rom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rom', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'rom', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rml', '\xe2', '\x80',  
'\x93', 'Baltic', 'Romani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'rml', '\xe2', '\x80', '\x93', 'Baltic', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmc', '\xe2', '\x80',  
'\x93', 'Carpathian', 'Romani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'rmc', '\xe2', '\x80', '\x93', 'Carpathian', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmo', '\xe2', '\x80',  
'\x93', 'Sinte', 'Romani']:

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rmo', '\xe2', '\x80', '\x93', 'Sinte', 'Type: ANY']], ['NUM', '639-3']], 11, True)
```

```
['Yue', 'Dialects', 'Yuehai', '(', 'Cantonese', ')']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)
```

```
['Yue', 'Dialects', 'Siyi', '(', 'Taishanese', ')']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)
```

```
['Yue', 'Dialects', 'Gao-Yang']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Yue', 'Dialects', 'Yong-Xun']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Yue', 'Dialects', 'Goulou']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Yue', 'Dialects', 'Luo-Guang']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Yue', 'Dialects', 'Qin-Lian']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Yue', 'Dialects', 'Wu-Hua']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Yue', 'Dialects', 'Hainan', 'Yue']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Yue', 'Writing', 'system', 'Traditional', 'Chinese']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Yue', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
```

ANY', 'of', 'Type: ANY']]], 6, True)

['Yue', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Yue', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Yue', 'Language', 'codes', 'ISO', '639-3', 'yue']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yue', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Siwi', 'Spoken', 'in', 'Egypt']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Siwi', 'Region', 'Siwa', 'Oasis', ',', 'Gara', 'Oasis']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Siwa',  
'Oasis', ',', 'Gara', 'Oasis']]], 7, True)

['Siwi', 'Language', 'codes', 'ISO', '639-3', 'siz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zenaga', 'Spoken', 'in', 'Mauritania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Zenaga', 'Region', 'Mederdra']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Zenaga', 'Language', 'codes', 'ISO', '639-2', 'zen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zenaga', 'Language', 'codes', 'ISO', '639-3', 'zen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Tuareg', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tuareg', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Tuareg', 'Spoken', 'in', 'Libya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tuareg', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tuareg', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tuareg', 'Region', 'Sahara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-2', 'tmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'tmh', '\xe2', '\x80',  
\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tmh', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-3', ':', 'thv', '\xe2', '\x80',  
\x93', 'Tahaggart', 'Tamahaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'thv', '\xe2', '\x80', '\x93', 'Tahaggart', 'Type:  
ANY']]], ['NUM', '639-3']], 12, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'taq', '\xe2', '\x80',  
\x93', 'Tamasheq']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'taq', '\xe2', '\x80', '\x93', 'Tamasheq']], ['NUM', '639-3']], 10, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'ttq', '\xe2', '\x80', '\x93', 'Tawallammat', 'Tamajaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ttq', '\xe2', '\x80', '\x93', 'Tawallammat', 'Tamajaq']], ['NUM', '639-3']], 11, True)
```

```
['Tuareg', 'Language', 'codes', 'ISO', '639-3', 'thz', '\xe2', '\x80', '\x93', 'Tayart', 'Tamajeq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'thz', '\xe2', '\x80', '\x93', 'Tayart', 'Tamajeq']], ['NUM', '639-3']], 11, True)
```

```
['Ancient', 'Spoken', 'in', 'eastern', 'Mediterranean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Ancient', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Ancient', 'Language', 'codes', 'ISO', '639-2', 'grc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Ancient', 'Language', 'codes', 'ISO', '639-3', 'grc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Black', 'Speech', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Speech', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']], 17, True)
```

```
['Black', 'Setting', 'and', 'usage', 'Mordor', 'in', 'Middle-earth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Mordor', 'in', 'Middle-earth']], 7, True)
```

```
['Black', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']], 10, True)
```



```
['Black', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Black', 'Category', '(', 'purpose', ')', 'artistic', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Black', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Polabian', 'Spoken', 'in', 'Poland', ',', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Polabian', 'Language', 'codes', 'ISO', '639-3', 'pox']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Polabian', 'Language', 'codes', 'Linguasphere', '53-AAA-bc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Slovincian', 'Spoken', 'in', 'Poland', ',', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Slovincian', 'Region', 'Pomerania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Slovincian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Slovincian', 'Language', 'codes', 'ISO', '639-3', 'csb-slo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Slovincian', 'Language', 'codes', 'Linguasphere', '53-AAA-ca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

['Knaanic', 'Spoken', 'in', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Knaanic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Knaanic', 'Language', 'codes', 'ISO', '639-3', 'czk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rusyn', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Slovakia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rusyn', 'Official', 'language', 'in', 'Vojvodina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Rusyn', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:

ANY', 'of', 'Type: ANY']]], 6, True)

['Rusyn', 'Language', 'codes', 'ISO', '639-3', 'rue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romano-Serbian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Romano-Serbian', 'Language', 'codes', 'ISO', '639-3', 'rsb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Manchu', 'Region', 'Heilongjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Manchu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Manchu', 'Writing', 'system', 'Manchu', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Manchu', 'Language', 'codes', 'ISO', '639-2', 'mnc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Manchu', 'Language', 'codes', 'ISO', '639-3', 'mnc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cimbrian', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cimbrian', 'Language', 'codes', 'ISO', '639-3', 'cim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Austro-Bavarian', 'Spoken', 'in', 'Austria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Austro-Bavarian', 'Spoken', 'in', 'Germany', ',', 'Bavaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Austro-Bavarian', 'Total', 'speakers', '12', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '12']], 5, True)
```

```
['Austro-Bavarian', 'Language', 'codes', 'ISO', '639-3', 'bar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Galindan,', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Galindan,', 'Language', 'codes', 'ISO', '639-3', 'xgl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Georgian', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Georgian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Georgian', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Georgian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Georgian', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Georgian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Georgian', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Georgian', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Georgian', 'Total', 'speakers', '6-7', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '6-7']], 5, True)

['Georgian', 'Writing', 'system', 'Georgian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Georgian', 'Official', 'language', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Georgian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Georgian', 'Language', 'codes', 'ISO', '639-2', 'geo', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'geo', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Georgian', 'Language', 'codes', 'ISO', '639-2', 'kat', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kat', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Georgian', 'Language', 'codes', 'ISO', '639-3', 'kat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Georgian', 'Language', 'codes', 'ISO', '639-1', 'ka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Mongolian', 'Region', 'All', 'of', 'Mongolia', 'and', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Liaoning']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'All', 'of', 'Mongolia', 'and', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Liaoning']]], 12, True)

['Mongolian', 'Region', 'Jilin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY']]], 3, True)

['Mongolian', 'Region', 'and', 'Heilongjiang', 'provinces', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Heilongjiang', 'provinces', 'in', 'China']]], 7, True)

['Mongolian', 'Pronunciation', '/', 'm', '\xc9', '\x94', '\xc5', '\x8b',  
'\xc9', '\xa2', '\xc9', '\x94', '\xcc', '\x86', '\xc9', '\xae', 'xi',  
'\xc9', '\xae', '/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',  
'm', '\xc9', '\x94', '\xc5', '\x8b', '\xc9', '\xa2', '\xc9', '\x94',  
'\xcc', '\x86', '\xc9', '\xae', 'xi', '\xc9', '\xae', '/']]], 20, True)

['Mongolian', 'Official', 'language', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-2', 'mon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'mon', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mon', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', ':', 'khk', '\xe2',  
'\x80', '\x93', 'Khalkha', 'Mongolian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'khk', '\xe2', '\x80', '\x93', 'Khalkha', 'Type: ANY']],  
['NUM', '639-3']], 12, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'mvf', '\xe2', '\x80',  
'\x93', 'Peripheral', 'Mongolian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mvf', '\xe2', '\x80', '\x93', 'Peripheral', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Mongolian', 'Language', 'codes', 'ISO', '639-1', 'mn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

```
['Manglish', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Manglish', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']], 6, True)
```

```
['Manglish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Rotokas', 'Region', 'Bougainville']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Rotokas', 'Language', 'codes', 'ISO', '639-3', 'roo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Luba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)
```

```
['Luba', 'Region', 'Kasai-Occidental', 'and', 'Kasai-Oriental',  
'provinces']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kasai-  
Occidental', 'and', 'Kasai-Oriental', 'provinces']], 6, True)
```

```
['Luba', 'Language', 'codes', 'ISO', '639-2', 'lua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Luba', 'Language', 'codes', 'ISO', '639-3', 'lua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Burushaski', 'Spoken', 'in', 'Gilgit', '\xe2', '\x80', '\x93',  
'Baltistan', ',', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Gilgit',  
'\xe2', '\x80', '\x93', 'Baltistan', ',', 'Pakistan']], 10, True)
```

```
['Burushaski', 'Spoken', 'in', 'Jammu', 'and', 'Kashmir', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Jammu',  
'and', 'Kashmir', ',', 'India']], 8, True)
```

```

['Burushaski', 'Language', 'codes', 'ISO', '639-3', 'bsk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Acadian', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Acadian', 'Spoken', 'in', ':', 'New', 'Brunswick']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',
'Type: ANY', 'Type: ANY']]], 6, True)

['Acadian', 'Spoken', 'in', 'Nova', 'Scotia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Acadian', 'Spoken', 'in', 'Prince', 'Edward', 'Island']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Prince',
'Edward', 'Island']]], 6, True)

['Acadian', 'Spoken', 'in', 'Magdalen', 'Islands', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Magdalen', 'Islands', ',']]], 6, True)

['Acadian', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Acadian', 'Spoken', 'in', ':', 'Maine']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':',
'Type: ANY']]], 5, True)

['Acadian', 'Language', 'codes', 'ISO', '639-3', 'fra-aca', '\xe2', '\x80',
'\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'fra-aca', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,
True)

['Acadian', 'Language', 'codes', 'Linguasphere', '51-AAA-ho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Norwegian', 'Total', 'speakers', '5', 'million', 'Norwegians']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'million', 'Norwegians']], ['NUM', '5']], 6, True)

```



['Norwegian', 'Writing', 'system', 'Latin', '(', 'Norwegian', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Norwegian', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']], 8, True)

['Norwegian', 'Official', 'language', 'in', 'Norway']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Norwegian', 'Official', 'language', 'in', 'Nordic', 'Council']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-2']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'nor', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

```
['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nob', '\xe2', '\x80',  
\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']],  
['NUM', '639-3']], 13, True)
```

```
['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nno', '\xe2', '\x80',  
\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-3']],  
10, True)
```

```
['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-cd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'no', '\xe2', '\x80',  
\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-1']], 10, True)
```

```
['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nb', '\xe2', '\x80',  
\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']],  
['NUM', '639-1']], 13, True)
```

```
['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nn', '\xe2', '\x80',  
\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-1']],  
10, True)
```

```
['Dungan', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)
```

```
['Dungan', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)
```

```
['Dungan', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)
```

['Dungan', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dungan', 'Spoken', 'in', 'and', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dungan', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dungan', 'Pronunciation', 'xw', '\xc9', '\x9b', 'itsu', 'jyj', '\xc9', '\x91', 'nneed', 'tone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'xw', '\xc9', '\x9b', 'itsu', 'jyj', '\xc9', '\x91', 'nneed', 'tone']]], 11, True)

['Dungan', 'Language', 'codes', 'ISO', '639-3', 'dng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Total', 'speakers', '5', 'million', 'Norwegians']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'Norwegians']], ['NUM', '5']], 6, True)

['Norwegian', 'Writing', 'system', 'Latin', '(', 'Norwegian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Norwegian', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\x94', '\xca', '\x82', 'k']]], 8, True)

['Norwegian', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Norwegian', 'Official', 'language', 'in', 'Nordic', 'Council']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-2']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-2', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-2']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', ':', 'nor', '\xe2', '\x80', '\x93', 'Norwegian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'nor', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nob', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-3']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'nno', '\xe2', '\x80', '\x93', 'Nynorsk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-3']], 10, True)

['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-be']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'no', '\xe2', '\x80', '\x93', 'Norwegian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'no', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-1']], 10, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nb', '\xe2', '\x80', '\x93', 'Bokm', '\xc3', '\xa5', 'l']], ['NUM', '639-1']], 13, True)

['Norwegian', 'Language', 'codes', 'ISO', '639-1', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nn', '\xe2', '\x80', '\x93', 'Nynorsk']], ['NUM', '639-1']], 10, True)

['Brandenburgisch', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Brandenburgisch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Brandenburgisch', 'Region', 'Brandenburg', ',', 'Saxony-Anhalt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Brandenburgisch', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Brandenburgisch', 'Language', 'codes', 'ISO', '639-3', 'nds-mar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Haz\xc4\x81rag\xc4\xab', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haz\xc4\x81rag\xc4\xab', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haz\xc4\x81rag\xc4\xab', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Region', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haz\\xc4\\x81rag\\xc4\\xab', 'Language', 'codes', 'ISO', '639-3', 'haz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Navajo', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Navajo', 'Region', 'Arizona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Navajo', 'Region', 'New', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Navajo', 'Region', 'Utah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Navajo', 'Region', 'Colorado']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Navajo', 'Language', 'codes', 'ISO', '639-2', 'nav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Navajo', 'Language', 'codes', 'ISO', '639-3', 'nav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Navajo', 'Language', 'codes', 'ISO', '639-1', 'nv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bosnian', 'Total', 'speakers', '3', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '3']], 5, True)

['Bosnian', 'Region', 'Southern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Bosnian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Spoken', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Spoken', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Spoken', 'in', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bosnian', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Bosnian', 'Pronunciation', 'bosanski', '\xcb', '\x90']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'bosanski', '\xcb', '\x90']], 5, True)

['Bosnian', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Bosnia', 'and', 'Herzegovina']]], 7, True)

['Bosnian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Recognised', 'minority', 'language', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)

['Bosnian', 'Language', 'codes', 'ISO', '639-2', 'bos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bosnian', 'Language', 'codes', 'ISO', '639-3', 'bos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bosnian', 'Language', 'codes', 'ISO', '639-1', 'bs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Okinawan', 'Region', 'The', 'Okinawa', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Okinawan', 'Spoken', 'in', 'Okinawa', 'Prefecture', ',', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Okinawa', 'Prefecture', ',', 'Japan']]], 7, True)

['Okinawan', 'Writing', 'system', 'Okinawan', 'writing', 'system']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Okinawan', 'Writing', 'system', 'Japanese', 'writing', 'system']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Okinawan', 'Pronunciation', '\xca', '\x94', 'ut', '\xcd', '\xa1', '\xc9', '\x95', 'ina', '\xcb', '\x90', '\xc9', '\xa1', 'ut', '\xcd', '\xa1', '\xc9', '\x95', 'i']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca', '\x94', 'ut', '\xcd', '\xa1', '\xc9', '\x95', 'ina', '\xcb', '\x90', '\xc9', '\xa1', 'ut', '\xcd', '\xa1', '\xc9', '\x95', 'i']]], 20, True)

['Okinawan', 'Language', 'codes', 'ISO', '639-3', 'ryu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Okinawan', 'Language', 'codes', 'Linguasphere', '45-CAC-ai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Okinawan', 'Language', 'codes', 'Linguasphere', '45-CAC-aj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Okinawan', 'Language', 'codes', 'Linguasphere', '45-CAC-ak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Texas', 'Spoken', 'in', 'Texas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Texas', 'Region', 'Texas', 'Hill', 'Country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Texas', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Hutterite', 'Spoken', 'in', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hutterite', 'Spoken', 'in', 'Saskatchewan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hutterite', 'Spoken', 'in', 'and', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hutterite', 'Spoken', 'in', 'Canada', ';', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada', ';', 'Washington']]], 6, True)

['Hutterite', 'Spoken', 'in', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hutterite', 'Language', 'codes', 'ISO', '639-3', 'geh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Plautdietsch', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Belize']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Plautdietsch', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Spoken', 'in', 'Uruguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Plautdietsch', 'Language', 'codes', 'ISO', '639-3', 'pdt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mohawk', 'Total', 'speakers', '3', ',', '350', '(', 'Ethnologue', ')', ';', '900', 'in', 'Canada', ';', '2', ',', '017', 'in', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)', ';', 'NUM', 'in', 'Canada', ';', 'NUM', ',', 'NUM', 'in', 'the', 'United', 'States']]], ['NUM', '3'], ['NUM', '350'], ['NUM', '900'], ['NUM', '2'], ['NUM', '017']], 21, True)

['Mohawk', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Mohawk', 'Pronunciation', 'kanj', '\\xca', '\\x8c', '\\xcc', '\\x83', '\\xca', '\\x94', '\\xcb', '\\x88', '\\xc9', '\\xa1', '\\xc9', '\\x9b', 'ha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kanj', '\\xca', '\\x8c', '\\xcc', '\\x83', '\\xca', '\\x94', '\\xcb', '\\x88', '\\xc9',

'\xa1', '\xc9', '\x9b', 'ha']]], 16, True)

['Mohawk', 'Language', 'codes', 'ISO', '639-2', 'moh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mohawk', 'Language', 'codes', 'ISO', '639-3', 'moh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Soran\xc3\xae', 'Total', 'speakers', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '5']], 5, True)

['Soran\xc3\xae', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Soran\xc3\xae', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Soran\xc3\xae', 'Region', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Soran\xc3\xae', 'Language', 'codes', 'ISO', '639-3', 'ckb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Soran\xc3\xae', 'Language', 'codes', 'Linguasphere', '58-AAA-cae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Chinook', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']], 7, True)

['Chinook', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Chinook', 'Language', 'codes', 'ISO', '639-2', 'chn', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chn', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Chinook', 'Language', 'codes', 'ISO', '639-2', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 8, True)

['Chinook', 'Language', 'codes', 'ISO', '639-3', 'chn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chechen', 'Region', 'Republic', 'of', 'Chechnya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Chechen', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chechen', 'Writing', 'system', 'Cyrillic', 'alphabet', ',', 'Latin',  
'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Chechen', 'Official', 'language', 'in', 'Chechnya', ',', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', ',', 'Type: ANY']], 7, True)

['Chechen', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Chechen', 'Language', 'codes', 'ISO', '639-2', 'che']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chechen', 'Language', 'codes', 'ISO', '639-3', 'che']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chechen', 'Language', 'codes', 'ISO', '639-1', 'ce']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['East', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['East', 'Region', 'East', 'Frisia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['East', 'Language', 'codes', 'ISO', '639-2', 'frs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['East', 'Language', 'codes', 'ISO', '639-3', 'frs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taiwanese', 'Signed', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Taiwanese', 'Total', 'signers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Gullah', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Gullah', 'Language', 'codes', 'ISO', '639-2', 'cpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gullah', 'Language', 'codes', 'ISO', '639-3', 'gul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gullah', 'Language', 'codes', 'Linguasphere', '52-ABB-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Arvanitika', 'Region', 'Attica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arvanitika', 'Region', 'Boeotia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arvanitika', 'Region', 'South', 'Euboea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Arvanitika', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Arvanitika', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Arvanitika', 'Pronunciation', 'a', '\xc9', '\xbe', 'b', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xbe', 'i', '\xca', '\x83', 't']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'a', '\xc9', '\xbe', 'b', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xbe', 'i', '\xca', '\x83', 't']], 16, True)

['Arvanitika', 'Language', 'codes', 'ISO', '639-2', 'alb', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'alb', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Arvanitika', 'Language', 'codes', 'ISO', '639-2', 'sqi', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sqi', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Arvanitika', 'Language', 'codes', 'ISO', '639-3', 'aat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Modern', 'Total', 'speakers', 'Second', 'language', ':', '199', '\xe2', '\x80', '\x93', '600', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Second', 'language', ':', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '199'], ['NUM', '600']], 12, True)

['Modern', 'Spoken', 'in', 'Listed', 'in', 'the', 'article']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Listed', 'in', 'the', 'article']], 7, True)

['Modern', 'Writing', 'system', 'Latin', '(', 'English', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Modern', 'Official', 'language', 'in', '53', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Modern', 'Official', 'language', 'in', 'United', 'Nations']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Modern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', 'eng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-3', 'eng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-1', 'en']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Assamese', 'Region', 'Assam', ',', 'Arunachal', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Assam', ',',  
'Arunachal', 'Pradesh']]], 6, True)

['Assamese', 'Writing', 'system', 'Assamese', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Assamese', 'Official', 'language', 'in', 'India', '(', 'Assam', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'India', '\\\\(', 'Type: ANY', '\\\)']]], 8, True)

['Assamese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Assamese', 'Language', 'codes', 'ISO', '639-2', 'asm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Assamese', 'Language', 'codes', 'ISO', '639-3', 'asm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Assamese', 'Language', 'codes', 'Linguasphere', '59-AAF-w']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)



```

['Assamese', 'Language', 'codes', 'ISO', '639-1', 'as']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Spanish', 'Total', 'speakers', '35', ',', '468', ',', '501', '(',
'United', 'States', 'only', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', ',', 'NUM', '\\(', 'United', 'States', 'only', '\\)']],
['NUM', '35'], ['NUM', '468'], ['NUM', '501']], 13, True)

['Spanish', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Spanish', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Spanish', 'Pronunciation', 'espa', '\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'espa',
'\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol']], 8, True)

['Spanish', 'Language', 'codes', 'ISO', '639-2', 'spa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Spanish', 'Language', 'codes', 'ISO', '639-3', 'spa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Spanish', 'Language', 'codes', 'ISO', '639-1', 'es']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Lydian', 'Spoken', 'in', 'Lydia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Lydian', 'Language', 'codes', 'ISO', '639-3', 'xld']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lycian', 'Spoken', 'in', 'Lycia', ',', 'Lycaonia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)

```

['Lycian', 'Region', 'Southwestern', 'Anatolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lycian', 'Language', 'codes', 'ISO', '639-3', 'xlc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tatar', 'Spoken', 'in', 'Russia', ',', 'other', 'former', 'Soviet', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Russia', ',', 'other', 'former', 'Soviet', 'Union']], 9, True)

['Tatar', 'Official', 'language', 'in', 'Tatarstan', '(', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Tatarstan', '\\(', 'Russia', '\\)']], 8, True)

['Tatar', 'Language', 'codes', 'ISO', '639-2', 'tat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tatar', 'Language', 'codes', 'ISO', '639-3', 'tat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tatar', 'Language', 'codes', 'ISO', '639-1', 'tt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Azerbaijani', 'Total', 'speakers', '20', '\\xe2', '\\x80', '\\x93', '30', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\xe2', '\\x80', '\\x93', 'NUM', 'million']], ['NUM', '20'], ['NUM', '30']], 9, True)

['Azerbaijani', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Uzbekistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Azerbaijani', 'Spoken', 'in', 'Turkmenistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Azerbaijani', 'Pronunciation', 'az', '\xc3', '\xa6', 'rbajd', '\xca', '\x92', 'an', 'dili']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'az', '\xc3', '\xa6', 'rbajd', '\xca', '\x92', 'an', 'dili']]], 10, True)

['Azerbaijani', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-2', 'aze']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', 'aze', '\xe2', '\x80', '\x93', 'Macrolanguage']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'aze', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', 'individual',

```

'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Azerbaijani', 'Language', 'codes', 'ISO', '639-1', 'az']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Alem\\xc3\\xa1n', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Alem\\xc3\\xa1n', 'Total', 'speakers', 'Unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Alem\\xc3\\xa1n', 'Language', 'codes', 'ISO', '639-3', 'gct']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ket', 'Total', 'speakers', '550']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Ket', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Ket', 'Region', 'Krasnoyarsk', 'Krai']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Ket', 'Language', 'codes', 'ISO', '639-3', 'ket']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shelta', 'Spoken', 'in', 'Ireland', ',', 'Irish', 'diaspora']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY', 'diaspora']], 7, True)

['Shelta', 'Writing', 'system', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Shelta', 'Language', 'codes', 'ISO', '639-3', 'sth']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Min', 'Total', 'speakers', '49', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', 'NUM', 'million']], ['NUM', '49']], 6, True)

['Southern', 'Dialects', 'Hokkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Southern', 'Dialects', 'Teochew']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Southern', 'Dialects', 'Hainanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Southern', 'Spoken', 'in', 'Republic', 'of', 'China', '(', 'Taiwan',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'China', '\\\\(', 'Taiwan', '\\\\)']], 9, True)

['Southern', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Regulated', 'by', 'None', '(', 'The', 'Republic', 'of', 'China', 'Ministry', 'of', 'Education', 'and', 'some', 'NGOs', 'are', 'influential', 'in', 'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'None', '\\(', 'The', 'Republic', 'of', 'China', 'Ministry', 'of', 'Education', 'and', 'some', 'NGOs', 'are', 'influential', 'in', 'Taiwan', '\\)']], 20, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'nan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ubykh', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ubykh', 'Region', 'Manyas', ',', 'Bal', '\\xc4', '\\xb1', 'kesir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Manyas', ',', 'Bal', '\\xc4', '\\xb1', 'kesir']], 8, True)

['Ubykh', 'Language', 'codes', 'ISO', '639-3', 'uby']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Westphalian', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Westphalian', 'Spoken', 'in', 'Germany', ',', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Westphalian', 'Region', 'Northwestern', ',', 'Westphalia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Westphalian', 'Language', 'codes', 'ISO', '639-3', 'wep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chagatai', 'Language', 'extinction', '1990s']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chagatai', 'Language', 'codes', 'ISO', '639-2', 'chg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chagatai', 'Language', 'codes', 'ISO', '639-3', 'chg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Proto-Germanic', 'Language', 'extinction', 'Gothic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Proto-Germanic', 'Spoken', 'in', 'Northern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Proto-Germanic', 'Writing', 'system', 'Elder', 'Futhark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Proto-Germanic', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Abkhaz', 'Dialects', 'Abzhywa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Abkhaz', 'Dialects', 'Bzyb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Abkhaz', 'Dialects', 'Sadz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

ANY']]], 3, True)

['Abkhaz', 'Region', 'Caucasus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abkhaz', 'Spoken', 'in', 'Abkhazia', '(', 'Recognized', 'as', 'independent', 'by', 'Russia', ',', 'Nicaragua', ',', 'Nauru', 'and', 'Venezuela', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Abkhazia', '\\(', 'Recognized', 'as', 'independent', 'by', 'Russia', ',', 'Nicaragua', ',', 'Nauru', 'and', 'Venezuela', '\\)']]], 17, True)

['Abkhaz', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abkhaz', 'Spoken', 'in', 'Iran', 'and', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Abkhaz', 'Official', 'language', 'in', 'Abkhazia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Abkhaz', 'Official', 'language', 'in', 'Georgia', '(', 'on', 'the',



'territory', 'of', 'the', 'Autonomous', 'Republic', 'of', 'Abkhazia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Georgia', '\\(', 'on', 'the', 'territory', 'of', 'the',  
'Autonomous', 'Republic', 'of', 'Abkhazia', '\\)']]], 16, True)

['Abkhaz', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Abkhaz', 'Language', 'codes', 'ISO', '639-2', 'abk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Abkhaz', 'Language', 'codes', 'ISO', '639-3', 'abk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Abkhaz', 'Language', 'codes', 'ISO', '639-1', 'ab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Northern', 'Region', 'Northern', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Pronunciation', 'kam', '\\xcb', '\\xa7', 'm', '\\xc9', '\\xaf',  
'\\xcb', '\\x90', '\\xc9', '\\x99', '\\xc5', '\\x8b', '\\xcb', '\\xa7']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kam',  
'\\xcb', '\\xa7', 'm', '\\xc9', '\\xaf', '\\xcb', '\\x90', '\\xc9', '\\x99',  
'\\xc5', '\\x8b', '\\xcb', '\\xa7']]], 16, True)

['Northern', 'Language', 'codes', 'ISO', '639-2', 'tai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'nod']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dyirbal', 'Spoken', 'in', 'Northeast', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Dyirbal', 'Language', 'codes', 'ISO', '639-3', 'dbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cheyenne', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Cheyenne', 'Region', 'Montana', 'and', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Cheyenne', 'Language', 'codes', 'ISO', '639-2', 'chy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cheyenne', 'Language', 'codes', 'ISO', '639-3', 'chy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pirah\xc3\xa3', 'Total', 'speakers', '250-380']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pirah\xc3\xa3', 'Region', 'Amazon', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pirah\xc3\xa3', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pirah\xc3\xa3', 'Pronunciation', '\xca', '\x94', '\xc3', '\xa0', 'p',

'\xc3', '\xa0', 'i', '\xcc', '\xaf', '\xcb', '\x88', 't', '\xcd', '\xa1',  
'\xca', '\x83', '\xc3', '\xae', '\xcb', '\x90', 's', '\xc3', '\xb2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',  
'\x94', '\xc3', '\xa0', 'p', '\xc3', '\xa0', 'i', '\xcc', '\xaf', '\xcb',  
'\x88', 't', '\xcd', '\xa1', '\xca', '\x83', '\xc3', '\xae', '\xcb',  
'\x90', 's', '\xc3', '\xb2']]], 26, True)

['Pirah\xc3\xa3', 'Language', 'codes', 'ISO', '639-3', 'myp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tanacross', 'Region', 'Alaska', '(', 'middle', 'Tanana', 'River', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\  
(', 'middle', 'Tanana', 'River', '\\)']]], 8, True)

['Tanacross', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tanacross', 'Language', 'codes', 'ISO', '639-3', 'tcb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nux\xc3\xa1lk', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nux\xc3\xa1lk', 'Region', 'Bella', 'Coola', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nux\xc3\xa1lk', 'Region', 'Central', 'Coast', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nux\xc3\xa1lk', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nux\xc3\xa1lk', 'Language', 'codes', 'ISO', '639-3', 'blc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Westron', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Westron', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Westron', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Westron', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Westron', 'Category', '(', 'purpose', ')', 'languages', 'of', 'Arda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'languages', 'of', 'Arda']]]], 8, True)

['Westron', 'Category', '(', 'purpose', ')', 'Taliskan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]]], 6, True)

['Westron', 'Category', '(', 'purpose', ')', 'Ad', '\\xc3', '\\xbb', 'naic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Ad', '\\xc3', '\\xbb', 'naic']]]], 9, True)

['Westron', 'Category', '(', 'purpose', ')', 'Westron']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]]], 6, True)

['Westron', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',  
'.', 'R', '.', 'R', '.', 'Tolkien']]]], 10, True)

['Westron', 'Category', '(', 'sources', ')', 'a', 'priori', 'language',  
,',', 'but', 'relative', 'to', 'other', 'languages', 'of', 'Arda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'a', 'priori', 'language', ',,', 'but', 'relative', 'to',  
'other', 'languages', 'of', 'Arda']]]], 16, True)

['Westron', 'Setting', 'and', 'usage', 'Fantasy', 'world', 'of', 'Middle-  
earth']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Fantasy', 'world', 'of', 'Middle-earth']]]], 8, True)

['Westron', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Westron', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Filipino', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Filipino', 'Total', 'speakers', 'First', 'language', ':', '25',  
'million', '.', 'Second', 'language', ':', 'over', '60', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'First', 'language', ':', 'NUM', 'million', '.', 'Second', 'language', ':',  
'over', 'NUM', 'million']], ['NUM', '25'], ['NUM', '60']], 15, True)

['Filipino', 'Total', 'speakers', 'Overall', ':', '90', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Overall', ':', 'NUM', 'million']], ['NUM', '90']], 7, True)

['Filipino', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Filipino', 'Official', 'language', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Filipino', 'Language', 'codes', 'ISO', '639-2', 'fil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Filipino', 'Language', 'codes', 'ISO', '639-3', 'fil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Filipino', 'Language', 'codes', 'ISO', '639-1', 'tl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Votic', 'Total', 'speakers', '20', '(', 'or', 'fewer', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'or', 'fewer', '\\)']], ['NUM', '20']], 8, True)

['Votic', 'Spoken', 'in', 'Russia']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Votic', 'Region', 'Ingria']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Votic', 'Language', 'codes', 'ISO', '639-2', 'vot']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Votic', 'Language', 'codes', 'ISO', '639-3', 'vot']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Middle', 'Region', 'Medieval', 'China']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Middle', 'Writing', 'system', 'Seal', 'Script']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Middle', 'Writing', 'system', 'Clerical', 'Script']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Middle', 'Writing', 'system', 'Kaishu']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Middle', 'Writing', 'system', 'Semi-cursive', 'script']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Middle', 'Writing', 'system', 'Grass', 'script']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Middle', 'Writing', 'system', 'Phagspa']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Middle', 'Writing', 'system', 'Hangul']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

['Middle', 'Official', 'language', 'in', 'Sui', 'Dynasty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Middle', 'Official', 'language', 'in', 'Tang', 'Dynasty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Middle', 'Official', 'language', 'in', 'Song', 'Dynasty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Middle', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'ltc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sona', 'language', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Sona', 'Date', 'founded', '1935']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sona', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sona', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sona', 'Category', '(', 'purpose', ')', 'Sona', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sona', 'Created', 'by', 'Kenneth', 'Searight']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Sona', 'Setting', 'and', 'usage', 'international', 'auxiliary',

'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Sona', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Old', 'Spoken', 'in', 'Latvia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Lithuania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'xcu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nivkh', 'Spoken', 'in', 'Russia', ',', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Nivkh', 'Region', 'Sakhalin', 'Island', ',', 'and', 'along', 'the', 'Amur', 'River']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sakhalin', 'Island', ',', 'and', 'along', 'the', 'Amur', 'River']]], 10, True)

['Nivkh', 'Language', 'codes', 'ISO', '639-3', 'niv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ga', 'Spoken', 'in', 'South-eastern', 'Ghana', ',', 'around', 'Accra']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'South-eastern', 'Ghana', ',', 'around', 'Accra']]], 8, True)

['Ga', 'Pronunciation', '\xc9', '\xa1', '\xc3', '\xa3']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9', '\xa1', '\xc3', '\xa3']]], 6, True)

['Ga', 'Official', 'language', 'in', 'Ghana']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Ga', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Ga', 'Language', 'codes', 'ISO', '639-2', 'gaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Ga', 'Language', 'codes', 'ISO', '639-3', 'gaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tulu', 'Region', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tulu', 'Region', 'Gulf', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tulu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tulu', 'Writing', 'system', 'Tulu', 'script', '(', 'originally', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Tulu', 'Writing', 'system', 'Kannada', 'script', '(', 'presently', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Tulu', 'Language', 'codes', 'ISO', '639-3', 'tcy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chukchi', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chukchi', 'Region', 'Chukotka', 'Autonomous', 'Okrug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY']]], 5, True)

```

['Chukchi', 'Language', 'codes', 'ISO', '639-3', 'ckt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karelian', 'Region', 'Republic', 'of', 'Karelia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Karelian', 'Region', 'Tver', 'Oblast']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karelian', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karelian', 'Spoken', 'in', 'Finland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karelian', 'Official', 'language', 'in', 'recognised', 'as', 'minority',
'language', 'in']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'recognised', 'as', 'minority', 'language', 'in']]], 9, True)

['Karelian', 'Official', 'language', 'in', ':', 'Finland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Karelian', 'Official', 'language', 'in', 'Republic', 'of', 'Karelia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Republic', 'of', 'Karelia']]], 7, True)

['Karelian', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Karelian', 'Language', 'codes', 'ISO', '639-2', 'krl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Karelian', 'Language', 'codes', 'ISO', '639-3', 'krl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

```

['Veps', 'Region', 'Republic', 'of', 'Karelia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Veps', 'Region', 'Vologda', 'Oblast']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Veps', 'Region', 'Veps', 'National', 'Volost']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Veps', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Veps', 'Official', 'language', 'in', 'recognised', 'as', 'minority',
'language', 'in']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'recognised', 'as', 'minority', 'language', 'in']]], 9, True)

['Veps', 'Official', 'language', 'in', ':', 'Republic', 'of', 'Karelia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', ':', 'Republic', 'of', 'Karelia']]], 8, True)

['Veps', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Veps', 'Language', 'codes', 'ISO', '639-3', 'vep']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turkmen', 'Spoken', 'in', 'Turkmenistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Turkmen', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Turkmen', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Turkmen', 'Spoken', 'in', 'Stavropol', 'krai', '(', 'Russia', ')']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Stavropol', 'krai', '\\\\(', 'Russia', '\\\\)']]], 8, True)

['Turkmen', 'Total', 'speakers', 'ca', '.', '4', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c', '.', 'NUM', 'million']], ['NUM', '4']], 7, True)

['Turkmen', 'Official', 'language', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Turkmen', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Turkmen', 'Language', 'codes', 'ISO', '639-2', 'tuk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Turkmen', 'Language', 'codes', 'ISO', '639-3', 'tuk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Turkmen', 'Language', 'codes', 'ISO', '639-1', 'tk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Proven\\xc3\\xa7al', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proven\\xc3\\xa7al', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proven\\xc3\\xa7al', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proven\\xc3\\xa7al', 'Spoken', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proven\\xc3\\xa7al', 'Spoken', 'in', 'small', 'community', 'in', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'small',

'community', 'in', 'California']]], 7, True)

['Proven\xc3\xa7al', 'Region', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Proven\xc3\xa7al', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Proven\xc3\xa7al', 'Language', 'codes', 'ISO', '639-3', 'oci-prv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kodava', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kodava', 'Region', 'Kodagu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kodava', 'Language', 'codes', 'ISO', '639-3', 'kfa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mirandese', 'Spoken', 'in', 'Portugal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mirandese', 'Official', 'language', 'in', 'Co-official', 'recognition', '.', 'Special', 'protection', 'status', 'in', 'Miranda', 'do', 'Douro', ',', 'Portugal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Co-official', 'recognition', '.', 'Special', 'protection', 'status', 'in', 'Miranda', 'do', 'Douro', ',', 'Portugal']]], 16, True)

['Mirandese', 'Regulated', 'by', 'Anstituto', 'de', 'la', 'Lh', '\xc3', '\xa9', 'ngua', 'Mirandesa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Anstituto', 'de', 'la', 'Lh', '\xc3', '\xa9', 'ngua', 'Mirandesa']]], 11, True)

['Mirandese', 'Language', 'codes', 'ISO', '639-2', 'mwl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mirandese', 'Language', 'codes', 'ISO', '639-3', 'mwl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mirandese', 'Language', 'codes', 'Linguasphere', '51-AAA-cb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Modern', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Modern', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Modern', 'Spoken', 'in', 'Greek', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Modern', 'Total', 'speakers', '12', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '12']], 5, True)

['Modern', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Modern', 'Official', 'language', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Modern', 'Official', 'language', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Modern', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Modern', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in', 'parts', 'of']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'recognised', 'as', 'minority', 'language', 'in', 'parts', 'of']]],  
11, True)

['Modern', 'Official', 'language', 'in', ':', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)

['Modern', 'Official', 'language', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Modern', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]]], 6, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', 'gre', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'gre', '\\(', 'B', '\\)']]], ['NUM', '639-2']]], 9, True)

['Modern', 'Language', 'codes', 'ISO', '639-2', 'ell', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ell', '\\(', 'T', '\\)']]], ['NUM', '639-2']]], 9, True)

['Modern', 'Language', 'codes', 'ISO', '639-1', 'el']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']]], 6, True)

['Upper', 'Region', 'Saxony', ',', 'Brandenburg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Upper', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Upper', 'Writing', 'system', 'Latin', '(', 'Sorbian', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]]], 8, True)

['Upper', 'Pronunciation', '\\xcb', '\\x88', 'h', '\\xc9', '\\x94', 'rnj',  
'\\xc9', '\\x94', '\\xcb', '\\x8c', 's', '\\xc9', '\\x9b', 'rb', '\\xca', '\\x83',  
't', '\\xca', '\\x83', 'ina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb',  
'\\x88', 'h', '\\xc9', '\\x94', 'rnj', '\\xc9', '\\x94', '\\xcb', '\\x8c', 's',  
'\\xc9', '\\x9b', 'rb', '\\xca', '\\x83', 't', '\\xca', '\\x83', 'ina']]]], 22,  
True)

['Upper', 'Official', 'language', 'in', 'regional', 'language', 'in', 'Germany', '(', 'Brandenburg', 'and', 'Saxony', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'regional', 'language', 'in', 'Germany', '\\\\(', 'Brandenburg', 'and', 'Saxony', '\\\\)']]], 13, True)

['Upper', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Upper', 'Language', 'codes', 'ISO', '639-2', 'hsb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Upper', 'Language', 'codes', 'ISO', '639-3', 'hsb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lower', 'Region', 'Brandenburg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lower', 'Spoken', 'in', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lower', 'Writing', 'system', 'Latin', '(', 'Sorbian', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Lower', 'Pronunciation', '\\xcb', '\\x88', 'd', '\\xc9', '\\x94', 'ln', '\\xc9', '\\x94', '\\xcb', '\\x8c', 's', '\\xc9', '\\x9b', 'rsk', '\\xca', '\\xb2', 'i']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb', '\\x88', 'd', '\\xc9', '\\x94', 'ln', '\\xc9', '\\x94', '\\xcb', '\\x8c', 's', '\\xc9', '\\x9b', 'rsk', '\\xca', '\\xb2', 'i']]], 19, True)

['Lower', 'Language', 'codes', 'ISO', '639-2', 'dsb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lower', 'Language', 'codes', 'ISO', '639-3', 'dsb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Abaza', 'Spoken', 'in', 'Russia', ',', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Abaza', 'Region', 'Karachay-Cherkess', 'Republic', 'of', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Abaza', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Abaza', 'Language', 'codes', 'ISO', '639-3', 'abq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sundanese', 'Total', 'speakers', '33', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '33']], 5, True)

['Sundanese', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sundanese', 'Region', 'West', 'Java']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sundanese', 'Region', 'Banten']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sundanese', 'Region', 'Jakarta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sundanese', 'Official', 'language', 'in', 'West', 'Java']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Sundanese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Sundanese', 'Language', 'codes', 'ISO', '639-2', 'sun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sundanese', 'Language', 'codes', 'ISO', '639-3', 'sun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sundanese', 'Language', 'codes', 'ISO', '639-1', 'su']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Walloon', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Walloon', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Walloon', 'Spoken', 'in', '(', 'also', ':', 'some', 'speakers', 'in',  
'the', 'United', 'States', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\\\(',  
'also', ':', 'some', 'speakers', 'in', 'the', 'United', 'States', '\\\\)']],  
13, True)

['Walloon', 'Language', 'codes', 'ISO', '639-2', 'wln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Walloon', 'Language', 'codes', 'ISO', '639-3', 'wln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Walloon', 'Language', 'codes', 'Linguasphere', '51-AAA-hf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Walloon', 'Language', 'codes', 'ISO', '639-1', 'wa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Orokolo', 'Region', 'Gulf', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Orokolo', 'Language', 'codes', 'ISO', '639-3', 'oro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tiwa', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tiwa', 'Region', 'Tiwa', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tiwa', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Proto-Esperanto', 'Date', 'founded', '1878-1881']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Proto-Esperanto', 'Language', 'extinction', '1887']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Proto-Esperanto', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Proto-Esperanto', 'Category', '(', 'purpose', ')', 'constructed',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Proto-Esperanto', 'Category', '(', 'purpose', ')', 'international',  
'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'international', 'auxiliary', 'language']]], 8, True)

['Proto-Esperanto', 'Category', '(', 'purpose', ')', 'Proto-Esperanto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)

['Proto-Esperanto', 'Created', 'by', 'L', '.', 'L', '.', 'Zamenhof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'L',  
'.', 'L', '.', 'Zamenhof']]], 8, True)

['Proto-Esperanto', 'Setting', 'and', 'usage', 'international',  
'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Proto-Esperanto', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Proto-Esperanto', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Old', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Southern', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Old', 'Writing', 'system', 'Anglo-Saxon', 'runes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'ofs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pictish', 'Language', 'extinction', 'by', '900', 'AD']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY', 'NUM', 'AD']], ['NUM', '900']], 6, True)

['Pictish', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Pictish', 'Language', 'codes', 'ISO', '639-3', 'xpi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cumbric', 'Spoken', 'in', 'Southern', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cumbric', 'Spoken', 'in', 'Cumberland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cumbric', 'Spoken', 'in', 'Westmorland', 'parts', 'of',  
'Northumberland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Westmorland', 'parts', 'of', 'Northumberland']]], 7, True)

['Cumbric', 'Language', 'codes', 'ISO', '639-3', 'xcb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Camuno', 'Region', 'Italian', 'Peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Camuno', 'Language', 'codes', 'ISO', '639-3', 'xcc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chemakum', 'Spoken', 'in', 'Olympic', 'Peninsula', ',', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Olympic', 'Peninsula', ',', 'Washington']]], 7, True)

['Chemakum', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chemakum', 'Language', 'codes', 'ISO', '639-3', 'cmk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Western', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Egypt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Western', 'Spoken', 'in', 'Uruguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Hungary']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3, True)

['Western', 'Writing', 'system', 'Armenian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'arm', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'arm', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

```
['Western', 'Language', 'codes', 'ISO', '639-2', 'hye', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'hye', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Western', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Western', 'Language', 'codes', 'ISO', '639-1', 'hy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Friulan', 'Region', 'Europe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Friulan', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Friulan', 'Writing', 'system', 'Roman', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Friulan', 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']], 5, True)
```

```
['Friulan', 'Regulated', 'by', 'Osservatori', 'Regjon', '\\xc3', '\\xa2',
'l', 'de', 'Lenghe', 'e', 'de', 'Culture', 'Furlanis']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Osservatori', 'Regjon', '\\xc3', '\\xa2', 'l', 'de', 'Lenghe', 'e', 'de',
'Culture', 'Furlanis']], 14, True)
```

```
['Friulan', 'Language', 'codes', 'ISO', '639-2', 'fur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Friulan', 'Language', 'codes', 'ISO', '639-3', 'fur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Friulan', 'Language', 'codes', 'Linguasphere', '51-AAA-m']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)
```



```
['Swabian', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Swabian', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)
```

```
['Swabian', 'Language', 'codes', 'ISO', '639-3', 'swg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'northern', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Old', 'Spoken', 'in', 'parts', 'of', 'Belgium', '(', 'Wallonia', ')',  
'and', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',  
'of', 'Belgium', '\\\\(', 'Wallonia', '\\\\)', 'and', 'Switzerland']], 11,  
True)
```

```
['Old', 'Spoken', 'in', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Old', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Old', 'Spoken', 'in', 'Kingdom', 'of', 'Sicily']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kingdom', 'of', 'Sicily']], 6, True)
```

```
['Old', 'Spoken', 'in', 'Principality', 'of', 'Antioch']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Principality', 'of', 'Antioch']], 6, True)
```

```
['Old', 'Spoken', 'in', 'Kingdom', 'of', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kingdom', 'of', 'Cyprus']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-2', 'fro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```

['Old', 'Language', 'codes', 'ISO', '639-3', 'fro']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sotho', 'Total', 'speakers', 'at', 'least', '5', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'at', 'least', 'NUM', 'million']], ['NUM', '5']], 7, True)

['Sotho', 'Pronunciation', 's', '\xc9', '\xaa', '\xcc', '\x80', 's',
'\xca', '\x8a', '\xcc', '\x81', 't', '\xca', '\xb0', '\xca', '\x8a',
'\xcc', '\x80']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's',
'\xc9', '\xaa', '\xcc', '\x80', 's', '\xca', '\x8a', '\xcc', '\x81', 't',
'\xca', '\xb0', '\xca', '\x8a', '\xcc', '\x80']]], 19, True)

['Sotho', 'Spoken', 'in', 'Lesotho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sotho', 'Spoken', 'in', 'South', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Sotho', 'Official', 'language', 'in', 'Lesotho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Sotho', 'Official', 'language', 'in', 'South', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Sotho', 'Language', 'codes', 'ISO', '639-2', 'sot']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sotho', 'Language', 'codes', 'ISO', '639-3', 'sot']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sotho', 'Language', 'codes', 'ISO', '639-1', 'st']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sindhi', 'Total', 'speakers', '35', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

```

'NUM', 'million']], ['NUM', '35']], 5, True)

['Sindhi', 'Dialects', 'Indo-European']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Indo-Iranian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Indo-Aryan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Northwestern', 'Zone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sindhi', 'Dialects', 'Sindhi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'India', '.', 'Also', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'India', '.', 'Also', 'Hong', 'Kong']], 7, True)

['Sindhi', 'Dialects', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sindhi', 'Dialects', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Dialects', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Dialects', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Dialects', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sindhi', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'India', '.', 'Also', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', '.', 'Also', 'Hong', 'Kong']]], 8, True)

['Sindhi', 'Spoken', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'UAE']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'UK']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Spoken', 'in', 'Sri', 'Lanka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sindhi', 'Writing', 'system', 'Arabic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Writing', 'system', 'Devanagari']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sindhi', 'Writing', 'system', 'La', '\xe1', '\xb9', '\x87', '\xe1', '\xb8', '\x8d', '\xc4', '\x81', 'scripts']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'La', '\xe1', '\xb9', '\x87', '\xe1', '\xb8', '\x8d', '\xc4', '\x81', 'scripts']]], 13, True)

['Sindhi', 'Writing', 'system', 'particularly', 'Gurumukhi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sindhi', 'Official', 'language', 'in', 'Pakistan', '(', 'Sindh', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Pakistan', '\\(', 'Sindh', '\\)']]], 8, True)

['Sindhi', 'Official', 'language', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Sindhi', 'Regulated', 'by', 'Indian', 'Institute', 'of', 'Sindhology', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Indian', 'Institute', 'of', 'Sindhology', '\\(', 'India', '\\)']], 10, True)

['Sindhi', 'Language', 'codes', 'ISO', '639-2', 'snd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sindhi', 'Language', 'codes', 'ISO', '639-3', 'snd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sindhi', 'Language', 'codes', 'ISO', '639-1', 'sd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Jamamad\\xc3\\xad', 'Spoken', 'in', 'Amazonas', 'State', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

['Jamamad\\xc3\\xad', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']], 6, True)

['Jamamad\\xc3\\xad', 'Dialects', 'Jarawara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Jamamad\\xc3\\xad', 'Dialects', 'Baniwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Jamamad\\xc3\\xad', 'Dialects', 'Kanamanti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Jamamad\\xc3\\xad', 'Dialects', 'etc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Jamamad\\xc3\\xad', 'Language', 'codes', 'ISO', '639-3', 'jaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tup\\xc3\\xad', 'Spoken', 'in', 'Brazil', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Tup\xc3\xad', 'Spoken', 'in', 'Tupinamb', '\xc3', '\xa1']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Tupinamb', '\xc3', '\xa1']]]], 6, True)

['Tup\xc3\xad', 'Dialects', 'L', '\xc3', '\xad', 'ngua', 'geral', 'amaz',  
'\xc3', '\xb4', 'nica']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'L', '\xc3',  
'\xad', 'ngua', 'geral', 'amaz', '\xc3', '\xb4', 'nica']]]], 11, True)

['Tup\xc3\xad', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Tup\xc3\xad', 'Language', 'codes', 'ISO', '639-2', 'sai', '(',  
'Collective', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sai', '\\(', 'Collective', '\\)']]], ['NUM', '639-2']], 9,  
True)

['Tup\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tup\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'tpn', '\xe2', '\x80',  
'\x93', 'Tupinamb', '\xc3', '\xa1']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tpn', '\xe2', '\x80', '\x93', 'Tupinamb', '\xc3', '\xa1']]],  
['NUM', '639-3']], 12, True)

['Rwanda', 'Spoken', 'in', 'Rwanda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Rwanda', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Rwanda', 'Spoken', 'in', 'DR', 'Congo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Rwanda', 'Total', 'speakers', '12', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']]], ['NUM', '12']], 5, True)

```
['Rwanda', 'Official', 'language', 'in', 'Rwanda']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Rwanda', 'Regulated', 'by', 'No', 'official', 'regulation']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Rwanda', 'Language', 'codes', 'ISO', '639-2', 'kin']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Rwanda', 'Language', 'codes', 'ISO', '639-3', 'kin']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Rwanda', 'Language', 'codes', 'ISO', '639-1', 'rw']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Wichita', 'Total', 'speakers', '1', 'fluent', ',', '4', 'conversant']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'fluent', ',', 'NUM', 'conversant']]], ['NUM', '1'], ['NUM', '4']],
8, True)
```

```
['Wichita', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Wichita', 'Region', 'West-central', 'Oklahoma']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Wichita', 'Language', 'codes', 'ISO', '639-3', 'wic']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tajik', 'Spoken', 'in', 'Tajikistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tajik', 'Spoken', 'in', 'Uzbekistan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```



['Tajik', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tajik', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tajik', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tajik', 'Writing', 'system', 'Perso-Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tajik', 'Official', 'language', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Tajik', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tajik', 'Language', 'codes', 'ISO', '639-2', 'tgk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tajik', 'Language', 'codes', 'ISO', '639-3', 'tgk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tajik', 'Language', 'codes', 'ISO', '639-1', 'tg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Uyghur', 'Total', 'speakers', '8-11', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '8-11']], 5, True)

['Uyghur', 'Region', 'Xinjiang', 'Uyghur', 'Autonomous']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Uyghur', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Kazakhstan', ';', 'also', 'spoken', 'in', 'Afghanistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kazakhstan', ';', 'also', 'spoken', 'in', 'Afghanistan']]], 9, True)

['Uyghur', 'Spoken', 'in', 'Albania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Belgium']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Kyrgyzstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Mongolia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uyghur', 'Spoken', 'in', 'Saudi', 'Arabia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Uyghur', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uyghur', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Uyghur', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Uyghur', 'Spoken', 'in', 'and', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Uyghur', 'Writing', 'system', 'Uyghur', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Uyghur', 'Pronunciation', '\xca', '\x94', '\xca', '\x8a', 'j', '\xcb', '\x88', '\xca', '\x81', '\xca', '\x8a', 'rt', '\xca', '\x83', '\xc9', '\x9b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca', '\x94', '\xca', '\x8a', 'j', '\xcb', '\x88', '\xca', '\x81', '\xca', '\x8a', 'rt', '\xca', '\x83', '\xc9', '\x9b']], 18, True)

['Uyghur', 'Language', 'codes', 'ISO', '639-2', 'uig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Uyghur', 'Language', 'codes', 'ISO', '639-3', 'uig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uyghur', 'Language', 'codes', 'ISO', '639-1', 'ug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Oromo', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Oromo', 'Spoken', 'in', 'Kenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Oromo', 'Spoken', 'in', 'Somalia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Oromo', 'Language', 'codes', 'ISO', '639-2', 'orm']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', 'orm', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'orm', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', ':', 'gax', '\xe2', '\x80',  
'\x93', 'Borana-Arsi-Guji-Walagaa', 'Oromo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'gax', '\xe2', '\x80', '\x93', 'Borana-Arsi-Guji-  
Walagaa', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Oromo', 'Language', 'codes', 'ISO', '639-3', 'orc', '\xe2', '\x80',  
'\x93', 'Orma']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'orc', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Oromo', 'Language', 'codes', 'ISO', '639-1', 'om']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Oriya', 'Region', 'Orissa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Oriya', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oriya', 'Writing', 'system', 'Oriya', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Oriya', 'Official', 'language', 'in', 'Orissa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Oriya', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-2', 'ori']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-3', 'ori']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oriya', 'Language', 'codes', 'Linguasphere', '59-AAF-x']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Oriya', 'Language', 'codes', 'ISO', '639-1', 'or']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Shona', 'Spoken', 'in', 'Zimbabwe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shona', 'Spoken', 'in', 'Mozambique']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Shona', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Shona', 'Region', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Shona', 'Language', 'codes', 'ISO', '639-2', 'sna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Shona', 'Language', 'codes', 'ISO', '639-3', 'sna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shona', 'Language', 'codes', 'ISO', '639-1', 'sn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Aleut', 'Spoken', 'in', 'Alaska', '(', 'Aleutian', 'and', 'Pribilof',  
'Islands', ')', ',', 'Kamchatka', 'Krai', '(', 'Commander', 'Islands',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Alaska',  
'\\(', 'Aleutian', 'and', 'Pribilof', 'Islands', '\\)', ',', 'Kamchatka',  
'Krai', '\\(', 'Commander', 'Islands', '\\)']]], 17, True)
```

```
['Aleut', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aleut', 'Writing', 'system', 'Cyrillic', 'alphabet', '(', 'historic',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Aleut', 'Language', 'codes', 'ISO', '639-2', 'ale']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Aleut', 'Language', 'codes', 'ISO', '639-3', 'ale']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Hindko', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Hindko', 'Spoken', 'in', 'Pakistan', '(', 'Khyber', 'Pakhtunkhwa', ',', 'Punjab', ',', 'Pakistan', 'Administered', 'Kashmir', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Pakistan', '\\\\(', 'Khyber', 'Pakhtunkhwa', ',', 'Punjab', ',', 'Pakistan', 'Administered', 'Kashmir', '\\\\)']]], 14, True)

['Hindko', 'Spoken', 'in', 'Afghanistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hindko', 'Writing', 'system', 'Nasta'liq', 'script', ',', 'Devanagari', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Hindko', 'Pronunciation', 'hindkou', '\\xcc', '\\xaf']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'hindkou', '\\xcc', '\\xaf']]], 5, True)

['Hindko', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hindko', 'Language', 'codes', 'ISO', '639-3', ':', 'hnd', '\\xe2', '\\x80', '\\x93', 'Southern']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'hnd', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Hindko', 'Language', 'codes', 'ISO', '639-3', 'hno', '\\xe2', '\\x80', '\\x93', 'Northern']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hno', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Lower', 'Spoken', 'in', 'Poland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lower', 'Spoken', 'in', 'Czech', 'Republic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lower', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lower', 'Language', 'codes', 'ISO', '639-3', 'sli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Potiguara', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Potiguara', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Potiguara', 'Language', 'codes', 'ISO', '639-3', 'pog']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mari', 'Spoken', 'in', 'Bashkortostan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Tatarstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Udmurtia', ';', 'oblasti', 'Nizhny', 'Novgorod']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Udmurtia', ';', 'oblasti', 'Nizhny', 'Novgorod']]], 8, True)

['Mari', 'Spoken', 'in', 'Kirov']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Sverdlovsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mari', 'Spoken', 'in', 'Orenburg', ';', 'Perm', 'Krai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Orenburg', ';', 'Perm', 'Krai']]], 7, True)

['Mari', 'Official', 'language', 'in', 'Mari', 'El']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)



```
['Mari', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Mari', 'Language', 'codes', 'ISO', '639-2', 'chm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Mari', 'Language', 'codes', 'ISO', '639-3', 'chm', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chm', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM',  
'639-3']], 10, True)
```

```
['Mari', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)
```

```
['Yerkish', 'Total', 'speakers', '3', '(', 'ape', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'ape', '\\)']]], ['NUM', '3']], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'Animal', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Yerkish', 'Category', '(', 'purpose', ')', 'Yerkish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)
```

```
['Yerkish', 'Created', 'by', 'Ernst', 'von', 'Glasersfeld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Ernst',  
'von', 'Glasersfeld']]], 6, True)
```

```
['Yerkish', 'Category', '(', 'sources', ')', 'Use', 'a', 'keyboard', 'to',  
'punch', 'keys', 'with', 'so', 'called', 'lexigrams', ',', 'symbols',  
'corresponding', 'to', 'objects', 'or', 'ideas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Use', 'a', 'keyboard', 'to', 'punch', 'keys', 'with',  
'so', 'called', 'lexigrams', ',', 'symbols', 'corresponding', 'to',  
'objects', 'or', 'ideas']]], 22, True)
```

```
['Yerkish', 'Setting', 'and', 'usage', 'Animal', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Type: ANY', 'language']]], 6, True)
```

```
['Yerkish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Sassarese', 'Region', 'Sardinia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sassarese', 'Spoken', 'in', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sassarese', 'Pronunciation', 'sas', '\xcb', '\x90', 'a', '\xcb', '\x88',  
'rezu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sas',  
'\xcb', '\x90', 'a', '\xcb', '\x88', 'rezu']]], 9, True)
```

```
['Sassarese', 'Official', 'language', 'in', 'Sardinia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Sassarese', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Sassarese', 'Language', 'codes', 'ISO', '639-2', 'srd', 'or', 'cos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'srd', 'or', 'cos']], ['NUM', '639-2']], 8, True)
```

```
['Sassarese', 'Language', 'codes', 'ISO', '639-3', 'sdc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sassarese', 'Language', 'codes', 'Linguasphere', '51-AAA-pe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Sassarese', 'Language', 'codes', 'ISO', '639-1', 'sc', 'or', 'co']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sc', 'or', 'co']], ['NUM', '639-1']], 8, True)
```

```
['Hittite', 'Spoken', 'in', 'Anatolia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Hittite', 'Language', 'codes', 'ISO', '639-2', 'hit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Hittite', 'Language', 'codes', 'ISO', '639-3', 'hit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Istriot', 'Spoken', 'in', 'Istria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Istriot', 'Language', 'codes', 'ISO', '639-3', 'ist']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Istriot', 'Language', 'codes', 'Linguasphere', '51-AAA-na']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Cebuano', 'Total', 'speakers', 'About', '20', 'million', 'people', 'and',  
'the', 'second', 'most', 'spoken', 'language', 'in', 'the', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'About', 'NUM', 'million', 'people', 'and', 'the', 'second', 'most',  
'spoken', 'language', 'in', 'the', 'Philippines']]], ['NUM', '20']], 16,  
True)
```

```
['Cebuano', 'Dialects', 'Cebuano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cebuano', 'Dialects', 'Boholano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cebuano', 'Dialects', 'Leytehanon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Cebuano', 'Dialects', 'Mindanao', 'Visayan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Cebuano', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cebuano', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cebuano', 'Writing', 'system', 'Latin', '(', 'Cebuano', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Cebuano', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Cebuano', 'Regulated', 'by', 'Visayan', 'Academy', 'of', 'Arts', 'and', 'Letters']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Visayan', 'Academy', 'of', 'Arts', 'and', 'Letters']], 9, True)

['Cebuano', 'Language', 'codes', 'ISO', '639-2', 'ceb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cebuano', 'Language', 'codes', 'ISO', '639-3', 'ceb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huilliche', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Huilliche', 'Language', 'codes', 'ISO', '639-3', 'huh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mam', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mam', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Mam', 'Region', 'Quetzaltenango', ',', '']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Mam', 'Region', 'Huehuetenango', ',', '']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Mam', 'Region', 'San', 'Marcos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Mam', 'Region', 'Chiapas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Mam', 'Language', 'codes', 'ISO', '639-3', 'mam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lingua', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'purpose', '\\\\)', 'Type: ANY', 'language']], 7, True)
```

```
['Lingua', 'Created', 'by', 'C', '.', 'George', 'Boeree']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'C',
'.', 'George', 'Boeree']], 7, True)
```

```
['Lingua', 'Category', '(', 'sources', ')', 'based', 'on', 'Romance',
'and', 'Creole', 'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
'sources', '\\\\)', 'based', 'on', 'Romance', 'and', 'Creole',
'languages']], 11, True)
```

```
['Lingua', 'Language', 'codes', 'ISO', '639-3', 'lfn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Silesian', 'Total', 'speakers', 'no', 'data']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'no', 'data']], 5, True)
```

```
['Silesian', 'Language', 'codes', 'ISO', '639-3', 'szl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Silesian', 'Language', 'codes', 'Linguasphere', '53-AAA-cck']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Cherokee', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cherokee', 'Writing', 'system', 'Cherokee', 'syllabary', ',', 'Latin',  
'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Cherokee', 'Language', 'codes', 'ISO', '639-2', 'chr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cherokee', 'Language', 'codes', 'ISO', '639-3', 'chr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yor\x3\xb9b\x3\xa1', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yor\x3\xb9b\x3\xa1', 'Spoken', 'in', 'Togo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yor\x3\xb9b\x3\xa1', 'Spoken', 'in', 'Benin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yor\x3\xb9b\x3\xa1', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Yor\x3\xb9b\x3\xa1', 'Official', 'language', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Yor\x3\xb9b\x3\xa1', 'Regulated', 'by', 'No', 'official',  
'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Yor\xc3\xb9b\xc3\xa1', 'Language', 'codes', 'ISO', '639-2', 'yor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Yor\xc3\xb9b\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'yor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yor\xc3\xb9b\xc3\xa1', 'Language', 'codes', 'ISO', '639-1', 'yo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Chinantec', 'Total', 'speakers', 'approx', '.', '93', ',', '000', '(',  
'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'approx', '.', 'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM',  
'93'], ['NUM', '000']], 11, True)

['Chinantec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chinantec', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', ':', 'cco', '\xe2',  
'\x80', '\x93', 'Comaltepec', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'cco', '\xe2', '\x80', '\x93', 'Comaltepec', 'Type:  
ANY']], ['NUM', '639-3']], 12, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'chj', '\xe2', '\x80',  
'\x93', 'Ojitl', '\xc3', '\xa1', 'n', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chj', '\xe2', '\x80', '\x93', 'Ojitl', '\xc3', '\xa1', 'n',  
'Type: ANY']], ['NUM', '639-3']], 14, True)

['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'chq', '\xe2', '\x80',  
'\x93', 'Quiotepec', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```
'ISO', 'NUM', 'chq', '\xe2', '\x80', '\x93', 'Quiotepec', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'chz', '\xe2', '\x80',  
\x93', 'Ozumac', '\xc3', '\xad', 'n', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chz', '\xe2', '\x80', '\x93', 'Ozumac', '\xc3', '\xad', 'n',  
'Type: ANY']], ['NUM', '639-3']], 14, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cle', '\xe2', '\x80',  
\x93', 'Lealao', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cle', '\xe2', '\x80', '\x93', 'Lealao', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cnl', '\xe2', '\x80',  
\x93', 'Lalana', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cnl', '\xe2', '\x80', '\x93', 'Lalana', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cnt', '\xe2', '\x80',  
\x93', 'Tepetotutla', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cnt', '\xe2', '\x80', '\x93', 'Tepetotutla', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cpa', '\xe2', '\x80',  
\x93', 'Palantla', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cpa', '\xe2', '\x80', '\x93', 'Palantla', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'csa', '\xe2', '\x80',  
\x93', 'Chiltepec', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'csa', '\xe2', '\x80', '\x93', 'Chiltepec', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cso', '\xe2', '\x80',  
\x93', 'Sochiapan', 'Chinantec']:  
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cso', '\xe2', '\x80', '\x93', 'Sochiapan', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cte', '\xe2', '\x80',
```



```
'\x93', 'Tepinapa', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cte', '\xe2', '\x80', '\x93', 'Tepinapa', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'ctl', '\xe2', '\x80',  
'\x93', 'Tlacoatzintepec', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ctl', '\xe2', '\x80', '\x93', 'Tlacoatzintepec', 'Type:  
ANY']], ['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cuc', '\xe2', '\x80',  
'\x93', 'Usila', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cuc', '\xe2', '\x80', '\x93', 'Usila', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Chinantec', 'Language', 'codes', 'ISO', '639-3', 'cvn', '\xe2', '\x80',  
'\x93', 'Valle', 'Nacional', 'Chinantec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cvn', '\xe2', '\x80', '\x93', 'Valle', 'Nacional', 'Type:  
ANY']], ['NUM', '639-3']], 12, True)
```

```
['Baoule', 'Language', 'codes', 'ISO', '639-3', 'bci']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['J\xc3\xa8rriais', 'Spoken', 'in', 'Jersey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['J\xc3\xa8rriais', 'Spoken', 'in', 'Sark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['J\xc3\xa8rriais', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['J\xc3\xa8rriais', 'Language', 'codes', 'Linguasphere', '51-AAA-hc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)
```

```
['Papiamento', 'Spoken', 'in', 'Aruba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Papiamento', 'Spoken', 'in', 'Cura', '\xc3', '\xa7', 'ao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Cura',  
'\xc3', '\xa7', 'ao']]], 7, True)

['Papiamento', 'Spoken', 'in', 'Caribbean', 'Netherlands', '(', 'Bonaire',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Caribbean', 'Netherlands', '\\(', 'Bonaire', '\\)']]], 8, True)

['Papiamento', 'Region', 'Caribbean', 'islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Papiamento', 'Official', 'language', 'in', 'Aruba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Papiamento', 'Official', 'language', 'in', 'Cura', '\xc3', '\xa7', 'ao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Cura', '\xc3', '\xa7', 'ao']]], 8, True)

['Papiamento', 'Official', 'language', 'in', 'Caribbean', 'Netherlands',  
'(', 'Bonaire', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Caribbean', 'Netherlands', '\\(', 'Bonaire', '\\)']]], 9, True)

['Papiamento', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Papiamento', 'Language', 'codes', 'ISO', '639-2', 'pap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Papiamento', 'Language', 'codes', 'ISO', '639-3', 'pap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papiamento', 'Language', 'codes', 'Linguasphere', '51-AAC-be']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Taa', 'Spoken', 'in', 'Botswana', ',', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', ',', 'Type: ANY']]], 6, True)

['Taa', 'Region', 'Southern', 'Ghanzi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taa', 'Region', 'northern', 'Kgalagadi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taa', 'Language', 'codes', 'ISO', '639-3', 'nmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vedic', 'Spoken', 'in', 'Bronze', 'Age', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Bronze', 'Age', 'India']]], 6, True)

['Vedic', 'Spoken', 'in', 'Iron', 'Age', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iron', 'Age', 'India']]], 6, True)

['Vedic', 'Spoken', 'in', 'Greater', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Vedic', 'Language', 'codes', 'ISO', '639-3', 'san']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balinese', 'Spoken', 'in', 'Bali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balinese', 'Spoken', 'in', 'Nusa', 'Penida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Balinese', 'Spoken', 'in', 'Lombok', 'and', 'Java']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Balinese', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Balinese', 'Writing', 'system', 'Latin', ',', 'Balinese']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Balinese', 'Language', 'codes', 'ISO', '639-2', 'ban']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Balinese', 'Language', 'codes', 'ISO', '639-3', 'ban']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Buginese', 'Spoken', 'in', 'Certain', 'Parts', 'of', 'Sumatra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Certain', 'Parts', 'of', 'Sumatra']]], 7, True)
```

```
['Buginese', 'Spoken', 'in', 'Riau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Buginese', 'Spoken', 'in', 'Kalimantan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Buginese', 'Spoken', 'in', 'Sabah']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Buginese', 'Spoken', 'in', 'Malay', 'peninsula']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Buginese', 'Writing', 'system', 'Lontara', ',', '']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', ',', '']]], 5, True)
```

```
['Buginese', 'Writing', 'system', 'Roman', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Buginese', 'Language', 'codes', 'ISO', '639-2', 'bug']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Buginese', 'Language', 'codes', 'ISO', '639-3', 'bug']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gayo', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gayo', 'Region', 'Sumatra']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Gayo', 'Language', 'codes', 'ISO', '639-2', 'gay']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gayo', 'Language', 'codes', 'ISO', '639-3', 'gay']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iban', 'Spoken', 'in', 'Malaysia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iban', 'Region', 'Borneo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Iban', 'Language', 'codes', 'ISO', '639-2', 'iba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Iban', 'Language', 'codes', 'ISO', '639-3', 'iba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Javanese', 'Spoken', 'in', 'Java', '(', 'Indonesia', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']], 7, True)

['Javanese', 'Spoken', 'in', 'Peninsular', 'Malaysia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Javanese', 'Spoken', 'in', 'Suriname']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Javanese', 'Spoken', 'in', 'New', 'Caledonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Javanese', 'Total', 'speakers', 'about', '80', 'million', 'total']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'about', 'NUM', 'million', 'total']], ['NUM', '80']], 7, True)
```

```
['Javanese', 'Writing', 'system', 'Javanese', 'script', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Javanese', 'Writing', 'system', 'Arabic', 'script', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Javanese', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-2', 'jav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', ':', 'jav', '\xe2',  
'\x80', '\x93', 'Javanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'jav', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 11, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', 'jvn', '\xe2', '\x80',  
'\x93', 'Caribbean', 'Javanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'jvn', '\xe2', '\x80', '\x93', 'Caribbean', 'Type: ANY']]],  
['NUM', '639-3']], 11, True)
```

```
['Javanese', 'Language', 'codes', 'ISO', '639-3', 'tes', '\xe2', '\x80',  
'\x93', 'Tenggerese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'tes', '\xe2', '\x80', '\x93', 'Tenggerese']]], ['NUM',  
'639-3']], 10, True)
```

['Javanese', 'Language', 'codes', 'ISO', '639-1', 'jv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Madurese', 'Spoken', 'in', 'Island', 'of', 'Madura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Island',  
'of', 'Madura']]], 6, True)

['Madurese', 'Spoken', 'in', 'Sapudi', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Madurese', 'Spoken', 'in', 'northern', 'coastal', 'area', 'of',  
'eastern', 'Java']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'northern', 'coastal', 'area', 'of', 'eastern', 'Java']]], 9, True)

['Madurese', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Madurese', 'Language', 'codes', 'ISO', '639-2', 'mad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Madurese', 'Language', 'codes', 'ISO', '639-3', 'mad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Makassar', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Makassar', 'Writing', 'system', 'Lontara', ',', 'Roman', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', ',', 'Type: ANY', 'alphabet']]], 7, True)

['Makassar', 'Language', 'codes', 'ISO', '639-2', 'mak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Makassar', 'Language', 'codes', 'ISO', '639-3', 'mak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minangkabau', 'Spoken', 'in', 'Indonesia', ',', 'Malaysia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Minangkabau', 'Region', 'West', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Minangkabau', 'Region', 'Riau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Minangkabau', 'Region', 'Jambi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Minangkabau', 'Region', 'Bengkulu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Minangkabau', 'Region', 'North', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Minangkabau', 'Region', 'Aceh', '(', 'Indonesia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Aceh', '\\(', 'Indonesia', '\\)']]], 6, True)

['Minangkabau', 'Language', 'codes', 'ISO', '639-2', 'min']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Minangkabau', 'Language', 'codes', 'ISO', '639-3', 'min']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nicaraguan', 'Region', 'The', 'Managua', 'region', ',', 'but', 'spreading', 'throughout', 'the', 'country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'The', 'Managua', 'region', ',', 'but', 'spreading', 'throughout', 'the', 'country']]], 11, True)

['Nicaraguan', 'Signed', 'in', 'Nicaragua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bashkir', 'Spoken', 'in', 'Russia', ',', 'Kazakhstan']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Bashkir', 'Total', 'speakers', 'approx', '.', '1', ',', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'approx', '.', 'NUM', ',', 'NUM', 'million']], ['NUM', '1'], ['NUM', '5']], 9, True)

['Bashkir', 'Official', 'language', 'in', 'Bashkortostan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Bashkir', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Bashkir', 'Language', 'codes', 'ISO', '639-2', 'bak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bashkir', 'Language', 'codes', 'ISO', '639-3', 'bak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bashkir', 'Language', 'codes', 'ISO', '639-1', 'ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Fijian', 'Total', 'speakers', '450', ',', '000', 'native', 'speakers', ',', '200', ',', '000', 'second-language', 'users']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'native', 'speakers', ',', 'NUM', ',', 'NUM', 'second-language', 'users']], ['NUM', '450'], ['NUM', '000'], ['NUM', '200'], ['NUM', '000']], 14, True)

['Fijian', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fijian', 'Region', 'Spoken', 'as', 'first', 'language', 'on', 'Vanua', 'Levu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Spoken', 'as', 'first', 'language', 'on', 'Vanua', 'Levu']]], 9, True)

['Fijian', 'Region', 'the', 'eastern', 'half', 'of', 'Viti', 'Levu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'the',

'eastern', 'half', 'of', 'Viti', 'Levu']]], 8, True)

['Fijian', 'Region', 'and', 'on', 'the', 'lesser', 'islands', 'of', 'Kadavu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'on', 'the', 'lesser', 'islands', 'of', 'Kadavu']]], 9, True)

['Fijian', 'Region', 'Nayau']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Lakeba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Oneata']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Moce']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Komo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Namuka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Kabara']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Region', 'Vulaga']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Fijian', 'Official', 'language', 'in', 'Fiji']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Fijian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

```

['Fijian', 'Language', 'codes', 'ISO', '639-2', 'fij']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Fijian', 'Language', 'codes', 'ISO', '639-3', 'fij']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fijian', 'Language', 'codes', 'ISO', '639-1', 'fj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Kashmiri', 'Spoken', 'in', 'Jammu', 'and', 'Kashmir', '(', 'India', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Jammu',
'and', 'Kashmir', '\\(', 'India', '\\)']], 9, True)

['Kashmiri', 'Spoken', 'in', 'Azad', 'Kashmir', '(', 'Pakistan', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Azad',
'Kashmir', '\\(', 'Pakistan', '\\)']], 8, True)

['Kashmiri', 'Writing', 'system', 'Perso-Arabic', 'script', '(',
'contemporary', ')', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Perso-Arabic', 'script', '\\(', 'contemporary', '\\)', ',']], 9, True)

['Kashmiri', 'Writing', 'system', 'Devanagari', 'script', '(',
'contemporary', ')', ',']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Devanagari', 'script', '\\(', 'contemporary', '\\)', ',']], 9, True)

['Kashmiri', 'Writing', 'system', 'Sharada', 'script', '(', 'ancient', '/',
'liturgical', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Sharada', 'script', '\\(', 'ancient', '/', 'liturgical', '\\)']], 10,
True)

['Kashmiri', 'Official', 'language', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']], 5, True)

['Kashmiri', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']], 6, True)

['Kashmiri', 'Language', 'codes', 'ISO', '639-2', 'kas']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kashmiri', 'Language', 'codes', 'ISO', '639-3', 'kas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kashmiri', 'Language', 'codes', 'ISO', '639-1', 'ks']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Rundi', 'Spoken', 'in', 'Burundi', ',', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Rundi', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Rundi', 'Official', 'language', 'in', 'Burundi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Rundi', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Rundi', 'Language', 'codes', 'ISO', '639-2', 'run']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Rundi', 'Language', 'codes', 'ISO', '639-3', 'run']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rundi', 'Language', 'codes', 'ISO', '639-1', 'rn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Latino', 'sine', 'flexione', 'Total', 'speakers', 'virtually', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'sine', 'flexione', 'Total', 'speakers', 'virtually', 'extinct']], 7, True)

['Latino', 'Date', 'founded', '1903']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Latino', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Latino', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Latino', 'Created', 'by', 'Giuseppe', 'Peano']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Latino', 'Setting', 'and', 'usage', 'international', 'auxiliary',  
'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Latino', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Ligurian', 'Spoken', 'in', 'Liguria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ligurian', 'Language', 'codes', 'ISO', '639-3', 'xlg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sango', 'Region', 'Central', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sango', 'Spoken', 'in', 'Central', 'African', 'Republic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', 'Republic']]], 6, True)

['Sango', 'Spoken', 'in', 'Chad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sango', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Sango', 'Pronunciation', 'j', '\xc3', '\xa1', '\xc5', '\x8b', '\xc9', '\xa1', '\xc3', '\xa1', 't', '\xc3', '\xad', 's', '\xc4', '\x81', '\xc5', '\x8b', '\xc9', '\xa1', '\xc5', '\x8d']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'j', '\xc3', '\xa1', '\xc5', '\x8b', '\xc9', '\xa1', '\xc3', '\xa1', 't', '\xc3', '\xad', 's', '\xc4', '\x81', '\xc5', '\x8b', '\xc9', '\xa1', '\xc5', '\x8d']]], 23, True)

['Sango', 'Regulated', 'by', 'Institute', 'of', 'Applied', 'Linguistics']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Institute', 'of', 'Applied', 'Linguistics']]], 7, True)

['Sango', 'Language', 'codes', 'ISO', '639-2', 'sag', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sag', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Sango', 'Language', 'codes', 'ISO', '639-2', 'sag', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sag', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Sango', 'Language', 'codes', 'ISO', '639-3', 'sag']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sango', 'Language', 'codes', 'ISO', '639-1', 'sg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Swazi', 'Spoken', 'in', 'Swaziland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swazi', 'Spoken', 'in', 'South', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Swazi', 'Spoken', 'in', 'Lesotho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swazi', 'Spoken', 'in', 'Mozambique']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Swazi', 'Total', 'speakers', '3', ',', '000', ',', '000', '(',

'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']], ['NUM', '3'],  
['NUM', '000'], ['NUM', '000']], 11, True)

['Swazi', 'Official', 'language', 'in', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Swazi', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Swazi', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Swazi', 'Language', 'codes', 'ISO', '639-2', 'ssw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Swazi', 'Language', 'codes', 'ISO', '639-3', 'ssw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Swazi', 'Language', 'codes', 'ISO', '639-1', 'ss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Tongan', 'Spoken', 'in', 'Tonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'American', 'Samoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tongan', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'Fiji']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tongan', 'Spoken', 'in', 'Niue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tongan', 'Official', 'language', 'in', 'Tonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Tongan', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tongan', 'Language', 'codes', 'ISO', '639-2', 'ton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tongan', 'Language', 'codes', 'ISO', '639-3', 'ton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tongan', 'Language', 'codes', 'ISO', '639-1', 'to']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tsonga', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsonga', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)



['Tsonga', 'Spoken', 'in', 'Swaziland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsonga', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsonga', 'Region', 'Limpopo', ',', 'Mpumalanga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tsonga', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Tsonga', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Tsonga', 'Language', 'codes', 'ISO', '639-2', 'tso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tsonga', 'Language', 'codes', 'ISO', '639-3', 'tso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsonga', 'Language', 'codes', 'ISO', '639-1', 'ts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Zhuang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zhuang', 'Total', 'speakers', '14', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '14']], 5, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-2', 'zha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zha', '\xe2', '\x80',

```

'\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zha', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zhd', '\xe2', '\x80',
'\x93', 'Dai', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zhd', '\xe2', '\x80', '\x93', 'Dai', 'Type: ANY']], ['NUM',
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zgb', '\xe2', '\x80',
'\x93', 'Guibei', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zgb', '\xe2', '\x80', '\x93', 'Guibei', 'Type: ANY']],
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zgn', '\xe2', '\x80',
'\x93', 'Guibian', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zgn', '\xe2', '\x80', '\x93', 'Guibian', 'Type: ANY']],
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zln', '\xe2', '\x80',
'\x93', 'Lianshan', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zln', '\xe2', '\x80', '\x93', 'Lianshan', 'Type: ANY']],
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zlj', '\xe2', '\x80',
'\x93', 'Liujiang', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zlj', '\xe2', '\x80', '\x93', 'Liujiang', 'Type: ANY']],
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zfq', '\xe2', '\x80',
'\x93', 'Liuqian', 'Zhuang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zfq', '\xe2', '\x80', '\x93', 'Liuqian', 'Type: ANY']],
['NUM', '639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zgm', '\xe2', '\x80',

```

'\x93', 'Minz', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zgm', '\xe2', '\x80', '\x93', 'Minz', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zhn', '\xe2', '\x80',  
'\x93', 'Nong', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zhn', '\xe2', '\x80', '\x93', 'Nong', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zqe', '\xe2', '\x80',  
'\x93', 'Qiubei', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zqe', '\xe2', '\x80', '\x93', 'Qiubei', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyg', '\xe2', '\x80',  
'\x93', 'Yang', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zyg', '\xe2', '\x80', '\x93', 'Yang', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyb', '\xe2', '\x80',  
'\x93', 'Yongbei', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zyb', '\xe2', '\x80', '\x93', 'Yongbei', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyn', '\xe2', '\x80',  
'\x93', 'Yongnan', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zyn', '\xe2', '\x80', '\x93', 'Yongnan', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zyj', '\xe2', '\x80',  
'\x93', 'Youjiang', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zyj', '\xe2', '\x80', '\x93', 'Youjiang', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zzj', '\xe2', '\x80',  
'\x93', 'Zuojiang', 'Zhuang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zzj', '\xe2', '\x80', '\x93', 'Zuojiang', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

```

['Zhuang', 'Language', 'codes', 'ISO', '639-1', 'za']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['N\x7\x8u', 'Spoken', 'in', 'South', 'Africa']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['N\x7\x8u', 'Total', 'speakers', '10']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['N\x7\x8u', 'Language', 'codes', 'ISO', '639-3', 'ngh']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auslan', 'Signed', 'in', 'Australia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Algonquin', 'Spoken', 'in', 'Canada']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Algonquin', 'Language', 'codes', 'ISO', '639-3', 'alq']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['French', 'Signed', 'in', 'France']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Switzerland']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Mali']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Rwanda']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['French', 'Signed', 'in', 'Congo']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['French', 'Signed', 'in', 'Togo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['French', 'Signed', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dacian', 'Spoken', 'in', 'Romania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dacian', 'Spoken', 'in', 'Moldova']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dacian', 'Spoken', 'in', 'parts', 'of', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts', 'of', 'Ukraine']]], 6, True)

['Dacian', 'Spoken', 'in', 'Hungary']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dacian', 'Spoken', 'in', 'Serbia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dacian', 'Language', 'codes', 'ISO', '639-3', 'xdc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wenedyk', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Wenedyk', 'Date', 'founded', '2002']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wenedyk', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'alternative', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'purpose', ')', 'Wenedyk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY']]], 6, True)
```

```
['Wenedyk', 'Created', 'by', 'Jan', 'van', 'Steenbergen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Jan',  
'van', 'Steenbergen']]], 6, True)
```

```
['Wenedyk', 'Category', '(', 'sources', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)'], 'Type: ANY', 'Type: ANY']]], 7, True)
```

```
['Wenedyk', 'Category', '(', 'sources', ')', 'a', 'posteriori',  
'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)'], 'a', 'posteriori', 'language']]], 8, True)
```

```
['Wenedyk', 'Setting', 'and', 'usage', 'A', 'thought', 'experiment', 'in',  
'the', 'alternate', 'history']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'A', 'thought', 'experiment', 'in', 'the', 'alternate',  
'history']]], 11, True)
```

```
['Wenedyk', 'Setting', 'and', 'usage', 'Ill', 'Bethisad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Ill', 'Bethisad']]], 6, True)
```

```
['Wenedyk', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Wenedyk', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Sranan', 'Spoken', 'in', 'Suriname']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Sranan', 'Language', 'codes', 'ISO', '639-2', 'srn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sranan', 'Language', 'codes', 'ISO', '639-3', 'srn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sranan', 'Language', 'codes', 'Linguasphere', '52-ABB-aw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

["Martha's", 'Signed', 'in', 'formerly', 'in', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in',  
'formerly', 'in', 'the', 'United', 'States']]], 8, True)

["Martha's", 'Language', 'codes', 'ISO', '639-2', 'sgn-US-MA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Avar', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Avar', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Avar', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Avar', 'Spoken', 'in', 'Georgia', 'and', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Avar', 'Region', 'Republic', 'of', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Avar', 'Language', 'codes', 'ISO', '639-2', 'ava']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```

['Avar', 'Language', 'codes', 'ISO', '639-3', 'ava']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Avar', 'Language', 'codes', 'ISO', '639-1', 'av']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ruthenian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Ruthenian', 'Language', 'codes', 'ISO', '639-3', 'orv-olr']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Ancient', 'Iran']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Old', 'Writing', 'system', 'Old', 'Persian', 'Cuneiform']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'peo']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'peo']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Europanto', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']], 6, True)

['Europanto', 'Setting', 'and', 'usage', 'European', 'Union',
'administration']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',
'usage', 'European', 'Union', 'administration']], 7, True)

['Europanto', 'Created', 'by', 'Diego', 'Marani']:
C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:
ANY', 'Type: ANY']], 5, True)

```



```

['Europanto', 'Category', '(', 'purpose', ')', 'jest']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY']]), 6, True)

['Europanto', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary',
'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'int', '.', 'auxiliary', 'languages']]), 9, True)

['Europanto', 'Category', '(', 'purpose', ')', 'Europanto']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)'], 'Type: ANY']]), 6, True)

['Europanto', 'Date', 'founded', '1996']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]), 4, True)

['Europanto', 'Language', 'codes', 'ISO', '639-2', 'art']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Europanto', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Middle', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]), 5, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'dum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'dum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'southern', 'Germany', '(', 'south', 'of',
'the', 'Benrath', 'line', ')', ',', 'parts', 'of', 'Austria', 'and',
'Switzerland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'southern', 'Germany', '\\(', 'south', 'of', 'the', 'Benrath', 'line', '\\
\\)', ',', 'parts', 'of', 'Austria', 'and', 'Switzerland']]), 18, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'gmh']:

```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Middle', 'Language', 'codes', 'ISO', '639-3', 'gmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Early', 'Spoken', 'in', 'southern', 'Germany', '(', 'south', 'of', 'the', 'Benrath', 'line', ')', ',', 'parts', 'of', 'Austria', 'and', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'southern', 'Germany', '\\(', 'south', 'of', 'the', 'Benrath', 'line', '\\)', ',', 'parts', 'of', 'Austria', 'and', 'Switzerland']]], 18, True)
```

```
['Early', 'Language', 'codes', 'ISO', '639-2', 'deu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Early', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
["Dena'ina", 'Total', 'speakers', '75']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
["Dena'ina", 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Dena'ina", 'Writing', 'system', 'Latin', '(', "Dena'ina", 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
["Dena'ina", 'Language', 'codes', 'ISO', '639-3', 'tfn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ilocano', 'Spoken', 'in', 'Philippines', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',']]], 5, True)
```

```
['Ilocano', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

['Ilocano', 'Region', 'Northern', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ilocano', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']], 7, True)

['Ilocano', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Ilocano', 'Language', 'codes', 'ISO', '639-2', 'ilo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ilocano', 'Language', 'codes', 'ISO', '639-3', 'ilo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Phoenician', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'Lebanon']], 7, True)

['Phoenician', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'Malta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'Sicily']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Phoenician', 'Spoken', 'in', 'and', 'other', 'coastal', 'outposts', 'and', 'islands', 'throughout', 'the', 'Mediterranean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'other', 'coastal', 'outposts', 'and', 'islands', 'throughout', 'the', 'Mediterranean']], 12, True)

['Phoenician', 'Writing', 'system', 'Phoenician', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Phoenician', 'Language', 'codes', 'ISO', '639-2', 'phn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Phoenician', 'Language', 'codes', 'ISO', '639-3', 'phn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zauruan', 'Spoken', 'in', 'Nauru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zauruan', 'Language', 'codes', 'ISO', '639-2', 'nau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Zauruan', 'Language', 'codes', 'ISO', '639-3', 'nau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zauruan', 'Language', 'codes', 'ISO', '639-1', 'na']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Brithenig', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

```
['Brithenig', 'Date', 'founded', '1996']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'alternative', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'purpose', ')', 'Brithenig']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)
```

```
['Brithenig', 'Created', 'by', 'Andrew', 'Smith']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Brithenig', 'Category', '(', 'sources', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Type: ANY', 'Type: ANY']]], 7, True)
```

```
['Brithenig', 'Category', '(', 'sources', ')', 'a', 'posteriori',  
'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'a', 'posteriori', 'language']]], 8, True)
```

```
['Brithenig', 'Category', '(', 'sources', ')', '(', 'Romance', 'language',  
'based', 'on', 'Celtic', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', '\\\\(', 'Romance', 'language', 'based', 'on', 'Celtic', '\\  
\\)']]], 12, True)
```

```
['Brithenig', 'Setting', 'and', 'usage', 'A', 'thought', 'experiment',  
'in', 'alternate', 'history']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'A', 'thought', 'experiment', 'in', 'alternate', 'history']]], 10,  
True)
```

```
['Brithenig', 'Setting', 'and', 'usage', 'Ill', 'Bethisad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Ill', 'Bethisad']]], 6, True)
```

```
['Brithenig', 'Language', 'codes', 'ISO', '639-3', 'bzt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zimakani', 'Total', 'speakers', '1500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zimakani', 'Language', 'codes', 'ISO', '639-3', 'zik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Palatine', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Palatine', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Palatine', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)
```

```
['Palatine', 'Language', 'codes', 'ISO', '639-3', 'pfl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'southern', 'Germany', '(', 'south', 'of', 'the',  
'Benrath', 'line', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'southern', 'Germany', '\\\\(', 'south', 'of', 'the', 'Benrath', 'line', '\\  
)']]], 12, True)
```

```
['Old', 'Spoken', 'in', 'parts', 'of', 'Austria', 'and', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'parts',  
'of', 'Austria', 'and', 'Switzerland']]], 8, True)
```

```
['Old', 'Spoken', 'in', 'Southern', 'Bohemia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```

['Old', 'Language', 'codes', 'ISO', '639-2', 'goh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'goh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xavante', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Xavante', 'Region', 'Mato', 'Grosso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Xavante', 'Language', 'codes', 'ISO', '639-3', 'xav']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mogholi', 'Total', 'speakers', '200']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Mogholi', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Mogholi', 'Region', 'near', 'Herat']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Mogholi', 'Language', 'codes', 'ISO', '639-3', 'mhj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marwari', 'Total', 'speakers', '13', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '13']], 5, True)

['Marwari', 'Region', 'Rajasthan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Marwari', 'Region', 'Gujarat']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

```

ANY'[]], 3, True)

['Marwari', 'Region', 'Haryana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Marwari', 'Region', 'Sindh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Marwari', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Marwari', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Marwari', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Marwari', 'Writing', 'system', 'Devanagari', 'script', 'and', 'Mahajani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Devanagari', 'script', 'and', 'Mahajani'[]]], 7, True)

['Marwari', 'Language', 'codes', 'ISO', '639-2', 'mwr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'[]], ['NUM', '639-2'[]]], 6, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'mwr', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mwr', '\xe2', '\x80', '\x93', 'Macrolanguage'[]], ['NUM', '639-3'[]]], 10, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code'[]], ['NUM', '639-3'[]]], 7, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', ':', 'dhd', '\xe2', '\x80', '\x93', 'Dhundari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'dhd', '\xe2', '\x80', '\x93', 'Dhundari'[]], ['NUM', '639-3'[]]], 11, True)



['Marwari', 'Language', 'codes', 'ISO', '639-3', 'wry', '\xe2', '\x80', '\x93', 'Merwari']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wry', '\xe2', '\x80', '\x93', 'Merwari']], ['NUM', '639-3']], 10, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'mtr', '\xe2', '\x80', '\x93', 'Mewari']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mtr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Marwari', 'Language', 'codes', 'ISO', '639-3', 'smv', '\xe2', '\x80', '\x93', 'Shekhawati']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'smv', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Mozarabic', 'Spoken', 'in', 'Iberia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mozarabic', 'Language', 'codes', 'ISO', '639-3', 'mxi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hattic', 'Spoken', 'in', 'Anatolia', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Hattic', 'Spoken', 'in', 'Hattians']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hattic', 'Language', 'codes', 'ISO', '639-3', 'xht']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Somali', 'Spoken', 'in', 'Somalia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Somali', 'Spoken', 'in', 'Somaliland', '(', 'not', 'recognized', 'internationally', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',

```
'Somaliland', '\\(', 'not', 'recognized', 'internationally', '\\)']]], 9, True)
```

```
['Somali', 'Spoken', 'in', 'Djibouti']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Yemen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Somali', 'Spoken', 'in', 'Kenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Somali', 'Total', 'speakers', '13-25', 'million', 'native', 'speakers',  
'and', 'maybe', '500', ',', '000', 'second', 'language', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', 'native', 'speakers', 'and', 'maybe', 'NUM', ',', 'NUM',  
'second', 'language', 'speakers']], ['NUM', '13-25'], ['NUM', '500'],  
['NUM', '000']]], 15, True)
```

```
['Somali', 'Writing', 'system', 'Latin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Somali', 'Writing', 'system', 'Arabic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Somali', 'Writing', 'system', 'Osmanya', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Somali', 'Writing', 'system']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system']]],  
3, True)
```

```
['Somali', 'Official', 'language', 'in', 'Somalia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Somali', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Somali', 'Language', 'codes', 'ISO', '639-2', 'som']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Somali', 'Language', 'codes', 'ISO', '639-3', 'som']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Somali', 'Language', 'codes', 'ISO', '639-1', 'so']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Gutnish', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gutnish', 'Region', 'Gotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Gutnish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)
```

```
['Gutnish', 'Language', 'codes', 'ISO', '639-3', 'non-gut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['East', 'Spoken', 'in', 'Germany', ',', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['East', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['East', 'Language', 'codes', 'ISO', '639-2', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

['East', 'Language', 'codes', 'ISO', '639-3', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Venda', 'Region', 'Limpopo', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Venda', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Venda', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Venda', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Venda', 'Official', 'language', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Venda', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Venda', 'Language', 'codes', 'ISO', '639-2', 'ven']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Venda', 'Language', 'codes', 'ISO', '639-3', 'ven']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Venda', 'Language', 'codes', 'ISO', '639-1', 've']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bemba', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Bemba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',

'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Bemba', 'Spoken', 'in', 'Tanzania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bemba', 'Official', 'language', 'in', 'Zambia', '(', 'regional', 'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Zambia', '\\(', 'regional', 'language', '\\)']]], 9, True)

['Bemba', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Bemba', 'Language', 'codes', 'ISO', '639-2', 'bem']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bemba', 'Language', 'codes', 'ISO', '639-3', 'bem']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hopi', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hopi', 'Region', 'Northeastern', 'Arizona']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hopi', 'Language', 'codes', 'ISO', '639-3', 'hop']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adele', 'Spoken', 'in', 'Ghana', 'and', 'Togo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Adele', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Adele', 'Language', 'codes', 'ISO', '639-3', 'ade']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaonde', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kaonde', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Kaonde', 'Language', 'codes', 'ISO', '639-3', 'kqn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lozi', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lozi', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lozi', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lozi', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lozi', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Lozi', 'Language', 'codes', 'ISO', '639-2', 'loz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lozi', 'Language', 'codes', 'ISO', '639-3', 'loz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ekajuk', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ekajuk', 'Language', 'codes', 'ISO', '639-3', 'eka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Region', 'Mainly', 'in', 'the', 'Punjab', 'area', 'of', 'Pakistan', '.', 'Also', 'spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mainly', 'in', 'the', 'Punjab', 'area', 'of', 'Pakistan', '.', 'Also', 'spoken', 'in', 'Afghanistan']]], 14, True)

['Western', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Writing', 'system', 'Shahmukhi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'lah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'pnb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lunda', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lunda', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lunda', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Lunda', 'Region', 'Copperbelt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lunda', 'Language', 'codes', 'ISO', '639-2', 'lun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lunda', 'Language', 'codes', 'ISO', '639-3', 'lun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Masaba', 'Spoken', 'in', 'Uganda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Masaba', 'Region', 'Eastern']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Masaba', 'Region', 'south', 'of', 'the', 'Kupsabiny']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'south', 'of', 'the', 'Kupsabiny']], 6, True)

['Masaba', 'Region', 'Bugisu', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Masaba', 'Language', 'codes', 'ISO', '639-3', 'myx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yorkshire', 'Spoken', 'in', 'England']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yorkshire', 'Region', 'Yorkshire']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Yorkshire', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Mauritian', 'Spoken', 'in', 'Mauritius']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mauritian', 'Language', 'codes', 'ISO', '639-3', 'mfe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mauritian', 'Language', 'codes', 'Linguasphere', '51-AAC-cec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)



['Tahitian', 'Spoken', 'in', 'French', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tahitian', 'Language', 'codes', 'ISO', '639-2', 'tah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Tahitian', 'Language', 'codes', 'ISO', '639-3', 'tah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tahitian', 'Language', 'codes', 'ISO', '639-1', 'ty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Palauan', 'Spoken', 'in', 'Palau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Palauan', 'Spoken', 'in', 'Guam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Palauan', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Japanese', 'katakana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Palauan', 'Official', 'language', 'in', 'Palau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Palauan', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Palauan', 'Language', 'codes', 'ISO', '639-2', 'pau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Palauan', 'Language', 'codes', 'ISO', '639-3', 'pau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Aosta', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Piedmont']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Foggia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Franche-Comt', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Savoie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Bresse']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Bugey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Dombes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Beaujolais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Dauphin', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Lyonnais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Franco-Proven\xc3\xa7a1,', 'Region', 'Forez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY'']], 3, True)

['Franco-Proven\xc3\xa7a', 'Region', 'Romandy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Franco-Proven\xc3\xa7a', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Franco-Proven\xc3\xa7a', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Franco-Proven\xc3\xa7a', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Franco-Proven\xc3\xa7a', 'Writing', 'system', 'Latin', 'alphabet', 'with', 'diacritical', 'marks']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Franco-Proven\xc3\xa7a', 'Pronunciation', 'patu', '\xcb', '\x88', 'e', ',', 'patu', '\xcb', '\x88', '\xc9', '\x91']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'patu', '\xcb', '\x88', 'e', ',', 'patu', '\xcb', '\x88', '\xc9', '\x91']], 12, True)

['Franco-Proven\xc3\xa7a', 'Official', 'language', 'in', 'protected', 'by', 'statute', 'in', 'Italy', 'and', 'Aosta', 'Valley', 'Autonomous', 'Region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'protected', 'by', 'statute', 'in', 'Italy', 'and', 'Aosta', 'Valley', 'Autonomous', 'Region']], 14, True)

['Franco-Proven\xc3\xa7a', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Franco-Proven\xc3\xa7a', 'Language', 'codes', 'ISO', '639-3', 'frp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Franco-Proven\xc3\xa7a', 'Language', 'codes', 'Linguasphere', '51-AAA-

j']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Franco-Proven\xc3\xa7a', 'Language', 'codes', 'Linguasphere', ':', 'Dark', 'Blue', ':', 'Protected', '.', '\xe2', '\x80', '\x94', 'Medium', 'Blue', ':', 'General', 'regions']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', ':', 'Dark', 'Blue', ':', 'Protected', '.', '\xe2', '\x80', '\x94', 'Medium', 'Blue', ':', 'General', 'regions']]], 18, True)

['Franco-Proven\xc3\xa7a', 'Language', 'codes', 'Linguasphere', 'Light', 'Blue', ':', 'Historical', 'transition', 'zone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Light', 'Blue', ':', 'Historical', 'transition', 'zone']]], 10, True)

['Sicilian', 'Spoken', 'in', 'Sicily']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicilian', 'Spoken', 'in', 'Apulia', '(', 'Salento', ',', 'Lecce', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Apulia', '\\(', 'Salento', ',', 'Lecce', '\\)']]], 9, True)

['Sicilian', 'Spoken', 'in', 'Campania', '(', 'Cilento', ',', 'Campania', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Campania', '\\(', 'Cilento', ',', 'Campania', '\\)']]], 9, True)

['Sicilian', 'Spoken', 'in', 'Canada', '(', 'Vancouver', ',', 'Qu', '\xc3', '\xa9', 'bec', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada', '\\(', 'Vancouver', ',', 'Qu', '\xc3', '\xa9', 'bec', '\\)']]], 12, True)

['Sicilian', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicilian', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicilian', 'Spoken', 'in', 'Argentina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sicilian', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sicilian', 'Language', 'codes', 'ISO', '639-2', 'scn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sicilian', 'Language', 'codes', 'ISO', '639-3', 'scn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Comorian', 'Spoken', 'in', 'Comoros', 'and', 'Mayotte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Comorian', 'Region', 'Throughout', 'Comoros', 'and', 'Mayotte', ';;', 'also', 'in', 'Madagascar', 'and', 'R', '\xc3', '\xa9', 'union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Throughout', 'Comoros', 'and', 'Mayotte', ';;', 'also', 'in', 'Madagascar', 'and', 'R', '\xc3', '\xa9', 'union']], 15, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', ':', 'zdj', '\xe2', '\x80', '\x93', 'Ngazidja', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'zdj', '\xe2', '\x80', '\x93', 'Ngazidja', 'dialect']], ['NUM', '639-3']], 12, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', 'wni', '\xe2', '\x80', '\x93', 'Ndzwani', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wni', '\xe2', '\x80', '\x93', 'Ndzwani', 'dialect']], ['NUM', '639-3']], 11, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', 'swb', '\xe2', '\x80', '\x93', 'Maore', 'Comorian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'swb', '\xe2', '\x80', '\x93', 'Maore', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Comorian', 'Language', 'codes', 'ISO', '639-3', 'wlc', '\xe2', '\x80',

'\x93', 'Mwali', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wlc', '\xe2', '\x80', '\x93', 'Mwali', 'dialect']], ['NUM', '639-3']], 11, True)

['Maithili', 'Region', 'Bihar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Maithili', 'Region', 'Jharkhand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Maithili', 'Region', 'Terai', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Maithili', 'Spoken', 'in', 'India', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Maithili', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maithili', 'Writing', 'system', 'Kaithi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maithili', 'Writing', 'system', 'Mithilakshar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maithili', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Maithili', 'Language', 'codes', 'ISO', '639-2', 'mai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Maithili', 'Language', 'codes', 'ISO', '639-3', 'mai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maithili', 'Language', 'codes', 'ISO', '639-1', 'bh', '(', 'Bihari',

[')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bh', '\\\\(', 'Bihari', '\\\\)']], ['NUM', '639-1']], 9, True)

['Dakhni', 'Spoken', 'in', 'South', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Dakhni', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Dakhni', 'Writing', 'system', 'Urdu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dakhni', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Dakhni', 'Language', 'codes', 'ISO', '639-3', 'hin-dak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Haitian', 'Spoken', 'in', 'Haiti', '(', 'official', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)

['Haitian', 'Spoken', 'in', 'Bahamas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'Cuba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'Dominican', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Haitian', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haitian', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Haitian', 'Official', 'language', 'in', 'Haiti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Haitian', 'Language', 'codes', 'ISO', '639-2', 'hat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Haitian', 'Language', 'codes', 'ISO', '639-3', 'hat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Haitian', 'Language', 'codes', 'Linguasphere', '51-AAC-cb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Haitian', 'Language', 'codes', 'ISO', '639-1', 'ht']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Old', 'Spoken', 'in', 'Eastern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'orv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bhojpuri', 'Dialects', 'Caribbean', 'Hindustani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Dialects', 'Sarnami']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bhojpuri', 'Dialects', 'Mauritian', 'Bhojpuri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



'Type: ANY']]], 4, True)

['Bhojpuri', 'Region', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bhojpuri', 'Region', 'Bihar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bhojpuri', 'Region', 'Uttar', 'Pradesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Region', 'Jharkhand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bhojpuri', 'Spoken', 'in', 'moribund', 'in', 'Guyana', 'and', 'Trinidad', 'and', 'Tobago']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'moribund', 'in', 'Guyana', 'and', 'Trinidad', 'and', 'Tobago']]], 10, True)

['Bhojpuri', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Spoken', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Spoken', 'in', 'Mauritius']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Spoken', 'in', 'Suriname']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bhojpuri', 'Writing', 'system', 'Devanagari', ',', 'Kaithi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bhojpuri', 'Language', 'codes', 'ISO', '639-2', 'bho']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bhojpuri', 'Language', 'codes', 'ISO', '639-3', 'bho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bhojpuri', 'Language', 'codes', 'Linguasphere', '59-AAF-sa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Bhojpuri', 'Language', 'codes', 'ISO', '639-1', 'bh', '(', 'Bihari',  
'languages', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bh', '\\\\(', 'Bihari', 'languages', '\\\\)']], ['NUM',  
'639-1']], 10, True)

['Cajun', 'Spoken', 'in', 'Louisiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cajun', 'Language', 'codes', 'ISO', '639-3', 'frc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cajun', 'Language', 'codes', 'Linguasphere', '51-AAA-hp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Picard', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Picard', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Picard', 'Official', 'language', 'in', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Picard', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Picard', 'Language', 'codes', 'ISO', '639-3', 'pcd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Picard', 'Language', 'codes', 'Linguasphere', '51-AAA-he']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Crimean', 'Spoken', 'in', 'formerly', 'Crimea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Crimean', 'Writing', 'system', 'originally']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Crimean', 'Writing', 'system', 'probably', 'Runic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Crimean', 'Writing', 'system', 'later', 'Gothic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Crimean', 'Writing', 'system', 'Greek']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Crimean', 'Writing', 'system', 'and', 'Latin', 'alphabets']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Crimean', 'Language', 'codes', 'ISO', '639-2', 'got']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Crimean', 'Language', 'codes', 'ISO', '639-3', 'got', '\xe2', '\x80',  
'\x93', 'Gothic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'got', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Selonian,', 'Spoken', 'in', 'Latvia', 'and', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Selonian,', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY']]], 3, True)

['Selonian,', 'Language', 'codes', 'ISO', '639-3', 'sxl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thracian', 'Language', 'codes', 'ISO', '639-3', 'txh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chewa,', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chewa,', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chewa,', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chewa,', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chewa,', 'Official', 'language', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Chewa,', 'Official', 'language', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Chewa,', 'Regulated', 'by', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chewa,', 'Language', 'codes', 'ISO', '639-2', 'nya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chewa,', 'Language', 'codes', 'ISO', '639-3', 'nya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Chewa,', 'Language', 'codes', 'ISO', '639-1', 'ny']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Slavey', 'Spoken', 'in', 'Canada']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Slavey', 'Region', 'Northwest', 'Territories']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Slavey', 'Official', 'language', 'in', 'Northwest', 'Territories']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Slavey', 'Regulated', 'by', 'No', 'official', 'regulation']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']], 6, True)

['Slavey', 'Language', 'codes', 'ISO', '639-2', 'den']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Slavey', 'Language', 'codes', 'ISO', '639-3', 'den', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'den', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)

['Slavey', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Wu', 'Dialects', 'Oujiang', '(', 'Wenzhou', ')']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
'\(', 'Type: ANY', '\)']], 6, True)

['Wu', 'Dialects', 'Taihu']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Wu', 'Dialects', 'Taizhou']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

```

```

['Wu', 'Dialects', 'Wuzhou']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wu', 'Dialects', 'Chuju']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wu', 'Dialects', 'Xuanzhou']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wu', 'Region', 'Shanghai', ';', 'most', 'of', 'Zhejiang', 'province',
';', 'southern', 'Jiangsu', 'province', ';', 'Xuancheng', 'prefecture-
level', 'city', 'of', 'Anhui', 'province', ';', 'Shangrao', 'County']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Shanghai',
';', 'most', 'of', 'Zhejiang', 'province', ';', 'southern', 'Jiangsu',
'province', ';', 'Xuancheng', 'prefecture-level', 'city', 'of', 'Anhui',
'province', ';', 'Shangrao', 'County']]], 22, True)

['Wu', 'Region', 'Guangfeng', 'County', 'and', 'Yushan', 'County']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Guangfeng',
'County', 'and', 'Yushan', 'County']]], 7, True)

['Wu', 'Region', 'Jiangxi', 'province', ';', 'Pucheng', 'County']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jiangxi',
'province', ';', 'Pucheng', 'County']]], 7, True)

['Wu', 'Region', 'Hong', 'Kong']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Wu', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Wu', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Wu', 'Language', 'codes', 'ISO', '639-3', 'wu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wu', 'Language', 'codes', 'ISO', '639-1', 'zh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Tuvan', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvan', 'Spoken', 'in', 'Mongolia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvan', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tuvan', 'Region', 'Tuva']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tuvan', 'Official', 'language', 'in', 'Tuva', '(', 'federal', 'subject', 'of', 'Russia', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Tuva', '\\(', 'federal', 'subject', 'of', 'Russia', '\\)']], 11, True)

['Tuvan', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Tuvan', 'Language', 'codes', 'ISO', '639-2', 'tyv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tuvan', 'Language', 'codes', 'ISO', '639-3', 'tyv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auvergnat', 'Spoken', 'in', 'France']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Auvergnat', 'Region', 'Auvergne']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Auvergnat', 'Region', 'Puy-de-D', '\\xc3', '\\xb4', 'me']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Puy-de-D',

'\xc3', '\xb4', 'me']]], 6, True)

['Auvergnat', 'Region', 'Haute-Loire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Auvergnat', 'Region', 'Allier']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Auvergnat', 'Region', 'Cantal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Auvergnat', 'Region', 'Communities', 'in', 'Limousin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Auvergnat', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Auvergnat', 'Language', 'codes', 'ISO', '639-3', 'oci-auv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Auvergnat', 'Language', 'codes', 'Linguasphere', '51-AAA-gi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Soqotri', 'Spoken', 'in', 'Yemen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Soqotri', 'Region', 'Socotra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Soqotri', 'Language', 'codes', 'ISO', '639-3', 'sq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['L\xc3\xa1adan', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',



'\xe2', '\x80', '\x93']]], 6, True)

['L\xc3\xa1adan', 'Date', 'founded', '1982']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'constructed',  
'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'artistic', 'and',  
'philosophical', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'artistic', 'and', 'philosophical', 'language']]], 9,  
True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'fictional',  
'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['L\xc3\xa1adan', 'Category', '(', 'purpose', ')', 'L', '\xc3', '\xa1',  
'adan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'L', '\xc3', '\xa1', 'adan']]], 9, True)

['L\xc3\xa1adan', 'Created', 'by', 'Suzette', 'Haden', 'Elgin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
'Suzette', 'Haden', 'Elgin']]], 6, True)

['L\xc3\xa1adan', 'Language', 'codes', 'ISO', '639-3', 'ldn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piedmontese', 'Spoken', 'in', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Piedmontese', 'Region', 'northwest', 'Italy', ',', 'Piedmont']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northwest',  
'Italy', ',', 'Piedmont']]], 6, True)

['Piedmontese', 'Language', 'codes', 'ISO', '639-3', 'pms']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piedmontese', 'Language', 'codes', 'Linguasphere', '51-AAA-of']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Suzhou', 'dialect', 'Total', 'speakers', 'approx', '.', '5-7',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'approx', '.', 'NUM', 'million']]], ['NUM', '5-7']], 8, True)

['Suzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)

['Suzhou', 'Language', 'codes', 'ISO', '639-3', 'wu-suh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['K\x3\xb6lsch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['K\x3\xb6lsch', 'Region', 'Cologne', 'and', 'environs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['K\x3\xb6lsch', 'Language', 'codes', 'ISO', '639-3', 'ksh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Low', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Low', 'Spoken', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Low', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Low', 'Total', 'speakers', '\xe2', '\x89', '\x88', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',

'\xe2', '\x89', '\x88', 'NUM', 'million']], ['NUM', '5']], 8, True)

['Low', 'Language', 'codes', 'ISO', '639-2', 'nds']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Low', 'Language', 'codes', 'Linguasphere', '52-ACB']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Wyandot', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']], 7, True)

['Wyandot', 'Region', 'northeastern', 'Oklahoma', ',', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northeastern', 'Oklahoma', ',', 'Type: ANY']], 6, True)

['Wyandot', 'Writing', 'system', 'modified', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Wyandot', 'Language', 'codes', 'ISO', '639-3', 'wya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Elamite', 'Region', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Elamite', 'Language', 'codes', 'ISO', '639-2', 'elx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Elamite', 'Language', 'codes', 'ISO', '639-3', 'elx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chiac', 'Shiac', ',', 'Shiak', ',', 'Chiak', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Shiac', ',', 'Shiak', ',', 'Chiak', 'Total', 'speakers', 'Unknown']], 9, True)

['Chiac', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chiac,', 'Region', 'mainly', 'around', 'Moncton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Chiac,', 'Region', 'Shediac', 'and', 'Memramcook']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-2', 'fre', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fre', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-2', 'fra', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'fra', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-3', 'fra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chiac,', 'Language', 'codes', 'Linguasphere', '51-AAA-am']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Chiac,', 'Language', 'codes', 'ISO', '639-1', 'fr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Isan', 'Total', 'speakers', '20', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '20']], 5, True)

['Isan', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Isan', 'Region', 'Isan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Isan', 'Language', 'codes', 'ISO', '639-3', 'tts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Norfolk', 'Spoken', 'in', 'Norfolk', 'Island']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Norfolk', 'Spoken', 'in', 'Pitcairn', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Norfolk', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Norfolk', 'Pronunciation', 'n', '\xc9', '\x94', '\xcb', '\x90', 'fuk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\x94', '\xcb', '\x90', 'fuk']]], 8, True)

['Norfolk', 'Official', 'language', 'in', 'Norfolk', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Norfolk', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Norfolk', 'Language', 'codes', 'ISO', '639-3', 'pih']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bundeli', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bundeli', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bundeli', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bundeli', 'Official', 'language', 'in', 'India', '(', 'Madhya', 'Pradesh', 'and', 'parts', 'of', 'Uttar', 'Pradesh', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'India', '\\(', 'Madhya', 'Pradesh', 'and', 'parts', 'of', 'Uttar', 'Pradesh', '\\)']]], 14, True)

['Bundeli', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Bundeli', 'Language', 'codes', 'ISO', '639-3', 'bns']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bagheli', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bagheli', 'Region', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bagheli', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bagheli', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Bagheli', 'Language', 'codes', 'ISO', '639-3', 'bfy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khwarezmian', 'Spoken', 'in', 'Khwarezm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khwarezmian', 'Writing', 'system', 'Aramaic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Khwarezmian', 'Writing', 'system', 'Sogdian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Khwarezmian', 'Writing', 'system', 'Pahlavi', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Khwarezmian', 'Language', 'codes', 'ISO', '639-3', 'xco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hawaii', 'Spoken', 'in', 'US']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hawaii', 'Language', 'codes', 'ISO', '639-3', 'hwc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hawaii', 'Language', 'codes', 'Linguasphere', '52-ABB-dc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Torres', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Torres', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Torres', 'Language', 'codes', 'ISO', '639-2', 'cpe', '(', '?', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'cpe', '\\\\(', '?', '\\\\)']], ['NUM', '639-2']], 9, True)

['Torres', 'Language', 'codes', 'ISO', '639-3', 'tcs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Korandje', 'Total', 'speakers', '3000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Korandje', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Korandje', 'Region', 'Tabelbala', ',', 'wilaya', 'of', 'Bechar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', ',', 'Type: ANY', 'of', 'Type: ANY']], 7, True)

['Korandje', 'Language', 'codes', 'ISO', '639-3', 'kcy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zarma', 'Spoken', 'in', 'southwest', 'Niger']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Zarma', 'Language', 'codes', 'ISO', '639-3', 'dje']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Zyrian', 'Region', 'Komi', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zyrian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zyrian', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Zyrian', 'Official', 'language', 'in', 'Komi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Zyrian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Zyrian', 'Language', 'codes', 'ISO', '639-2', 'kom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Zyrian', 'Language', 'codes', 'ISO', '639-3', 'kpv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Zyrian', 'Language', 'codes', 'ISO', '639-1', 'kv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Saraiki', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Saraiki', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



```
['Saraiki', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Saraiki', 'Writing', 'system', 'Perso-Arabic', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Saraiki', 'Writing', 'system', 'Gurmukhi', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Saraiki', 'Writing', 'system', 'Devanagari', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Saraiki', 'Language', 'codes', 'ISO', '639-3', 'skr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shanghainese', 'Total', 'speakers', '14', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '14']], 5, True)
```

```
['Shanghainese', 'Pronunciation', 'z', '\xcc', '\xa5', '\xc9', '\x91',
'\xcc', '\x83', '\xcc', '\x80', 'h', '\xc3', '\xa9', '\xc9', '\xa6',
'\xc9', '\x9b', '\xcc', '\x80', '\xc9', '\xa6', '\xca', '\x8a', '\xcc',
'\x80', ',', '\xc9', '\xa6', 'u', '\xc9', '\xb2', 'y']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'z',
'\xcc', '\xa5', '\xc9', '\x91', '\xcc', '\x83', '\xcc', '\x80', 'h',
'\xc3', '\xa9', '\xc9', '\xa6', '\xc9', '\x9b', '\xcc', '\x80', '\xc9',
'\xa6', '\xca', '\x8a', '\xcc', '\x80', ',', '\xc9', '\xa6', 'u', '\xc9',
'\xb2', 'y']], 33, True)
```

```
['Shanghainese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93', 'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)
```

```
['Shanghainese', 'Language', 'codes', 'ISO', '639-3', 'wuu-sha']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lakota', 'Region', 'but', 'also', 'northern', 'Nebraska']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'but', 'also', 'northern', 'Nebraska']]], 6, True)

['Lakota', 'Region', 'southern', 'Minnesota']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lakota', 'Region', 'and', 'northern', 'Montana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lakota', 'Spoken', 'in', 'United', 'States', ',', 'with', 'some', 'speakers', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', ',', 'with', 'some', 'speakers', 'in', 'Canada']]], 11, True)

['Lakota', 'Pronunciation', 'la', '\xcb', '\x88', 'k', '\xcb', '\xa3', 'otijapi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'la', '\xcb', '\x88', 'k', '\xcb', '\xa3', 'otijapi']]], 9, True)

['Lakota', 'Language', 'codes', 'ISO', '639-3', 'lkt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rumsen', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Rumsen', 'Total', 'speakers', 'extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rumsen', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Rumsen', 'Language', 'codes', 'ISO', '639-2', 'nai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Rumsen', 'Language', 'codes', 'ISO', '639-3', 'css']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rumsen', 'Language', 'codes', 'ISO', '639-1', '-']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Komi-Permyak', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Komi-Permyak', 'Region', 'Komi-Permyak', 'Okrug']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Komi-Permyak', 'Language', 'codes', 'ISO', '639-2', 'kom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Komi-Permyak', 'Language', 'codes', 'ISO', '639-3', 'koi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Komi-Permyak', 'Language', 'codes', 'ISO', '639-1', 'kv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Erzya', 'Total', 'speakers', 'Ethnologue']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Armenia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Estonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Region', 'Mordovia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Nizhny', 'Novgorod']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erzya', 'Region', 'Chuvashia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Ulyanovsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Samara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Penza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Saratov']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Erzya', 'Region', 'Orenburg']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Erzya', 'Region', 'Tatarstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Erzya', 'Region', 'Bashkortostan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Erzya', 'Language', 'codes', 'ISO', '639-2', 'myv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Erzya', 'Language', 'codes', 'ISO', '639-3', 'myv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Evenki', 'Region', 'Inner', 'Mongolia', 'and', 'Heilongjiang', 'in', 'China', ';', 'Selenge', 'Province', 'in', 'Mongolia', ';', 'Krasnoyarsk', 'Krai', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Inner', 'Mongolia', 'and', 'Heilongjiang', 'in', 'China', ';', 'Selenge', 'Province', 'in', 'Mongolia', ';', 'Krasnoyarsk', 'Krai', 'in', 'Russia']]], 18, True)
```

```
['Evenki', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Evenki', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Evenki', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Evenki', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Evenki', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

['Evenki', 'Writing', 'system', 'Mongolian', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Evenki', 'Language', 'codes', 'ISO', '639-3', 'evn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shoshoni', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Shoshoni', 'Region', 'Wyoming']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Shoshoni', 'Region', 'Utah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Shoshoni', 'Region', 'Nevada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Shoshoni', 'Region', 'Idaho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Shoshoni', 'Language', 'codes', 'ISO', '639-3', 'shh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balti', 'Spoken', 'in', 'Baltistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Balti', 'Spoken', 'in', 'Ladakh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Balti', 'Language', 'codes', 'ISO', '639-3', 'bft']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalmyk', 'Region', 'Kalmykia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kalmyk', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalmyk', 'Writing', 'system', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalmyk', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kalmyk', 'Writing', 'system', 'Clear', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Kalmyk', 'Official', 'language', 'in', 'Kalmykia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Kalmyk', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Kalmyk', 'Language', 'codes', 'ISO', '639-2', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kalmyk', 'Language', 'codes', 'ISO', '639-3', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Beja', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beja', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beja', 'Spoken', 'in', 'Egypt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beja', 'Language', 'codes', 'ISO', '639-2', 'bej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Beja', 'Language', 'codes', 'ISO', '639-3', 'bej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quiripi', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Quiripi', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Quiripi', 'Language', 'codes', 'ISO', '639-3', 'qyp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kabardian', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kabardian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kabardian', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kabardian', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kabardian', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kabardian', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kabardian', 'Spoken', 'in', 'USA']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kabardian', 'Region', 'Kabardino-Balkaria', ',', 'Karachay-Cherkessia', 'and', 'north', 'east', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kabardino-Balkaria', ',', 'Karachay-Cherkessia', 'and', 'north', 'east', 'Turkey']]], 9, True)

['Kabardian', 'Language', 'codes', 'ISO', '639-2', 'kbd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kabardian', 'Language', 'codes', 'ISO', '639-3', 'kbd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Blackfoot', 'Total', 'speakers', '5', ',', '100', '/']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '/']]], ['NUM', '5'], ['NUM', '100']], 7, True)

['Blackfoot', 'Spoken', 'in', 'United', 'States', ',', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Blackfoot', 'Region', 'Blackfeet', 'Indian', 'Reservation', 'in', 'Montana', 'and', 'Piikani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Blackfeet', 'Indian', 'Reservation', 'in', 'Montana', 'and', 'Piikani']]], 9, True)

['Blackfoot', 'Region', 'Siksika']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Blackfoot', 'Region', 'and', 'Kainai', 'Reserves', 'in', 'southern', 'Alberta']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'Kainai', 'Reserves', 'in', 'southern', 'Alberta']]], 8, True)

['Blackfoot', 'Language', 'codes', 'ISO', '639-2', 'bla']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Blackfoot', 'Language', 'codes', 'ISO', '639-3', 'bla']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kongo', 'Total', 'speakers', '7', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '7']], 5, True)

['Kongo', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kongo', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Kongo', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'the', 'Congo']]], 7, True)

['Kongo', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kongo', 'Official', 'language', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Kongo', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of',  
'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 9, True)

['Kongo', 'Official', 'language', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Kongo', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Kongo', 'Language', 'codes', 'ISO', '639-2', 'kon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', 'kon', '\xe2', '\x80',  
'\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kon', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',

'639-3']], 10, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', ':', 'kng', '\xe2', '\x80',  
\x93', 'Koongo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'kng', '\xe2', '\x80', '\x93', 'Koongo']], ['NUM',  
'639-3']], 11, True)

['Kongo', 'Language', 'codes', 'ISO', '639-3', 'ldi', '\xe2', '\x80',  
\x93', 'Laari']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ldi', '\xe2', '\x80', '\x93', 'Laari']], ['NUM', '639-3']],  
10, True)

['Kongo', 'Language', 'codes', 'ISO', '639-1', 'kg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Lombard', 'Spoken', 'in', 'Italy', ',', 'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Lombard', 'Region', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Lombard', 'Region', ':', 'Lombardy', ',', 'Piedmont']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', ':',  
'Lombardy', ',', 'Piedmont']], 6, True)

['Lombard', 'Region', 'Switzerland']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Lombard', 'Region', ':', 'Ticino']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lombard', 'Region', 'Graub', '\xc3', '\xbc', 'nden']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Graub',  
'\xc3', '\xbc', 'nden']], 6, True)

['Lombard', 'Region', 'Val', 'Poschiavo', ',', 'Val', 'Mesolcina', 'and', 'Val', 'Bregaglia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Val', 'Poschiavo', ',', 'Val', 'Mesolcina', 'and', 'Val', 'Bregaglia']]], 10, True)

['Lombard', 'Language', 'codes', 'ISO', '639-3', 'lmo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Moabite', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern', 'Jordan']]], 8, True)

['Moabite', 'Language', 'codes', 'ISO', '639-3', 'obm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["O'odham", 'Spoken', 'in', 'United', 'States', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

["O'odham", 'Region', 'Primarily', 'south-central', 'Arizona', 'and', 'northern', 'Sonora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Primarily', 'south-central', 'Arizona', 'and', 'northern', 'Sonora']]], 8, True)

["O'odham", 'Official', 'language', 'in', 'One', 'of', 'the', 'national', 'languages', 'of', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'One', 'of', 'the', 'national', 'languages', 'of', 'Mexico']]], 11, True)

["O'odham", 'Regulated', 'by', 'Secretary', 'of', 'Public', 'Education', 'in', 'Mexico', ';', 'various', 'tribal', 'agencies', 'in', 'the', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Secretary', 'of', 'Public', 'Education', 'in', 'Mexico', ';', 'various', 'tribal', 'agencies', 'in', 'the', 'USA']]], 16, True)

["O'odham", 'Language', 'codes', 'ISO', '639-3', 'ood']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Phrygian', 'Language', 'codes', 'ISO', '639-3', 'xpg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adyghe', 'Spoken', 'in', 'Russia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Turkey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Jordan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Syria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'Israel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adyghe', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Adyghe', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Macedonia']]], 6, True)

['Adyghe', 'Spoken', 'in', 'Iraq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Adyghe', 'Region', 'Russia', ':', 'Republic', 'of', 'Adygea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Russia', ':',  
'Republic', 'of', 'Adygea']]], 7, True)

['Adyghe', 'Official', 'language', 'in', 'Republic', 'of', 'Adygea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Republic', 'of', 'Adygea']]], 7, True)

['Adyghe', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

```

['Adyghe', 'Language', 'codes', 'ISO', '639-2', 'ady']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Adyghe', 'Language', 'codes', 'ISO', '639-3', 'ady']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['H\x3\xa4n', 'Total', 'speakers', '10']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['H\x3\xa4n', 'Region', 'Yukon', ',', 'Alaska']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['H\x3\xa4n', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY', 'S']], 7, True)

['H\x3\xa4n', 'Language', 'codes', 'ISO', '639-3', 'haa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ojibwe', 'Region', 'Canada', ':', 'Quebec']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['Ojibwe', 'Region', 'Ontario']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Ojibwe', 'Region', 'Manitoba']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Ojibwe', 'Region', 'Saskatchewan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Ojibwe', 'Region', 'groups', 'in', 'Alberta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['Ojibwe', 'Region', 'Wisconsin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

```

ANY']]], 3, True)

['Ojibwe', 'Region', 'Minnesota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ojibwe', 'Region', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ojibwe', 'Spoken', 'in', 'Canada', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', '']]], 5, True)

['Ojibwe', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ojibwe', 'Pronunciation', 'an', '\xc9', '\xaa', '\xca', '\x83', '\xc9', '\xaa', 'na', '\xcb', '\x90', 'pe', '\xcb', '\x90', 'mow', '\xc9', '\xaa', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'an', '\xc9', '\xaa', '\xca', '\x83', '\xc9', '\xaa', 'na', '\xcb', '\x90', 'pe', '\xcb', '\x90', 'mow', '\xc9', '\xaa', 'n']]], 19, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-2', 'oji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'oji', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'oji', '\xe2', '\x80', '\x93', 'Macrolanguage']]], ['NUM', '639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']]], ['NUM', '639-3']], 7, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ciw', '\xe2', '\x80', '\x93', 'Chippewa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ciw', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'otw', '\xe2', '\x80',

```

'\x93', 'Ottawa']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'otw', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'alq', '\xe2', '\x80',
'\x93', 'Algonquin']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'alq', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Ojibwe', 'Language', 'codes', 'ISO', '639-1', 'oj']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Gen', 'Spoken', 'in', 'Togo', ',', 'Benin']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)

['Gen', 'Language', 'codes', 'ISO', '639-3', 'gej']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Stadsfries', 'Spoken', 'in', 'Netherlands']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Stadsfries', 'Writing', 'system', 'Latin', 'alphabet']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

['Stadsfries', 'Pronunciation', '\xcb', '\x88', 'st', '\xc9', '\x91',
'tsfris']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',
'\x88', 'st', '\xc9', '\x91', 'tsfris']], 8, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-2', 'dut', '(', 'B', ')']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'dut', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-2', 'nld', '(', 'T', ')']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'nld', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-3', 'nld']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Stadsfries', 'Language', 'codes', 'ISO', '639-1', 'nl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['West', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['West', 'Region', 'Friesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['West', 'Region', 'Groningen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['West', 'Official', 'language', 'in', 'Province', 'of', 'Friesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Province', 'of', 'Friesland']]], 7, True)

['West', 'Regulated', 'by', 'Fryske', 'Akademy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['West', 'Language', 'codes', 'ISO', '639-2', 'fry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'fry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['West', 'Language', 'codes', 'Linguasphere', '52-ACA-b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['West', 'Language', 'codes', 'Linguasphere', 'Map', 'showing', 'the',  
'Frisian-speaking', 'areas', 'in', 'the', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Map', 'showing', 'the', 'Frisian-speaking', 'areas', 'in',  
'the', 'Netherlands']]], 12, True)

['West', 'Language', 'codes', 'ISO', '639-1', 'fy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Saterland', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saterland', 'Region', 'Lower', 'Saxony']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saterland', 'Language', 'codes', 'ISO', '639-2', 'frs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Saterland', 'Language', 'codes', 'ISO', '639-3', 'stq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['North', 'Region', 'Schleswig-Holstein']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['North', 'Official', 'language', 'in', 'Nordfriesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['North', 'Official', 'language', 'in', 'Heligoland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['North', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-2', 'frr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'frr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sireniki', 'Writing', 'system', 'Transcribed', 'with', 'Cyrillic', 'in', 'old', 'monographs', '(', 'extended', 'with', 'diacritics', ')', ',', 'but', 'new', 'publications', 'may', 'appear', 'also', 'romanized']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Transcribed', 'with', 'Cyrillic', 'in', 'old', 'monographs', '\\\\(', 'extended', 'with', 'diacritics', '\\\\)'], ',', 'but', 'new', 'publications', 'may', 'appear', 'also', 'romanized']]], 22, True)

['Sireniki', 'Language', 'codes', 'ISO', '639-3', 'ysr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Soga', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Soga', 'Region', 'Mainly', 'in', 'Busoga', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mainly', 'in', 'Busoga', 'region']]], 6, True)

['Soga', 'Language', 'codes', 'ISO', '639-3', 'xog']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pangasinense', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pangasinense', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Pangasinense', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Pangasinense', 'Language', 'codes', 'ISO', '639-2', 'pag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pangasinense', 'Language', 'codes', 'ISO', '639-3', 'pag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Ottoman', 'Spoken', 'in', 'Ottoman', 'Empire-Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Ottoman', 'Writing', 'system', 'Ottoman', 'Turkish', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Ottoman', 'Language', 'codes', 'ISO', '639-2', 'ota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Ottoman', 'Language', 'codes', 'ISO', '639-3', 'ota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Finnish', 'Signed', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bulgar', 'Language', 'codes', 'ISO', '639-3', 'xbo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Limousin', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Limousin', 'Total', 'speakers', 'no', 'reliable', 'data']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'no', 'reliable', 'data']]], 6, True)
```

```
['Limousin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Limousin', 'Language', 'codes', 'ISO', '639-3', 'oci-lim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Limousin', 'Language', 'codes', 'Linguasphere', '51-AAA-gj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```

['Meitei', 'Spoken', 'in', 'Northeast', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Meitei', 'Spoken', 'in', 'Bangladesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Meitei', 'Spoken', 'in', 'Burma']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Meitei', 'Writing', 'system', 'Eastern', 'Nagari', '(', 'currently',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Meitei', 'Writing', 'system', 'Meitei', 'Mayek', '(', 'formerly', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Meitei', 'Official', 'language', 'in', 'India', '(', 'Manipur', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'India', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Meitei', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Meitei', 'Language', 'codes', 'ISO', '639-2', 'mni']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Meitei', 'Language', 'codes', 'ISO', '639-3', 'mni']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sylheti', 'Spoken', 'in', 'Including', 'the', 'Bangladeshi', 'diaspora',
'communities', 'in', 'Britain', 'and', 'the', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Including', 'the', 'Bangladeshi', 'diaspora', 'communities', 'in',
'Britain', 'and', 'the', 'United', 'States']]], 14, True)

['Sylheti', 'Spoken', 'in', 'Primarily', 'spoken', 'in', 'the', 'Sylhet',
'region', 'of', 'Bangladesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',

```

'Primarily', 'spoken', 'in', 'the', 'Sylhet', 'region', 'of',  
'Bangladesh']]], 11, True)

['Sylheti', 'Spoken', 'in', 'and', 'the', 'Indian', 'districts', 'of', ':',  
'Barak', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and',  
'the', 'Indian', 'districts', 'of', ':', 'Barak', 'Valley']]], 11, True)

['Sylheti', 'Spoken', 'in', 'Silchar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sylheti', 'Spoken', 'in', 'Hailakandi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sylheti', 'Spoken', 'in', 'Karimganj', 'and', 'Tripura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Sylheti', 'Writing', 'system', 'Sylheti', 'Nagari', ',', 'Bengali',  
'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Sylheti', 'Language', 'codes', 'ISO', '639-3', 'syl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kemi', 'Spoken', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kemi', 'Language', 'codes', 'ISO', '639-3', 'sjk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papi\x3\xa1', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Papi\x3\xa1', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Papi\x3\xa1', 'Language', 'codes', 'ISO', '639-2', 'cpp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Papi\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'mcm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papi\xc3\xa1', 'Language', 'codes', 'Linguasphere', '51-AAC-aha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Niuean', 'Spoken', 'in', 'Niue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Niuean', 'Spoken', 'in', 'Cook', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Niuean', 'Spoken', 'in', 'New', 'Zealand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Niuean', 'Spoken', 'in', 'Tonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Niuean', 'Language', 'codes', 'ISO', '639-2', 'niu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Niuean', 'Language', 'codes', 'ISO', '639-3', 'niu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Mi'kmaq", 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']], 7, True)

["Mi'kmaq", 'Region', 'Nova', 'Scotia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Mi'kmaq", 'Region', 'New', 'Brunswick']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

["Mi'kmaq", 'Region', 'Newfoundland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

["Mi'kmaq", 'Region', 'northern', 'Maine']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Mi'kmaq", 'Region', 'Boston', '(', 'Massachusetts', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Boston', '\\\\(', 'Massachusetts', '\\\\)']], 6, True)

["Mi'kmaq", 'Language', 'codes', 'ISO', '639-2', 'mic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

["Mi'kmaq", 'Language', 'codes', 'ISO', '639-3', 'mic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mandaic', 'Total', 'speakers', '5500']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'Iraq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'USA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mandaic', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ammonite', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'northwestern',

```



'Jordan']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
 'Formerly', 'spoken', 'in', 'northwestern', 'Jordan']]], 8, True)

['Ammonite', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
 'Linguist', 'List']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
 '639-3']], 10, True)

['Ammonite', 'Language', 'codes', 'ISO', '639-3', 'agg']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Meroitic', 'Spoken', 'in', 'Sudan']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY']], 4, True)

['Meroitic', 'Region', 'Mero', '\xc3', '\xab']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
 'Type: ANY', 'Type: ANY']], 5, True)

['Meroitic', 'Writing', 'system', 'Meroitic', 'alphabet']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
 'Type: ANY', 'Type: ANY']], 5, True)

['Meroitic', 'Language', 'codes', 'ISO', '639-3', 'xmr']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Santhali', 'Spoken', 'in', 'India']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY']], 4, True)

['Santhali', 'Spoken', 'in', 'Bangladesh']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY']], 4, True)

['Santhali', 'Spoken', 'in', 'Nepal']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY']], 4, True)

['Santhali', 'Spoken', 'in', 'Bhutan']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
 'Type: ANY']], 4, True)

['Santhali', 'Writing', 'system', 'Latin', 'alphabet', ',', '01', 'Chiki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Santhali', 'Language', 'codes', 'ISO', '639-2', 'sat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Santhali', 'Language', 'codes', 'ISO', '639-3', 'sat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Galician\xe2\x80\x93Portuguese', 'Spoken', 'in', 'Kingdom', 'of',  
'Galicia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kingdom', 'of', 'Galicia']]], 6, True)

['Galician\xe2\x80\x93Portuguese', 'Region', 'North-west', 'Iberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Galician\xe2\x80\x93Portuguese', 'Language', 'codes', 'ISO', '639-3',  
'\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)

['Galician\xe2\x80\x93Portuguese', 'Language', 'codes', 'ISO', '639-3',  
'079']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Carolinian', 'Region', 'Saipan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Carolinian', 'Region', 'Anatahan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Carolinian', 'Region', 'and', 'Agrihan', 'islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Carolinian', 'Region', 'Carolines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY'']], 3, True)

['Carolinian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Carolinian', 'Language', 'codes', 'ISO', '639-3', 'cal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dene', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dene', 'Region', 'Northern', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dene', 'Region', 'Saskatchewan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dene', 'Region', 'Manitoba', ';', 'southern', 'Northwest', 'Territories',  
'and', 'Nunavut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Manitoba',  
';', 'southern', 'Northwest', 'Territories', 'and', 'Nunavut']]], 9, True)

['Dene', 'Language', 'codes', 'ISO', '639-2', 'chp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dene', 'Language', 'codes', 'ISO', '639-3', 'chp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Forro', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9',  
'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'S',  
'\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad',  
'ncipe']]], 15, True)

['Forro', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Forro', 'Language', 'codes', 'ISO', '639-3', 'cri']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Forro', 'Language', 'codes', 'Linguasphere', '51-AAC-aca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Kristi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kristi', 'Total', 'speakers', '750']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kristi', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kristi', 'Language', 'codes', 'ISO', '639-3', 'vvp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kristi', 'Language', 'codes', 'Linguasphere', '51-AAC-agc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Angolar', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']], 15, True)

['Angolar', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Angolar', 'Language', 'codes', 'ISO', '639-3', 'aoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Angolar', 'Language', 'codes', 'Linguasphere', '51-AAC-ad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Principense', 'Spoken', 'in', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3',

'\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'S',  
'\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad',  
'ncipe']], 15, True)

['Principense', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Principense', 'Language', 'codes', 'ISO', '639-3', 'pre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Principense', 'Language', 'codes', 'Linguasphere', '51-AAC-acb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']], 5, True)

['Macanese', 'Spoken', 'in', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Macanese', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', '\\(', 'Type: ANY', '\\)']], 8, True)

['Macanese', 'Spoken', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Macanese', 'Spoken', 'in', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macanese', 'Spoken', 'in', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Macanese', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Macanese', 'Language', 'codes', 'ISO', '639-3', 'mzs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Macanese', 'Language', 'codes', 'Linguasphere', '51-AAC-ai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Gwich\x00\x99in', 'Region', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gwich\x00\x99in', 'Region', 'Yukon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gwich\x00\x99in', 'Region', 'Alaska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gwich\x00\x99in', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Gwich\x00\x99in', 'Official', 'language', 'in', 'Northwest', 'Territories', '(' , 'Canada', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Northwest', 'Territories', '\\(' , 'Canada', '\\)']]], 9, True)

['Gwich\x00\x99in', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Gwich\x00\x99in', 'Language', 'codes', 'ISO', '639-2', 'gwi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gwich\u00e2\u0080\u0099in', 'Language', 'codes', 'ISO', '639-3', 'gwi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Penang', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Penang', 'Region', 'Penang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Penang', 'Region', 'Kedah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Penang', 'Region', 'Perak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Penang', 'Region', 'Medan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Penang', 'Region', 'Phuket']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Penang', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Penang', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Penang', 'Language', 'codes', 'ISO', '639-3', 'nan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Penang', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Amorite', 'Spoken', 'in', 'ancient', 'Mesopotamia', ',', 'by', 'the', 'Amorites']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'ancient', 'Mesopotamia', ',', 'by', 'the', 'Amorites']]], 9, True)

['Amorite', 'Language', 'codes', 'ISO', '639-3', 'None']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dogrib', 'Spoken', 'in', 'Canada']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dogrib', 'Region', 'Northwest', 'Territories']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dogrib', 'Official', 'language', 'in', 'Northwest', 'Territories']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['Dogrib', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Dogrib', 'Language', 'codes', 'ISO', '639-2', 'dgr']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dogrib', 'Language', 'codes', 'ISO', '639-3', 'dgr']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dura', 'Language', 'extinction', '?']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dura', 'Spoken', 'in', 'East', 'Asia']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Dura', 'Language', 'codes', 'ISO', '639-3', 'drq']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Daman', 'Spoken', 'in', 'India']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Daman', 'Total', 'speakers', '4000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Daman', 'Language', 'codes', 'ISO', '639-2', 'cpp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Daman', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Daman', 'Language', 'codes', 'Linguasphere', '51-AAC-agb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Leponitic', 'Spoken', 'in', 'Cisalpine', 'Gaul']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Leponitic', 'Language', 'codes', 'ISO', '639-3', 'xlp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Messapian,', 'Spoken', 'in', 'Puglia', 'region', 'of', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Puglia', 'region', 'of', 'Italy']]], 7, True)

['Messapian,', 'Language', 'codes', 'ISO', '639-3', 'cms']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hurrian', 'Spoken', 'in', 'Mitanni']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hurrian', 'Region', 'Mesopotamia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hurrian', 'Language', 'codes', 'ISO', '639-3', 'xhu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ongota', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ongota', 'Language', 'codes', 'ISO', '639-3', 'bxé']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Menominee', 'Total', 'speakers', '130']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Menominee', 'Region', 'Northeastern', 'Wisconsin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Menominee', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Menominee', 'Pronunciation', 'om', '\xc3', '\xa6', '\xcb', '\x90', '\xca', '\x94', 'nomenew']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'om', '\xc3', '\xa6', '\xcb', '\x90', '\xca', '\x94', 'nomenew']], 10, True)

['Menominee', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Menominee', 'Language', 'codes', 'ISO', '639-2', 'alg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Menominee', 'Language', 'codes', 'ISO', '639-3', 'mez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Urantian', 'Spoken', 'in', 'Urartu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Urantian', 'Language', 'codes', 'ISO', '639-3', 'xur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inuvialuktun', 'Spoken', 'in', 'Canada', '(', 'Northwest', 'Territories', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada', '\\(', 'Northwest', 'Territories', '\\)']], 8, True)

['Inuvialuktun', 'Region', 'North', 'America']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Inuvialuktun', 'Official', 'language', 'in', 'Northwest', 'Territories', '(', 'Canada', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Northwest', 'Territories', '\\(', 'Canada', '\\)']], 9, True)

['Inuvialuktun', 'Language', 'codes', 'ISO', '639-2', 'iku']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inuvialuktun', 'Language', 'codes', 'ISO', '639-3', 'ikt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inuvialuktun', 'Language', 'codes', 'ISO', '639-1', 'iu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Palestinian', 'Spoken', 'in', 'Palestine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Palestinian', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Palestinian', 'Spoken', 'in', 'Jordan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Palestinian', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']], 6, True)

['Palestinian', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Palestinian', 'Language', 'codes', 'ISO', '639-3', 'ajp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Semigallian', 'Spoken', 'in', 'Latvia', 'and', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Semigallian', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Semigallian', 'Language', 'codes', 'ISO', '639-3', 'xzm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Extremaduran', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Extremaduran', 'Region', 'Autonomous', 'community', 'of', 'Extremadura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Extremaduran', 'Language', 'codes', 'ISO', '639-3', 'ext']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lezgian', 'Region', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Region', 'western', 'Caspian', 'Sea', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western', 'Caspian', 'Sea', 'coast']], 6, True)

['Lezgian', 'Region', 'central', 'Caucasus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Spoken', 'in', 'Russia', 'and', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Lezgian', 'Spoken', 'in', 'also', 'spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'also',

'spoken', 'in', 'Georgia']]], 7, True)

['Lezgian', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Spoken', 'in', 'Turkmenistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Spoken', 'in', 'Ukraine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lezgian', 'Spoken', 'in', 'and', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Lezgian', 'Pronunciation', 'lez', '\xc9', '\xa1', 'i', 't', '\xca', '\x83', '\xca', '\xbc', 'al']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'lez', '\xc9', '\xa1', 'i', 't', '\xca', '\x83', '\xca', '\xbc', 'al']], 12, True)

['Lezgian', 'Language', 'codes', 'ISO', '639-2', 'lez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lezgian', 'Language', 'codes', 'ISO', '639-3', 'lez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lushootseed', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lushootseed', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Lushootseed', 'Language', 'codes', 'ISO', '639-3', 'lut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Punic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Punic', 'Language', 'codes', 'ISO', '639-3', 'xpu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['N'Ko', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['N'Ko', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['N'Ko', 'Region', 'West', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['N'Ko', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['N'Ko', 'Language', 'codes', 'ISO', '639-2', 'nqo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['N'Ko', 'Language', 'codes', 'ISO', '639-3', 'nqo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fang', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fang', 'Spoken', 'in', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'the', 'Congo']]], 7, True)

```

['Fang', 'Spoken', 'in', 'Equatorial', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Fang', 'Spoken', 'in', 'Gabon', 'and', 'S', '\xc3', '\xa3', 'o', 'Tom',
'\xc3', '\xa9', 'and', 'Pr', '\xc3', '\xad', 'ncipe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Gabon',
'and', 'S', '\xc3', '\xa3', 'o', 'Tom', '\xc3', '\xa9', 'and', 'Pr',
'\xc3', '\xad', 'ncipe']]], 17, True)

['Fang', 'Language', 'codes', 'ISO', '639-2', 'fan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Fang', 'Language', 'codes', 'ISO', '639-3', 'fan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Swiss', 'Spoken', 'in', 'Switzerland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Swiss', 'Spoken', 'in', 'northeast', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Swiss', 'Spoken', 'in', 'northern', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Swiss', 'Regulated', 'by']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Swiss', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Uripiv', 'Total', 'speakers', '9', ',', '000', 'for', 'dialect',
'continuum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'for', 'dialect', 'continuum']]], ['NUM', '9'], ['NUM',
'000']], 9, True)

['Uripiv', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Uripiv', 'Region', 'Malakula']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Uripiv', 'Language', 'codes', 'ISO', '639-2', 'map']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Uripiv', 'Language', 'codes', 'ISO', '639-3', 'upv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laz', 'Spoken', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Laz', 'Spoken', 'in', 'Georgia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Laz', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Laz', 'Total', 'speakers', '33', ',', '250', 'in', '1980']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '33'], ['NUM', '250']], 8, True)

['Laz', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Georgian', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Laz', 'Language', 'codes', 'ISO', '639-3', 'lzz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mingrelian', 'Region', 'Samegrelo', ',', 'Abkhazia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Mingrelian', 'Spoken', 'in', 'Georgia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',



'Type: ANY']]], 4, True)

['Mingrelian', 'Writing', 'system', 'Georgian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mingrelian', 'Language', 'codes', 'ISO', '639-3', 'xmf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Moksha', 'Spoken', 'in', 'Russia', 'and', 'as', 'a', 'minority', 'language', 'also', 'in', 'Armenia', ',', 'Australia', 'and', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Russia', 'and', 'as', 'a', 'minority', 'language', 'also', 'in', 'Armenia', ',', 'Australia', 'and', 'the', 'United', 'States']]], 18, True)

['Moksha', 'Region', 'European', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Moksha', 'Official', 'language', 'in', 'Mordovia', '(', 'Russia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Mordovia', '\\\\(', 'Russia', '\\\\)']]], 8, True)

['Moksha', 'Regulated', 'by', 'Mordovian', 'Research', 'Institute', 'of', 'Language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Mordovian', 'Research', 'Institute', 'of', 'Type: ANY']]], 8, True)

['Moksha', 'Regulated', 'by', 'Literature']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Moksha', 'Regulated', 'by', 'History', 'and', 'Economics']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'History', 'and', 'Economics']]], 6, True)

['Moksha', 'Language', 'codes', 'ISO', '639-2', 'mdf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Moksha', 'Language', 'codes', 'ISO', '639-3', 'mdf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Svan', 'Total', 'speakers', 'less', 'than', '30', ',', '000', ';']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'less', 'than', 'NUM', ',', 'NUM', ';']], ['NUM', '30'], ['NUM', '000']],  
9, True)

['Svan', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Svan', 'Pronunciation', 'lu', '\xca', '\x83', 'nu', 'nin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'lu',  
'\xca', '\x83', 'nu', 'nin']]], 7, True)

['Svan', 'Language', 'codes', 'ISO', '639-3', 'sva']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Me\xc3\xa4nkieli', 'Spoken', 'in', 'Sweden', ',', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Me\xc3\xa4nkieli', 'Region', 'Torne', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Me\xc3\xa4nkieli', 'Official', 'language', 'in', 'Recognized',  
'minority', 'language', 'of', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Recognized', 'minority', 'language', 'of', 'Sweden']]], 9, True)

['Me\xc3\xa4nkieli', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Me\xc3\xa4nkieli', 'Language', 'codes', 'ISO', '639-3', 'fit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Han'er", 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Han'er", 'Region', 'Medieval', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
["Han'er", 'Writing', 'system', 'Seal', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Han'er", 'Writing', 'system', 'Clerical', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Han'er", 'Writing', 'system', 'Kaishu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Han'er", 'Writing', 'system', 'Semi-cursive', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Han'er", 'Writing', 'system', 'Grass', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Han'er", 'Writing', 'system', 'Phagspa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Han'er", 'Writing', 'system', 'Hangul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Han'er", 'Official', 'language', 'in', 'Yuan', 'Dynasty']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
["Han'er", 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
["Han'er", 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Arawak', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arawak', 'Spoken', 'in', 'French', 'Guiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'Type: ANY']]], 5, True)

['Arawak', 'Spoken', 'in', 'Guyana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arawak', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arawak', 'Region', 'Guianas', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arawak', 'Region', 'Arawak', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arawak', 'Language', 'codes', 'ISO', '639-3', 'arw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khariboli', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Khariboli', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khariboli', 'Writing', 'system', 'Devanagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Khariboli', 'Language', 'codes', 'ISO', '639-2', 'hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Khariboli', 'Language', 'codes', 'ISO', '639-3', 'hin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khariboli', 'Language', 'codes', 'ISO', '639-1', 'hi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

```
['Philippine', 'Region', 'Zamboanga', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Philippine', 'Region', 'Zamboanga', 'del', 'Norte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Philippine', 'Region', 'Zamboanga', 'Sibugay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Philippine', 'Region', 'Zamboanga', 'del', 'Sur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Philippine', 'Region', 'Basilan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Philippine', 'Region', 'Cavite', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Philippine', 'Region', 'Ternate']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Philippine', 'Region', 'Cavite']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Philippine', 'Region', 'Cotabato']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Philippine', 'Region', 'Davao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Philippine', 'Region', 'Jolo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Philippine', 'Region', 'Tawi-Tawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Philippine', 'Region', 'Semporna', 'in', 'Sabah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Philippine', 'Region', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Philippine', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Philippine', 'Writing', 'system', 'Latin', 'and', 'Spanish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Philippine', 'Official', 'language', 'in', 'Recognised', 'as', 'a',  
'regional', 'language', 'in', 'the', 'Philippines', ';', 'recognized',  
'as', 'a', 'minority', 'language', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Recognised', 'as', 'a', 'regional', 'language', 'in', 'the',  
'Philippines', ';', 'recognized', 'as', 'a', 'minority', 'language', 'in',  
'Malaysia']]], 20, True)

['Philippine', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Philippine', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Philippine', 'Language', 'codes', 'ISO', '639-3', 'cbk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Philippine', 'Language', 'codes', 'Linguasphere', '51-AAC-ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Massachusetts', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Massachusett', 'Region', 'Southeast', 'Massachusetts', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Massachusett', 'Language', 'codes', 'ISO', '639-3', 'wam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Greenlandic', 'Spoken', 'in', 'Greenland', ',', 'Kingdom', 'of',  
'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Greenland', ',', 'Kingdom', 'of', 'Denmark']]], 8, True)

['Greenlandic', 'Region', 'North', 'America']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Greenlandic', 'Official', 'language', 'in', 'Greenland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Greenlandic', 'Language', 'codes', 'ISO', '639-2', 'kal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Greenlandic', 'Language', 'codes', 'ISO', '639-3', 'kal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Greenlandic', 'Language', 'codes', 'ISO', '639-1', 'kl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Seychellois', 'Spoken', 'in', 'Seychelles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Seychellois', 'Official', 'language', 'in', 'Seychelles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Seychellois', 'Regulated', 'by', 'Lenstiti', 'Kreol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Seychellois', 'Language', 'codes', 'ISO', '639-3', 'crs']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Riograndenser', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Riograndenser', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)
```

```
['Riograndenser', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Riograndenser', 'Language', 'codes', 'ISO', '639-3', 'hrx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Venetic', 'Spoken', 'in', 'Northeastern', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Venetic', 'Language', 'codes', 'ISO', '639-3', 'xve']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jin', 'Total', 'speakers', '45', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '45']], 5, True)
```

```
['Jin', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Jin', 'Region', 'most', 'of', 'Shanxi', 'province', ';', 'central', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Hebei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'most', 'of', 'Shanxi', 'province', ';', 'central', 'Inner', 'Mongolia', ';', 'parts', 'of', 'Hebei']], 14, True)
```

```
['Jin', 'Region', 'Henan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Henan']], 3, True)
```

```
['Jin', 'Region', 'Shaanxi']:
```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Jin', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Jin', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Jin', 'Language', 'codes', 'ISO', '639-3', 'cjl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jin', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bouyei', 'Spoken', 'in', 'China', '(', 'Guizhou', ',', 'Yunnan', ',', 'and', 'Sichuan', 'Provinces', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'China', '\\\\(', 'Guizhou', ',', 'Yunnan', ',', 'and', 'Sichuan', 'Provinces', '\\\\)']], 13, True)

['Bouyei', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bouyei', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Bouyei', 'Language', 'codes', 'ISO', '639-3', 'pcc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inuktitut', 'Total', 'speakers', '35', ',', '000', '(', 'approx', '.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'approx', '.', '\\\\)']], ['NUM', '35'], ['NUM', '000']], 10, True)

['Inuktitut', 'Writing', 'system', 'Inuktitut', 'syllabics']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Inuktitut', 'Official', 'language', 'in', 'Nunavut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Inuktitut', 'Official', 'language', 'in', 'Nunavik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Inuktitut', 'Official', 'language', 'in', 'Northwest', 'Territories']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Inuktitut', 'Official', 'language', 'in', 'Nunatsiavut', '(', 'Canada', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Nunatsiavut', '\\(', 'Canada', '\\)']]], 8, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-2', 'iku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-3', 'iku', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'iku', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-3', 'individual', 'code']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Inuktitut', 'Language', 'codes', 'ISO', '639-1', 'iu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Hiligaynon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hiligaynon', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Hiligaynon', 'Official', 'language', 'in', 'Regional', 'language', 'in',

'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Hiligaynon', 'Language', 'codes', 'ISO', '639-2', 'hil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hiligaynon', 'Language', 'codes', 'ISO', '639-3', 'hil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sedang', 'Spoken', 'in', 'Vietnam', 'and', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Sedang', 'Region', 'Kontum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Sedang', 'Region', 'Quang', 'Nam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sedang', 'Region', 'Quang', 'Ngai', '(', 'Vietnam', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Quang',  
'Ngai', '\\\\(', 'Vietnam', '\\\\)']]], 7, True)

['Sedang', 'Language', 'codes', 'ISO', '639-3', 'sed']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kapampangan', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kapampangan', 'Region', 'Central', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kapampangan', 'Writing', 'system', 'Historically', 'written', 'in',  
'Matuang', 'Kudlitan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Kapampangan', 'Official', 'language', 'in', 'Regional', 'language', 'in',

```

'the', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Kapampangan', 'Language', 'codes', 'ISO', '639-2', 'pam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kapampangan', 'Language', 'codes', 'ISO', '639-3', 'pam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Milanese', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Milanese', 'Total', 'speakers', 'n', '/', 'a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'n',
 '/', 'a']]], 6, True)

['Milanese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',
'\x93', 'Lombard']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Gallo', 'Spoken', 'in', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gallo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Gallo', 'Language', 'codes', 'ISO', '639-3', 'fra-gal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gallo', 'Language', 'codes', 'Linguasphere', '51-AAA-hb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Middle', 'Spoken', 'in', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Middle', 'Language', 'codes', 'ISO', '639-2', 'frm']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'frm']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Judeo-Arabic', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Judeo-Arabic', 'Spoken', 'in', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2',  
'\x80', '\x93']]], 6, True)

['Judeo-Arabic', 'Region', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Judeo-Arabic', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-2', 'jrb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'jrb', '\xe2',  
'\x80', '\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'jrb', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)

['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'individual',  
'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Guern\xc3\xa9siais', 'Spoken', 'in', 'Guernsey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Guern\xc3\xa9siais', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',

'\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Guern\x3\xa9siais', 'Language', 'codes', 'Linguasphere', '51-AAA-hc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Ingrian', 'Total', 'speakers', '500']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ingrian', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ingrian', 'Region', 'Ingria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ingrian', 'Language', 'codes', 'ISO', '639-3', 'izh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Elfdalian/\xc3\x96vdalian', 'Spoken', 'in', 'Sweden']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Elfdalian/\xc3\x96vdalian', 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Elfdalian/\xc3\x96vdalian', 'Regulated', 'by', 'R', '\xc3', '\xa5', '\xc3', '\xb0', 'dj', '\xc3', '\xa4', 'rum', ',', '\xc3', '\xa4', 'lvdalska', 'spr', '\xc3', '\xa5', 'kr', '\xc3', '\xa5', 'det']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'R', '\xc3', '\xa5', '\xc3', '\xb0', 'dj', '\xc3', '\xa4', 'rum', ',', '\xc3', '\xa4', 'lvdalska', 'spr', '\xc3', '\xa5', 'kr', '\xc3', '\xa5', 'det']], 23, True)

['Elfdalian/\xc3\x96vdalian', 'Regulated', 'by', '(', 'semi-official', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', '\(', 'semi-official', '\)']], 6, True)

['Elfdalian/\xc3\x96vdalian', 'Language', 'codes', 'ISO', '639-3', 'qer',

'\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qer', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Bats', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Bats', 'Region', 'Zemo-Alvani', 'in', 'Kakheti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Bats', 'Language', 'codes', 'ISO', '639-3', 'bbl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Andalusian', 'Spoken', 'in', 'Al-Andalus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Andalusian', 'Region', 'Iberian', 'Peninsula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Andalusian', 'Language', 'codes', 'ISO', '639-3', 'xaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Egyptian', 'Spoken', 'in', 'Egypt', 'and', 'a', 'few', 'other',  
'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Egypt',  
'and', 'a', 'few', 'other', 'countries']], 9, True)

['Egyptian', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Egyptian', 'Pronunciation', '\xcb', '\x88', 'm', '\xc9', '\x91', 's',  
'\xcb', '\xa4', '\xc9', '\xbe', 'i']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'm', '\xc9', '\x91', 's', '\xcb', '\xa4', '\xc9', '\xbe', 'i']],  
13, True)

['Egyptian', 'Language', 'codes', 'ISO', '639-3', 'arz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Yemeni', 'Spoken', 'in', 'Yemen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Yemeni', 'Spoken', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Yemeni', 'Spoken', 'in', 'Somaliland', 'de', 'facto', 'state', 'not',  
'currently', 'recognized', 'internationally']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Somaliland', 'de', 'facto', 'state', 'not', 'currently', 'recognized',  
'internationally']]], 11, True)
```

```
['Yemeni', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Yemeni', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Yemeni', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Yemeni', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Washo', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Washo', 'Region', 'Nevada', '(', 'Washoe', 'County', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nevada', '\\  
(', 'Washoe', 'County', '\\\')']]], 7, True)
```

```
['Washo', 'Language', 'codes', 'ISO', '639-2', 'was']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Washo', 'Language', 'codes', 'ISO', '639-3', 'was']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```



['Levantine', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Levantine', 'Spoken', 'in', 'Lebanon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Levantine', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Levantine', 'Spoken', 'in', 'Palestine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Levantine', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Levantine', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Levantine', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Levantine', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['(Eastern)', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(Eastern)', 'Language', 'codes', 'ISO', '639-3', 'huc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x81Xegwi', 'Language', 'extinction', '1988']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['\xc7\x81Xegwi', 'Spoken', 'in', 'South', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['\xc7\x81Xegwi', 'Language', 'codes', 'ISO', '639-3', 'xeg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x80Xam', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['\xc7\x80Xam', 'Language', 'codes', 'ISO', '639-3', 'xam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Seroa', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Seroa', 'Language', 'codes', 'ISO', '639-3', 'kqu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ju\xc7\x80\xca\xbcchoan', 'Spoken', 'in', 'Namibia', ',', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Ju\xc7\x80\xca\xbcchoan', 'Region', 'near', 'border', 'with', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'near', 'border', 'with', 'Angola']]], 6, True)

['Ju\xc7\x80\xca\xbcchoan', 'Language', 'codes', 'ISO', '639-3', 'ktz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["K'iche'", 'Region', 'Central', 'highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["K'iche'", 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["K'iche'", 'Pronunciation', '/', 'k', '\xca', '\xbc', 'it', '\xca', '\x83', 'e', '\xca', '\x94', '/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/',

'k', '\xca', '\xbc', 'it', '\xca', '\x83', 'e', '\xca', '\x94', '/' ] ] ] ], 13, True)

["K'iche'", 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY'] ] ] ], 5, True)

["K'iche'", 'Regulated', 'by', 'Academia', 'de', 'Lenguas', 'Mayas', 'de', 'Guatemala', '(' , 'ALMG', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Academia', 'de', 'Lenguas', 'Mayas', 'de', 'Guatemala', '\\(' , 'ALMG', '\\)' ] ] ] ], 12, True)

["K'iche'", 'Recognised', 'minority', 'language', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'Type: ANY'] ] ] ], 6, True)

["K'iche'", 'Language', 'codes', 'ISO', '639-3', 'quc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'] ] ], ['NUM', '639-3'] ], 6, True)

['Jiwarli', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY'] ] ] ], 5, True)

['Jiwarli', 'Language', 'codes', 'ISO', '639-3', 'djl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'] ] ], ['NUM', '639-3'] ], 6, True)

['Warrungu', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'] ] ] ], 4, True)

['Warrungu', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'] ] ] ], 4, True)

['Warrungu', 'Language', 'codes', 'ISO', '639-3', 'wrg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'] ] ], ['NUM', '639-3'] ], 6, True)

['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Spoken', 'in', 'Namibia', ',', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY'] ] ] ], 6, True)

['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Region', 'Okavango', 'and',  
'Ovamboland', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Okavango',  
'and', 'Ovamboland', 'Territory']]], 6, True)

['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Language', 'codes', 'ISO', '639-3',  
'oun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbc0', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbc0', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbc0', 'Language', 'codes', 'ISO',  
'639-3', 'aue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Naro', 'Spoken', 'in', 'Botswana', ',', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Naro', 'Language', 'codes', 'ISO', '639-3', 'nhr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['G\xc7\x80wi', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['G\xc7\x80wi', 'Total', 'speakers', '2500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['G\xc7\x80wi', 'Language', 'codes', 'ISO', '639-3', 'gwj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kxoe', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kxoe', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kxoe', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kxoe', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Kxoe', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kxoe', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kxoe', 'Language', 'codes', 'ISO', '639-3', ':', 'xuu', '\xe2', '\x80', '\x93', 'Khwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'xuu', '\xe2', '\x80', '\x93', 'Khwe']], ['NUM', '639-3']], 11, True)

['Kxoe', 'Language', 'codes', 'ISO', '639-3', 'hnh', '\xe2', '\x80', '\x93', '\xc7', '\x81', 'Ani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hnh', '\xe2', '\x80', '\x93', '\xc7', '\x81', 'Ani']], ['NUM', '639-3']], 12, True)

['G\xc7\x81ana', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['G\xc7\x81ana', 'Total', 'speakers', '2000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['G\xc7\x81ana', 'Language', 'codes', 'ISO', '639-3', 'gnk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsoa', 'Spoken', 'in', 'Botswana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsoa', 'Total', 'speakers', '9300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsoa', 'Dialects', 'Hiechware']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tsoa', 'Dialects', 'Cua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tsoa', 'Dialects', 'Cire', 'Cire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsoa', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tsoa', 'Language', 'codes', 'ISO', '639-3', ':', 'hio', '\xe2', '\x80', '\x93', 'Tsoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'hio', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM', '639-3']], 11, True)

['Tsoa', 'Language', 'codes', 'ISO', '639-3', 'tyu', '\xe2', '\x80', '\x93', 'Kua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tyu', '\xe2', '\x80', '\x93', 'Kua']]], ['NUM', '639-3']], 10, True)

["A'tong", 'Spoken', 'in', 'India', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

["A'tong", 'Region', 'India', ',', 'region', 'of', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', ',', 'Type: ANY', 'of', 'Type: ANY']]], 7, True)

["A'tong", 'Official', 'language', 'in', 'No', 'official', 'status']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'No', 'official', 'status']]], 7, True)

```
["A'tong", 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
["A'tong", 'Language', 'codes', 'ISO', '639-3', 'aot']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['A-Pucikwar', 'Total', 'speakers', 'extinct', '(', 'disputed', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'extinct', '\\(', 'disputed', '\\)']]], 7, True)
```

```
['A-Pucikwar', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['A-Pucikwar', 'Region', 'Andaman', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['A-Pucikwar', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['A-Pucikwar', 'Language', 'codes', 'ISO', '639-3', ':', 'apq', '\\xe2',  
'\\x80', '\\x93', 'A-Pucikwar']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'apq', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 11, True)
```

```
['A-Pucikwar', 'Language', 'codes', 'ISO', '639-3', 'gac', '\\xe2', '\\x80',  
'\\x93', 'Great', 'Andamanese', 'creole']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'gac', '\\xe2', '\\x80', '\\x93', 'Great', 'Andamanese',  
'creole']]], ['NUM', '639-3']], 12, True)
```

```
['Aari', 'Total', 'speakers', '129', ',', '350', 'monolinguals']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'monolinguals']]], ['NUM', '129'], ['NUM', '350']], 7,  
True)
```

```
['Aari', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aari', 'Region', 'north', 'central', 'Omo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Aari', 'Language', 'codes', 'ISO', '639-3', 'aiw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Abanyom', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Abanyom', 'Region', 'Cross', 'River', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Abanyom', 'Language', 'codes', 'ISO', '639-2', 'bnt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Abanyom', 'Language', 'codes', 'ISO', '639-3', 'abm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Abau', 'Region', 'Sandaun']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Abau', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Abau', 'Language', 'codes', 'ISO', '639-3', 'aau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Oirat', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Oirat', 'Region', 'Xinjiang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Oirat', 'Region', 'Gansu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```



ANY']]], 3, True)

['Oirat', 'Region', 'Qinghai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Oirat', 'Region', 'Kalmykia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Oirat', 'Region', 'Khovd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Oirat', 'Region', 'Uvs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Oirat', 'Region', 'Bayan-', '\xc3', '\x96', 'lgii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Bayan-', '\xc3', '\x96', 'lgii']]], 6, True)

['Oirat', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oirat', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oirat', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oirat', 'Language', 'codes', 'ISO', '639-2', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Oirat', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oirat', 'Language', 'codes', 'ISO', '639-3', ':', 'xal', '\xe2', '\x80', '\x93', 'Oirat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'xal', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',

'639-3']], 11, True)

['Oirat', 'Language', 'codes', 'ISO', '639-3', 'xwo', '\xe2', '\x80', '\x93', 'Written', 'Oirat']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'xwo', '\xe2', '\x80', '\x93', 'Written', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Salar', 'Region', 'Qinghai', ',', 'Gansu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Salar', 'Spoken', 'in', 'China']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Salar', 'Writing', 'system', 'Arabic', 'alphabet', ',', 'Chinese', 'characters']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Salar', 'Language', 'codes', 'ISO', '639-3', 'slr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'China']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Region', 'Gansu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'ybe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuman', 'Region', 'Simbu', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kuman', 'Region', 'from', 'Kundiawa', 'to', 'beyond', 'Kerowagi', 'in', 'the', 'west', 'and', 'Gembogl', 'in', 'the', 'north']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'from', 'Kundiawa', 'to', 'beyond', 'Kerowagi', 'in', 'the', 'west', 'and', 'Gembogl', 'in', 'the', 'north']], 15, True)

['Kuman', 'Region', 'at', 'the', 'foot', 'of', 'Mount', 'Wilhelm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'at', 'the',  
'foot', 'of', 'Mount', 'Wilhelm']]], 8, True)

['Kuman', 'Language', 'codes', 'ISO', '639-3', 'kue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lule', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lule', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lule', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Lule', 'Language', 'codes', 'ISO', '639-3', 'ule']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Achang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Achang', 'Language', 'codes', 'ISO', '639-3', 'acn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kurama', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kurama', 'Region', 'Kano', 'and', 'Kaduna', 'states']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kano', 'and',  
'Kaduna', 'states']]], 6, True)

['Kurama', 'Language', 'codes', 'ISO', '639-3', 'krh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sungor', 'Spoken', 'in', 'Chad', 'and', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Sungor', 'Region', 'western', 'Sudan', 'and', 'eastern', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western', 'Sudan', 'and', 'eastern', 'Chad']]], 7, True)

['Sungor', 'Language', 'codes', 'ISO', '639-3', 'sjg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sungor', 'Language', 'codes', 'Linguasphere', '05-PEA-aa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Kurrama', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kurrama', 'Language', 'codes', 'ISO', '639-3', 'vku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Taiwanese', 'Region', 'Taiwan', 'Area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taiwanese', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taiwanese', 'Writing', 'system', 'Traditional', 'Chinese', 'characters']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Taiwanese', 'Official', 'language', 'in', 'Republic', 'of', 'China', '(', 'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'China', '\\\\(', 'Taiwan', '\\\\)']]], 10, True)

['Taiwanese', 'Regulated', 'by', 'National', 'Languages', 'Committee', '(', 'Ministry', 'of', 'Education', ', ', 'ROC', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'National', 'Languages', 'Committee', '\\\\(', 'Ministry', 'of', 'Education',

',', 'ROC', '\\\')]]], 13, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\(', 'B', '\\\')']], ['NUM', '639-2']], 9, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\(', 'T', '\\\')']], ['NUM', '639-2']], 9, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-3', 'cmn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taiwanese', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Oro', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Oro', 'Total', 'speakers', '5']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Oro', 'Language', 'codes', 'ISO', '639-3', 'orw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Canadian', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Canadian', 'Dialects', 'Canadian', 'Ukrainian']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Canadian', 'Language', 'codes', 'ISO', '639-2', 'ukr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Canadian', 'Language', 'codes', 'ISO', '639-3', 'ukr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Canadian', 'Language', 'codes', 'ISO', '639-1', 'uk']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bohemian', 'Spoken', 'in', 'Czech', 'Republic']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Bohemian', 'Total', 'speakers', 'Extinct']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bohemian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Lingua', 'sistemfrater', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',
'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Lingua', 'Date', 'founded', '1957']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Lingua', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Lingua', 'Created', 'by', 'Pham', 'Xuan', 'Thai']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Pham',
'Xuan', 'Type: ANY']]], 6, True)

['Lingua', 'Category', '(', 'sources', ')', 'A', 'posteriori', 'language',
'with', 'a', 'predominantly', 'Romance', 'word', 'stock', 'and', 'an',
'Asian-based', 'grammar']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)', 'A', 'posteriori', 'language', 'with', 'a',
'predominantly', 'Romance', 'word', 'stock', 'and', 'an', 'Asian-based',
'grammar']]], 18, True)

['Lingua', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Chimariko', 'Spoken', 'in', 'USA']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Chimariko', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chimariko', 'Language', 'codes', 'ISO', '639-3', 'cid']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chimariko', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'Chimariko']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Pre-contact', 'distribution', 'of', 'Type: ANY']]], ['NUM', '639-3']], 9, True)

['Polish', 'Signed', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polish', 'Total', 'signers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yinggarda', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yinggarda', 'Language', 'codes', 'ISO', '639-3', 'ya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mentawai', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mentawai', 'Language', 'codes', 'ISO', '639-3', 'mwv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chilean', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chilean', 'Total', 'speakers', '17', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'NUM', 'million']], ['NUM', '17']], 5, True)

['Chilean', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Chilean', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Chuj', 'Spoken', 'in', 'Guatemala']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chuj', 'Language', 'codes', 'ISO', '639-3', 'cac']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Solano', 'Spoken', 'in', 'NE', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Solano', 'Language', 'codes', 'ISO', '639-3', 'xso']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Timucua', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Timucua', 'Dialects', 'Acuera']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Timucua', 'Dialects', 'Aqua', 'Fresca']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Timucua', 'Dialects', 'Aqua', 'Salada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Timucua', 'Dialects', 'Itafi', '/', 'Icafui']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Itafi', '/',  
'Icafui']]], 5, True)



```

['Timucua', 'Dialects', 'Mocama']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Dialects', 'Oconi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Dialects', 'Potano']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Dialects', 'Tawasa', '(', '?', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
'\\(', 'Type: ANY', '\\)']]], 6, True)

['Timucua', 'Dialects', 'Timucua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Dialects', 'Tucururu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Dialects', 'Yufera']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Region', 'Florida']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Timucua', 'Region', 'Southeastern', 'Georgia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Timucua', 'Region', 'Eastern', 'Alabama']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Timucua', 'Pronunciation', 'ti', '\\xcb', '\\x88', 'mu', '\\xcb', '\\x90',
'kwa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ti',
'\\xcb', '\\x88', 'mu', '\\xcb', '\\x90', 'kwa']]], 9, True)

['Timucua', 'Language', 'codes', 'ISO', '639-3', 'tjm']:

```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Timucua', 'Language', 'codes', 'ISO', '639-3', 'if', 'it', 'were', 'truly', 'Timucua']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'if', 'it', 'were', 'truly', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Timucua', 'Language', 'codes', 'ISO', '639-3', 'would', 'have', 'been', 'geographically', 'isolated', 'in', 'Alabama']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'would', 'have', 'been', 'geographically', 'isolated', 'in', 'Type: ANY']], ['NUM', '639-3']], 12, True)
```

```
['Tamahaq', 'Spoken', 'in', 'Algeria']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tamahaq', 'Spoken', 'in', 'Libya']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tamahaq', 'Spoken', 'in', 'Niger']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tamahaq', 'Region', 'Sahara']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Tamahaq', 'Language', 'codes', 'ISO', '639-2', 'tmh']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Tamahaq', 'Language', 'codes', 'ISO', '639-3', 'thv']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wappo', 'Spoken', 'in', 'United', 'States']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Wappo', 'Region', 'Alexander', 'Valley', ',', 'California', '(', 'formerly', ')']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alexander',
```

'Valley', ',', 'California', '\\(', 'formerly', '\\)']]], 9, True)

['Wappo', 'Language', 'codes', 'ISO', '639-3', 'wao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Juba', 'Spoken', 'in', 'Southern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Juba', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']]], 6, True)

['Juba', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Juba', 'Language', 'codes', 'ISO', '639-3', 'pga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iwaidja', 'Total', 'speakers', '150']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Iwaidja', 'Language', 'codes', 'ISO', '639-3', 'ibd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yana', 'Language', 'extinction', '1916']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yana', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yana', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Yana', 'Dialects', 'Yahi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

```

['Yana', 'Language', 'codes', 'ISO', '639-3', 'ynn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['M\xc3\xa1ku', 'Spoken', 'in', 'Brazil', '\xe2', '\x80', '\x93',
'Venezuela', 'border']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Brazil',
'\xe2', '\x80', '\x93', 'Venezuela', 'border']]], 9, True)

['M\xc3\xa1ku', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93', 'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['M\xc3\xa1ku', 'Language', 'codes', 'ISO', '639-3', '099']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guinea-Bissau', 'Spoken', 'in', 'Guinea-Bissau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Guinea-Bissau', 'Spoken', 'in', 'Senegal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Guinea-Bissau', 'Total', 'speakers', '159', ',', '000', ';', '600', ',', '000', 'as', 'second', 'language', 'users']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', ';', 'NUM', ',', 'NUM', 'as', 'second', 'language',
'users']]], ['NUM', '159'], ['NUM', '000'], ['NUM', '600'], ['NUM', '000']],
14, True)

['Guinea-Bissau', 'Language', 'codes', 'ISO', '639-2', 'cpp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Guinea-Bissau', 'Language', 'codes', 'ISO', '639-3', 'pov']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guinea-Bissau', 'Language', 'codes', 'Linguasphere', '51-AAC-ab']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

```

```
['Diu', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Diu', 'Total', 'speakers', '180']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Diu', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Diu', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Diu', 'Language', 'codes', 'Linguasphere', '51-AAC-aga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Kitanemuk', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Kitanemuk', 'Region', 'Southern', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kitanemuk', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Kitanemuk', 'Language', 'codes', 'ISO', '639-3', 'qe8']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Papi\xc3\xa1', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Papi\xc3\xa1', 'Language', 'codes', 'ISO', '639-2', 'cpp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```

['Papi\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Papi\xc3\xa1', 'Language', 'codes', 'Linguasphere', '51-AAC-ahd']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

['Portugis', 'Spoken', 'in', 'Indonesia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Portugis', 'Language', 'codes', 'ISO', '639-2', 'cpp']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Portugis', 'Language', 'codes', 'ISO', '639-3', 'tmg']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bidau', 'Spoken', 'in', 'Bidau']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Bidau', 'Spoken', 'in', 'Dili']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Bidau', 'Spoken', 'in', 'East', 'Timor']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Bidau', 'Language', 'codes', 'ISO', '639-2', 'cpp']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bidau', 'Language', 'codes', 'ISO', '639-3', 'tvv']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bidau', 'Language', 'codes', 'Linguasphere', '51-AAC-ahf']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

['Dutch', 'Signed', 'in', 'Netherlands']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dutch', 'Language', 'codes', 'ISO', '639-2', 'sgn-NL']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Aklanon', 'Region', 'Aklan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Aklanon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aklanon', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Aklanon', 'Language', 'codes', 'ISO', '639-3', 'akl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Khuzestani', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khuzestani', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Khuzestani', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Khuzestani', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Khuzestani', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Bodo', 'Spoken', 'in', 'India', ',', 'with', 'a', 'few', 'small',

'communities', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
,',', 'with', 'a', 'few', 'small', 'communities', 'in', 'Type: ANY']]], 12,  
True)

['Bodo', 'Language', 'codes', 'ISO', '639-3', 'brx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kven', 'Spoken', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kven', 'Official', 'language', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Kven', 'Language', 'codes', 'ISO', '639-2', 'fin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kven', 'Language', 'codes', 'ISO', '639-3', 'fkv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kven', 'Language', 'codes', 'ISO', '639-1', 'fi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Ifugao', 'Spoken', 'in', 'Ifugao', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ifugao', 'Spoken', 'in', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ifugao', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ifugao', 'Language', 'codes', 'ISO', '639-3', 'ifb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bontoc', 'Spoken', 'in', 'Philippines', ',,', 'Mountain', 'Province']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Philippines', ',', 'Mountain', 'Province']]], 7, True)

['Bontoc', 'Language', 'codes', 'ISO', '639-3', 'bnc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karkin', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Karkin', 'Language', 'codes', 'ISO', '639-3', 'krb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Plains', 'Signed', 'in', 'USA', 'and', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'USA',  
'and', 'Canada']]], 6, True)

['Plains', 'Total', 'signers', 'Few']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Quzhou', 'dialect', 'Total', 'speakers', 'approx', '.', '2', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '2']], 8, True)

['Quzhou', 'Pronunciation', 'd', '\\xca', '\\x92', 'y', 't', '\\xc9', '\\x95',  
'i', '\\xc9', '\\xaf', '\\xca', '\\x94', 'u', '\\xc9', '\\x91']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'd',  
'\\xca', '\\x92', 'y', 't', '\\xc9', '\\x95', 'i', '\\xc9', '\\xaf', '\\xca',  
'\\x94', 'u', '\\xc9', '\\x91']]], 17, True)

['Quzhou', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Quzhou', 'Language', 'codes', 'ISO', '639-3', 'wu-chh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Italiano', 'centrale', 'Total', 'speakers', 'approx', '.', '5',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'centrale', 'Total',

'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '5']], 8, True)

['Italiano', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Italiano', 'Region', 'Lazio', 'and', 'east']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Italiano', 'Language', 'codes', 'ISO', '639-2', 'ita']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Italiano', 'Language', 'codes', 'ISO', '639-3', 'ita-cen', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ita-cen', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Italiano', 'Language', 'codes', 'ISO', '639-1', 'it']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Nisga\xe2\x80\x99a', 'Total', 'speakers', '700']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nisga\xe2\x80\x99a', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nisga\xe2\x80\x99a', 'Region', 'northwest', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Nisga\xe2\x80\x99a', 'Language', 'codes', 'ISO', '639-2', 'tsi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Nisga\xe2\x80\x99a', 'Language', 'codes', 'ISO', '639-3', 'ncg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alur', 'Spoken', 'in', 'Orientale', 'Province', '(', 'Democratic', 'Republic', 'of', 'Congo', ')', 'and', 'Uganda']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Orientale', 'Province', '\\(', 'Democratic', 'Republic', 'of', 'Congo', '\\)', 'and', 'Uganda']]], 13, True)

['Alur', 'Language', 'codes', 'ISO', '639-3', 'alz']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saanich', 'Total', 'speakers', '20']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Saanich', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Saanich', 'Language', 'codes', 'ISO', '639-3', 'str']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Puquina', 'Spoken', 'in', 'south', 'shore', 'of', 'Lake', 'Titicaca']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'south', 'shore', 'of', 'Lake', 'Titicaca']]], 8, True)

['Puquina', 'Language', 'codes', 'ISO', '639-3', 'puq']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Spoken', 'in', 'Mongolia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Classical', 'Spoken', 'in', 'China']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Classical', 'Spoken', 'in', 'Russia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Classical', 'Total', 'speakers', 'extinct']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'cmg']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nakanai', 'Language', 'codes', 'ISO', '639-3', 'nak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Me'en", 'Total', 'speakers', '5', ',', '770', 'of', 'Bodi', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'of', 'Bodi', 'dialect']], ['NUM', '5'], ['NUM',  
'770']], 9, True)

["Me'en", 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

["Me'en", 'Region', 'Eastern', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

["Me'en", 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

["Me'en", 'Language', 'codes', 'ISO', '639-3', 'mym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chaha', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chaha', 'Language', 'codes', 'ISO', '639-3', 'sgw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Semai', 'Spoken', 'in', 'Peninsular', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Semai', 'Language', 'codes', 'ISO', '639-3', 'sea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Libido', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```

['Libido', 'Region', 'Gurage']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Libido', 'Region', 'Hadiya']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Libido', 'Region', 'Kembata', 'regions']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Libido', 'Language', 'codes', 'ISO', '639-3', 'liq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maria', 'Language', 'codes', 'ISO', '639-3', 'mds']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Seto', 'Spoken', 'in', 'Estonia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Seto', 'Region', 'Southern', 'Estonia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Seto', 'Language', 'codes', 'ISO', '639-3', 'vro']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Takelma', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Takelma', 'Region', 'Oregon', ',', 'Rogue', 'Valley', 'along', 'the',
'middle', 'course', 'of', 'the', 'Rogue', 'River']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Oregon', ',',
'Rogue', 'Valley', 'along', 'the', 'middle', 'course', 'of', 'the',
'Rogue', 'River']]], 14, True)

['Takelma', 'Language', 'codes', 'ISO', '639-3', 'tkm']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['West', 'Region', 'Central', 'highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['West', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['West', 'Pronunciation', '/', 'k', '\xca', '\xbc', 'it', '\xcd', '\xa1', '\xca', '\x83', 'e', '\xca', '\x94', '/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', 'k', '\xca', '\xbc', 'it', '\xcd', '\xa1', '\xca', '\x83', 'e', '\xca', '\x94', '/']], 15, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'qut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ottawa', 'Total', 'speakers', '8', ',', '000', 'or', 'less', '(', 'estimated', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'or', 'less', '\\(', 'estimated', '\\)']], ['NUM', '8'], ['NUM', '000']], 11, True)

['Ottawa', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY', 'S']], 7, True)

['Ottawa', 'Region', 'Ontario']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ottawa', 'Region', 'Michigan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ottawa', 'Region', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ottawa', 'Language', 'codes', 'ISO', '639-3', 'otw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Israeli', 'Signed', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Israeli', 'Language', 'codes', 'ISO', '639-2', 'sgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tayart,', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tayart,', 'Region', 'Sahara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tayart,', 'Language', 'codes', 'ISO', '639-2', 'tmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tayart,', 'Language', 'codes', 'ISO', '639-3', 'thz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sirion\xc3\xbb', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sirion\xc3\xbb', 'Total', 'speakers', '520']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sirion\xc3\xbb', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sirion\xc3\xbb', 'Language', 'codes', 'ISO', '639-3', ':', 'srq', '\xe2', '\x80', '\x93', 'Sirion', '\xc3', '\xbb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'srq', '\xe2', '\x80', '\x93', 'Sirion', '\xc3', '\xbb']], ['NUM', '639-3']], 13, True)

['Sirion\xc3\xbb', 'Language', 'codes', 'ISO', '639-3', 'yuq', '\xe2', '\x80', '\x93', 'Yuqui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'yuq', '\xe2', '\x80', '\x93', 'Yuqui']], ['NUM', '639-3']], 10, True)

['Buyeo', 'Spoken', 'in', 'Buyeo', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Buyeo', 'Dialects', 'Goguryeo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Buyeo', 'Dialects', '?', 'Baekje']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Buyeo', 'Region', 'Manchuria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Buyeo', 'Language', 'codes', 'ISO', '639-3', 'xpy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Voro', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Voro', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Voro', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Voro', 'Language', 'codes', 'ISO', '639-3', 'vor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Philistine', 'Spoken', 'in', 'Philistia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Philistine', 'Spoken', 'in', 'Philistines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Philistine', 'Language', 'codes', 'ISO', '639-3', 'None']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Z\xc3\xa1paro', 'Spoken', 'in', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Z\xc3\xa1paro', 'Spoken', 'in', 'Ecuador']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Z\xc3\xa1paro', 'Language', 'codes', 'ISO', '639-3', 'zro']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['New', 'Total', 'speakers', 'Non-native', 'speakers', ':', '20', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Non-native', 'speakers', ':', 'NUM', 'million']], ['NUM', '20']], 8, True)

['New', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['New', 'Spoken', 'in', 'Austria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['New', 'Spoken', 'in', 'Liechtenstein']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['New', 'Spoken', 'in', 'Switzerland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['New', 'Official', 'language', 'in', 'Austria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['New', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)

['New', 'Official', 'language', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['New', 'Official', 'language', 'in', 'Liechtenstein']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['New', 'Official', 'language', 'in', 'Luxembourg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['New', 'Official', 'language', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['New', 'Official', 'language', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['New', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['New', 'Language', 'codes', 'ISO', '639-2', 'ger', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ger', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['New', 'Language', 'codes', 'ISO', '639-2', 'deu', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'deu', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['New', 'Language', 'codes', 'ISO', '639-3', 'deu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['New', 'Language', 'codes', 'ISO', '639-1', 'de']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Seediq', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Seediq', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Seediq', 'Region', 'eastern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'eastern']]],  
3, True)

['Seediq', 'Region', 'and', 'coastal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Seediq', 'Language', 'codes', 'ISO', '639-3', 'trv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hunde', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'Congo']]], 7, True)

['Hunde', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hunde', 'Language', 'codes', 'ISO', '639-3', 'hke']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wirangu', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wirangu', 'Language', 'codes', 'ISO', '639-3', 'wiw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Krio', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Krio', 'Region', 'West', 'Kalimantan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Krio', 'Dialects', 'Seputan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Krio', 'Language', 'codes', 'ISO', '639-3', 'xke']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['P\xc3\xa1lez', 'Spoken', 'in', 'Colombia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['P\xc3\xa1lez', 'Region', 'Andes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['P\xc3\xa1lez', 'Language', 'codes', 'ISO', '639-3', 'pbb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Koti', 'Spoken', 'in', 'Mozambique']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koti', 'Language', 'codes', 'ISO', '639-3', 'eko']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sonjo', 'Spoken', 'in', 'Tanzania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sonjo', 'Region', 'Arusha']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sonjo', 'Region', 'Ngorongoro', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sonjo', 'Language', 'codes', 'ISO', '639-3', 'soz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mbo', 'Spoken', 'in', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2', '\x80', '\x93']]], 6, True)

['Mbo', 'Region', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

```

['Mbo', 'Language', 'codes', 'ISO', '639-3', 'mbo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hwana', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Hwana', 'Region', 'Adamawa', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Hwana', 'Language', 'codes', 'ISO', '639-3', 'hwo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Azari', 'Spoken', 'in', 'Iran', '(', 'Persia', ')', ',', 'Azerbaijan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iran',
'\\(', 'Persia', '\\)', ',', 'Azerbaijan']]], 9, True)

['Azari', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Cueva', 'Spoken', 'in', 'Panama']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cueva', 'Region', 'Cueva', 'people']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cueva', 'Language', 'codes', 'ISO', '639-3', 'cuk', '\\xe2', '\\x80',
'\\x93', '(', 'erroneous', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'cuk', '\\xe2', '\\x80', '\\x93', '\\(', 'erroneous', '\\)']],
['NUM', '639-3']], 12, True)

['Colonial', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Colonial', 'Region', 'State', 'of', 'Louisiana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

```

```

['Colonial', 'Language', 'codes', 'ISO', '639-2', 'fra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Colonial', 'Language', 'codes', 'ISO', '639-3', 'fra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Colonial', 'Language', 'codes', 'ISO', '639-1', 'fr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Abinomn', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Abinomn', 'Region', 'Papua', ',', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Abinomn', 'Language', 'codes', 'ISO', '639-3', 'bsa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Isirawa', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Isirawa', 'Region', 'Papua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Isirawa', 'Language', 'codes', 'ISO', '639-3', 'srl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karkar-Yuri', 'Region', 'Sandaun', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karkar-Yuri', 'Region', 'Amanab', 'District']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karkar-Yuri', 'Region', 'along', 'the', 'Papua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',

```

'Type: ANY', 'Type: ANY']]], 5, True)

['Karkar-Yuri', 'Region', 'Indonesia', 'border', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Indonesia',  
'border', '.', 'Coordinates']]], 6, True)

['Karkar-Yuri', 'Language', 'codes', 'ISO', '639-3', 'yuj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yal\x3\xab,', 'Region', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yal\x3\xab,', 'Language', 'codes', 'ISO', '639-3', 'nce']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Busa', 'Region', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Busa', 'Region', 'Amanab', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Busa', 'Region', 'west', 'of', 'Namia', '.', '3', 'villages', '.',  
'Yare', 'is', 'north', 'and', 'east']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'west', 'of',  
'Namia', '.', 'NUM', 'villages', '.', 'Yare', 'is', 'north', 'and',  
'east']]], ['NUM', '3']], 14, True)

['Busa', 'Region', 'Biaka', 'is', 'northwest', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Biaka', 'is',  
'northwest', '.', 'Coordinates']]], 7, True)

['Busa', 'Language', 'codes', 'ISO', '639-3', 'bhf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sulka', 'Spoken', 'in', 'New', 'Britain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Sulka', 'Language', 'codes', 'ISO', '639-3', 'sua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kol', 'Spoken', 'in', 'New', 'Britain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Kol', 'Language', 'codes', 'ISO', '639-3', 'kol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuot,', 'Region', 'New', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kuot,', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kuot,', 'Language', 'codes', 'ISO', '639-3', 'kto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Porome', 'Language', 'codes', 'ISO', '639-3', 'prm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Burmeso', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Burmeso', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Burmeso', 'Language', 'codes', 'ISO', '639-3', 'bzu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yuracar\xc3\xa1', 'Total', 'speakers', '2500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yuracar\xc3\xa1', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



```
['Yuracar\\xc3\\xa9', 'Region', 'Cochabamba', 'Department']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Yuracar\\xc3\\xa9', 'Language', 'codes', 'ISO', '639-3', 'yuz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tubar', 'Spoken', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tubar', 'Total', 'speakers', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tubar', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-2']], 8, True)
```

```
['Tubar', 'Language', 'codes', 'ISO', '639-3', 'tbu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Riverense', 'Portu', '\\xc3', '\\xb1', 'ol', 'Total', 'speakers',  
'approximately', '100', ',', '000', 'people']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Portu', '\\xc3', '\\xb1',  
'ol', 'Total', 'speakers', 'approximately', 'NUM', ',', 'NUM', 'people']]],  
['NUM', '100'], ['NUM', '000']], 12, True)
```

```
['Riverense', 'Pronunciation', 'portu', '\\xcb', '\\x88', '\\xc9', '\\xb2',  
'ol']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'portu',  
'\\xcb', '\\x88', '\\xc9', '\\xb2', 'ol']]], 8, True)
```

```
['Riverense', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Riverense', 'Language', 'codes', 'Linguasphere', '51-AAA-am']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Spokil', 'Total', 'speakers', 'none']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Spokil', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY', 'language']]], 7, True)

['Spokil', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary',  
'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'int', '.', 'auxiliary', 'languages']]], 9, True)

['Spokil', 'Category', '(', 'purpose', ')', 'Spokil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)'], 'Type: ANY']]], 6, True)

['Spokil', 'Created', 'by', 'Adolphe', 'Nicolas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Spokil', 'Category', '(', 'sources', ')', 'mixed', 'a', 'priori', '/',  
'a', 'posteriori']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)'], 'mixed', 'a', 'priori', '/', 'a', 'posteriori']]], 11,  
True)

['Spokil', 'Setting', 'and', 'usage', 'international', 'auxiliary',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Spokil', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Nkore', 'Standard', 'forms', 'Kitara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nkore', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nkore', 'Language', 'codes', 'ISO', '639-3', 'nyn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['An\\xc3\\xaam', 'Total', 'speakers', '550']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['An\xc3\xaam', 'Region', 'New', 'Britain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['An\xc3\xaam', 'Language', 'codes', 'ISO', '639-3', 'anz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ata', 'Region', 'New', 'Britain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ata', 'Language', 'codes', 'ISO', '639-3', 'ata']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Y\xc3\xa9l\xc3\xae', 'Region', 'Louiade', 'Archipelago']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Y\xc3\xa9l\xc3\xae', 'Language', 'codes', 'ISO', '639-3', 'yle']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gitxsan', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gitxsan', 'Language', 'codes', 'ISO', '639-2', 'tsi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gitxsan', 'Language', 'codes', 'ISO', '639-3', 'git']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bora', 'Spoken', 'in', 'Peru', ',', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Bora', 'Language', 'codes', 'ISO', '639-3', 'boa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Tacana', 'Spoken', 'in', 'Bolivia', 'Ethnicity']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Tacana', 'Spoken', 'in', '5', ',', '000']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Tacana', 'Language', 'codes', 'ISO', '639-3', 'tna']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Phong', 'Spoken', 'in', 'Vietnam']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Phong', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Phong', 'Language', 'codes', 'ISO', '639-3', 'pnx']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bahrani', 'Writing', 'system', 'Arabic', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Bahrani', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Bahrani', 'Language', 'codes', 'ISO', '639-3', 'abv']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Upper', 'Region', 'Alaska', '(', 'middle', 'Yukon', 'River', ',', '
Koyukuk', 'River', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\
(', 'middle', 'Yukon', 'River', ',', 'Koyukuk', 'River', '\\)']]], 11,
True)
```

```
['Upper', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```

['Upper', 'Language', 'codes', 'ISO', '639-2', 'ath']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Upper', 'Language', 'codes', 'ISO', '639-3', 'kuu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tolowa', 'Spoken', 'in', 'USA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tolowa', 'Region', 'southwest', 'Oregon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tolowa', 'Language', 'codes', 'ISO', '639-3', 'tol']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsetsaut', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tsetsaut', 'Region', 'Northern', 'British', 'Columbia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['Tsetsaut', 'Language', 'codes', 'ISO', '639-3', 'txc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yuki', 'Spoken', 'in', 'Eel', 'River', 'area', '(', 'formerly', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Eel',
'River', 'area', '\\(', 'formerly', '\\)']], 9, True)

['Yuki', 'Language', 'codes', 'ISO', '639-3', 'yuk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Konkow', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Konkow', 'Region', 'California']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Konkow', 'Language', 'codes', 'ISO', '639-3', 'mjd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tause', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tause', 'Total', 'speakers', '350']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tause', 'Language', 'codes', 'ISO', '639-3', 'tad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chalon', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', '\\\\(', 'Type: ANY', '\\\\)']]], 8, True)

['Chalon', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chalon', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chalon', 'Language', 'codes', 'ISO', '639-3', 'cst']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wambule', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wambule', 'Language', 'codes', 'ISO', '639-3', 'wme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsotsitaal', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Tsotsitaal', 'Total', 'speakers']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Tsotsitaal', 'Language', 'codes', 'ISO', '639-3', 'either']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tsotsitaal', 'Language', 'codes', 'ISO', '639-3', ':', 'fly', '\xe2',
'\x80', '\x93', 'Tsotsitaal']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'fly', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 11, True)
```

```
['Tsotsitaal', 'Language', 'codes', 'ISO', '639-3', 'cmt', '\xe2', '\x80',
'\x93', 'Camtho']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'cmt', '\xe2', '\x80', '\x93', 'Camtho']], ['NUM', '639-3']],
10, True)
```

```
['Pentlatch', 'Language', 'extinction', '1940s']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pentlatch', 'Spoken', 'in', 'Vancouver', 'Island']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Pentlatch', 'Language', 'codes', 'ISO', '639-3', 'ptw']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Culina', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',
'\xc3', '\xba', ',', 'Brazil']]], 8, True)
```

```
['Culina', 'Language', 'codes', 'ISO', '639-3', 'cul']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Deni', 'Spoken', 'in', 'Brazil']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Deni', 'Total', 'speakers', '750']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Deni', 'Language', 'codes', 'ISO', '639-3', 'dny']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kathlamet', 'Spoken', 'in', 'Washington', ',', 'Oregon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Kathlamet', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kathlamet', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Kathlamet', 'Language', 'codes', 'ISO', '639-3', 'qp8', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qp8', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Mysian', 'Spoken', 'in', 'Mysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mysian', 'Language', 'codes', 'ISO', '639-3', 'yms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Picenum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['South', 'Region', 'Marche', ',', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['South', 'Writing', 'system', 'Picene', 'alphabet', 'variants']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Naba', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



'Type: ANY']]], 4, True)

['Naba', 'Region', 'Chad', ',', 'Batha', 'and', 'Chari-Baguirmi',  
'prefectures']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Chad', ',',  
'Batha', 'and', 'Chari-Baguirmi', 'prefectures']]]], 8, True)

['Naba', 'Language', 'codes', 'ISO', '639-3', 'mne']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cof\xc3\xa1n', 'Spoken', 'in', 'Ecuador', ',', 'Colombia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)

['Cof\xc3\xa1n', 'Official', 'language', 'in', 'Ecuador', ':',  
'indigenous', 'languages', 'official', 'in', 'own', 'territories']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Ecuador', ':', 'indigenous', 'languages', 'official', 'in', 'own',  
'territories']]]], 12, True)

['Cof\xc3\xa1n', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]]], 6, True)

['Cof\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'con']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yugambeh', 'Spoken', 'in', 'Queensland', ',', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)

['Yugambeh', 'Total', 'speakers', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Yugambeh', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Yugambeh', 'Language', 'codes', 'ISO', '639-3', 'yub']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uab', 'Region', 'Southeast', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uab', 'Official', 'language', 'in', 'West', 'Timor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Uab', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Uab', 'Language', 'codes', 'ISO', '639-3', 'aoz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka-Bea', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Bea', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Bea', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Bea', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Aka-Bea', 'Language', 'codes', 'ISO', '639-3', 'abj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka-Bale', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Bale', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Bale', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Bale', 'Language', 'codes', 'ISO', '639-3', 'acl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oroqen', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Oroqen', 'Region', 'Inner', 'Mongolia', 'and', 'Heilongjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Inner',  
'Mongolia', 'and', 'Heilongjiang']], 6, True)

['Oroqen', 'Language', 'codes', 'ISO', '639-3', 'orh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Makalero', 'Spoken', 'in', 'East', 'Timor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Makalero', 'Region', 'Timor', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Makalero', 'Region', 'eastern', 'end', 'around', 'Baucau', 'and',  
'inland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'eastern',  
'end', 'around', 'Baucau', 'and', 'inland']], 8, True)

['Makalero', 'Region', 'west', 'of', 'Fataluku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Makalero', 'Region', 'from', 'northern', 'to', 'southern', 'coast', 'in',  
'a', 'dialect', 'chain', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'from',  
'northern', 'to', 'southern', 'coast', 'in', 'a', 'dialect', 'chain', '.',  
'Coordinates']], 13, True)

['Makalero', 'Language', 'codes', 'ISO', '639-3', 'mkz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Babine\xe2\x80\x93Witsuwit'en", 'Dialects', 'Babine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY']]], 3, True)

["Babine\xe2\x80\x93Witsuwit'en", 'Dialects', "Witsuwit'en"]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

["Babine\xe2\x80\x93Witsuwit'en", 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Babine\xe2\x80\x93Witsuwit'en", 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Babine\xe2\x80\x93Witsuwit'en", 'Language', 'codes', 'ISO', '639-3', 'bcr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ulch', 'Total', 'speakers', '730', 'according', 'to', 'some', 'authors', ',', '2', ',', '000', 'according', 'to', 'others']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'according', 'to', 'some', 'authors', ',', 'NUM', ',', 'NUM', 'according', 'to', 'others']], ['NUM', '730'], ['NUM', '2'], ['NUM', '000']], 15, True)

['Ulch', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ulch', 'Region', 'Russian', 'Far', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ulch', 'Language', 'codes', 'ISO', '639-3', 'ulc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sichuanese', 'Region', 'Sichuan', ',', 'Chongqing', 'and', 'their', 'neighboring', 'provinces']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sichuan', ',', 'Chongqing', 'and', 'their', 'neighboring', 'provinces']]], 9, True)

['Sichuanese', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sichuanese', 'Pronunciation', 'locally', 's', '\xc9', '\xa8', '\xcb', '\xa8', '\xcb', '\xa9', '\xcb', '\xa7', 'ts', '\xca', '\xb0', 'wan', '\xcb', '\xa5', 'xwa', '\xcb', '\xa8', '\xcb', '\xa9', '\xcb', '\xa7']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'locally', 's', '\xc9', '\xa8', '\xcb', '\xa8', '\xcb', '\xa9', '\xcb', '\xa7', 'ts', '\xca', '\xb0', 'wan', '\xcb', '\xa5', 'xwa', '\xcb', '\xa8', '\xcb', '\xa9', '\xcb', '\xa7']]], 25, True)

['Sichuanese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Talysh', 'Total', 'speakers', 'ca', '.', '500', ',', '000', 'to', '1', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'ca', '.', 'NUM', ',', 'NUM', 'to', 'NUM', 'million']], ['NUM', '500'], ['NUM', '000'], ['NUM', '1']], 11, True)

['Talysh', 'Spoken', 'in', 'Iran', ',', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Talysh', 'Writing', 'system', 'Perso-Arabic', 'script', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Perso-Arabic', 'script', 'in', 'Iran']]], 7, True)

['Talysh', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Talysh', 'Language', 'codes', 'ISO', '639-3', 'tly']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Talysh', 'Language', 'codes', 'Linguasphere', '58-AAC-ed']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Arwi', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Arwi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Kissi', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kissi', 'Spoken', 'in', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kissi', 'Spoken', 'in', 'Sierra', 'Leone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kissi', 'Region', 'West', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kissi', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kono', 'Spoken', 'in', 'Sierra', 'Leone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kono', 'Region', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kono', 'Language', 'codes', 'ISO', '639-3', 'kno']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dalecarlian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Dalecarlian', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dalecarlian', 'Region', 'Dalarna', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Dalecarlian', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Dalecarlian', 'Regulated', 'by', 'Dalm', '\xc3', '\xa5', 'lsakademin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Dalm', '\xc3', '\xa5', 'lsakademin']]], 7, True)
```

```
['Dalecarlian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Dalecarlian', 'Language', 'codes', 'ISO', '639-3', 'dlc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tocodede', 'Spoken', 'in', 'East', 'Timor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Tocodede', 'Dialects', 'Keha', '(', 'Keia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\(', 'Type: ANY', '\\)']]], 6, True)
```

```
['Tocodede', 'Language', 'codes', 'ISO', '639-3', 'tkd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['T\xc3\xbcbatulabal', 'Spoken', 'in', 'Kern', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['T\xc3\xbcbatulabal', 'Spoken', 'in', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['T\xc3\xbcbatulabal', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['T\xc3\xbcbatulabal', 'Language', 'codes', 'ISO', '639-3', 'tub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Cham', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cham', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cham', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cham', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cham', 'Spoken', 'in', 'various', 'countries', 'with', 'recent',  
'immigrants']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'various', 'countries', 'with', 'recent', 'immigrants']]], 8, True)

['Cham', 'Pronunciation', 'cam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Cham', 'Official', 'language', 'in', 'none', ',', 'recognised', 'as',  
'a', 'minority', 'language', 'in', 'Cambodia', 'and', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'none', ',', 'recognised', 'as', 'a', 'minority', 'language', 'in',  
'Cambodia', 'and', 'Vietnam']]], 15, True)

['Cham', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Cham', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cham', 'Language', 'codes', 'ISO', '639-3', 'huq', '\xe2', '\x80',  
'\x93', 'Tsat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'huq', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Etchemin', 'Spoken', 'in', 'United', 'States']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Etchemin', 'Region', 'Maine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Etchemin', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Etchemin', 'Language', 'codes', 'ISO', '639-3', 'etc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kpelle', 'Spoken', 'in', 'Liberia', ',', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Kpelle', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Kpelle', 'Language', 'codes', 'ISO', '639-2', 'kpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kpelle', 'Language', 'codes', 'ISO', '639-3', 'kpe', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kpe', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Kpelle', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Kpelle', 'Language', 'codes', 'ISO', '639-3', ':', 'gkp', '\xe2', '\x80', '\x93', 'Guinea', 'Kpelle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'gkp', '\xe2', '\x80', '\x93', 'Guinea', 'Type: ANY']], ['NUM', '639-3']], 12, True)

['Kpelle', 'Language', 'codes', 'ISO', '639-3', 'xpe', '\xe2', '\x80', '\x93', 'Liberia', 'Kpelle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```
'ISO', 'NUM', 'xpe', '\xe2', '\x80', '\x93', 'Liberia', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Zapotec', 'Spoken', 'in', 'Mexico', '(', 'Oaxaca', ',', 'Puebla', ',',  
'Guerrero', ')', ';', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  
'\\(', 'Oaxaca', ',', 'Puebla', ',', 'Guerrero', '\\)', ';', 'USA']]]], 13,  
True)
```

```
['Zapotec', 'Official', 'language', 'in', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['Zapotec', 'Regulated', 'by', 'Instituto', 'Nacional', 'de', 'Lenguas',  
'Ind', '\xc3', '\xad', 'genas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Instituto', 'Nacional', 'de', 'Lenguas', 'Ind', '\xc3', '\xad',  
'genas']]]], 11, True)
```

```
['Zapotec', 'Language', 'codes', 'ISO', '639-2', 'zap']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Zapotec', 'Language', 'codes', 'ISO', '639-3', 'Zapotecan', 'languages',  
'in', 'Oaxaca']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Zapotecan', 'languages', 'in', 'Oaxaca']]], ['NUM',  
'639-3']], 9, True)
```

```
['Bunak', 'Region', 'central', 'Timor']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Bunak', 'Language', 'codes', 'ISO', '639-3', 'bfn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kemak', 'Spoken', 'in', 'East', 'Timor']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)
```

```
['Kemak', 'Language', 'codes', 'ISO', '639-3', 'kem']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Chiwere', 'Spoken', 'in', 'United', 'States']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chiwere', 'Region', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chiwere', 'Region', 'Missouri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Chiwere', 'Region', 'and', 'Kansas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chiwere', 'Language', 'codes', 'ISO', '639-3', 'iow']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'io']], ['NUM', '639-3']], 6, True)

['Kitsai', 'Total', 'speakers', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kitsai', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kitsai', 'Region', 'previously', 'west-central', 'Oklahoma', 'and', 'eastern', 'Texas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'previously', 'west-central', 'Oklahoma', 'and', 'eastern', 'Type: ANY']]], 8, True)

['Kitsai', 'Language', 'codes', 'ISO', '639-3', 'kii']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paumar\xc3\xad', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']]], 8, True)

['Paumar\xc3\xad', 'Total', 'speakers', '700']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Paumar\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'pad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chittagonian', 'Total', 'speakers', '14', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '14']], 5, True)

['Chittagonian', 'Spoken', 'in', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chittagonian', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chittagonian', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chittagonian', 'Region', 'Eastern', 'South', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Chittagonian', 'Language', 'codes', 'ISO', '639-3', 'ctg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nicola', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nicola', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nicola', 'Region', 'British', 'Columbia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Nicola', 'Region', 'Nicola', 'Athapaskans']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nicola', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',

'639-3']], 10, True)

['Nicola', 'Language', 'codes', 'ISO', '639-3', 'qs7']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tagol', 'Spoken', 'in', 'Malaysia', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Tagol', 'Region', 'Borneo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Tagol', 'Language', 'codes', 'ISO', '639-3', 'mvv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yavapai', 'Spoken', 'in', 'Arizona', ',', 'USA', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Arizona', ',', 'USA', 'Ethnicity']], 7, True)

['Yavapai', 'Total', 'speakers', '163']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yavapai', 'Language', 'codes', 'ISO', '639-3', 'yuf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Galoli', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Galoli', 'Language', 'codes', 'ISO', '639-3', 'gal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wetarese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']],  
3, True)

['Wetarese', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Wetarese', 'Region', 'Southeast', 'Asia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Wetarese', 'Language', 'codes', 'ISO', '639-3', 'wet', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'wet', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Wetarese', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Wetarese', 'Language', 'codes', 'ISO', '639-3', ':', 'apx', '\xe2',
'\x80', '\x93', 'Aputai']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'apx', '\xe2', '\x80', '\x93', 'Aputai']], ['NUM',
'639-3']], 11, True)
```

```
['Wetarese', 'Language', 'codes', 'ISO', '639-3', 'wet', '\xe2', '\x80',
'\x93', 'Parai']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'wet', '\xe2', '\x80', '\x93', 'Parai']], ['NUM', '639-3']],
10, True)
```

```
['Wetarese', 'Language', 'codes', 'ISO', '639-3', 'ilw', '\xe2', '\x80',
'\x93', 'Talur']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'ilw', '\xe2', '\x80', '\x93', 'Talur']], ['NUM', '639-3']],
10, True)
```

```
['Wetarese', 'Language', 'codes', 'ISO', '639-3', 'tzn', '\xe2', '\x80',
'\x93', 'Tugun']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'tzn', '\xe2', '\x80', '\x93', 'Tugun']], ['NUM', '639-3']],
10, True)
```

```
['Quichua', 'Spoken', 'in', 'Ecuador', ',', 'Colombia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Quichua', 'Language', 'codes', 'ISO', '639-2', 'que']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Quichua', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Quichua', 'Language', 'codes', 'ISO', '639-1', 'qu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Sidetic', 'Writing', 'system', 'Sidetic', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Sidetic', 'Language', 'codes', 'ISO', '639-3', 'xsd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pisidian', 'Writing', 'system', 'Pisidian', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Pisidian', 'Language', 'codes', 'ISO', '639-3', 'xps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Asturian', 'Region', 'Autonomous', 'Community', 'of', 'Asturias']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'of', 'Type: ANY']], 6, True)
```

```
['Asturian', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Asturian', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Asturian', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['Asturian', 'Language', 'codes', 'ISO', '639-2', 'ast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Asturian', 'Language', 'codes', 'ISO', '639-3', 'ast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Asturian', 'Language', 'codes', 'Linguasphere', '51-AAA-ca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Bezhta', 'Region', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bezhta', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bezhta', 'Pronunciation', '\xcb', '\x88', 'be', '\xca', '\x92', 't',  
'\xcd', '\xa1', '\xc9', '\xac', '\xca', '\xbc', '\xc9', '\x91', 'l',  
'\xc9', '\x91', 's', 'mit', '\xcd', '\xa1', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'be', '\xca', '\x92', 't', '\xcd', '\xa1', '\xc9', '\xac', '\xca',  
'\xbc', '\xc9', '\x91', 'l', '\xc9', '\x91', 's', 'mit', '\xcd', '\xa1',  
's']]], 24, True)

['Bezhta', 'Language', 'codes', 'ISO', '639-3', 'kap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuanua', 'Language', 'codes', 'ISO', '639-3', 'ksd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minigir', 'Language', 'codes', 'ISO', '639-3', 'vmg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Patpatar', 'Dialects', 'Pala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Patpatar', 'Dialects', 'Sokirik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Patpatar', 'Region', 'New', 'Ireland', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)



```

['Patpatar', 'Language', 'codes', 'ISO', '639-3', 'gfk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamviri', 'Total', 'speakers', '5', ',', '500', '(', 'Ethnologue', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '5'], ['NUM',
'500']], 9, True)

['Kamviri', 'Spoken', 'in', 'Afghanistan', ',', 'Pakistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)

['Kamviri', 'Region', 'Bashgal', 'Valley']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kamviri', 'Region', 'Langorbat']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Kamviri', 'Region', 'Badrugal', 'and', 'the', 'Urtsun', 'Valley']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Badrugal',
'and', 'the', 'Urtsun', 'Valley']], 7, True)

['Kamviri', 'Language', 'codes', 'ISO', '639-3', 'xvi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamviri', 'Language', 'codes', 'Linguasphere', '58-ACB-ad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

['Thai', 'Signed', 'in', 'Thailand']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Thai', 'Total', 'signers', '(', 'unknown', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\
(', 'unknown', '\\)']], 6, True)

['Sydney', 'Spoken', 'in', 'New', 'South', 'Wales']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', 'Wales']], 6, True)

['Sydney', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',

```

```

'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Sydney', 'Language', 'codes', 'ISO', '639-3', '0hl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bonan', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Bonan', 'Region', 'Gansu', ',', 'Qinghai']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['Bonan', 'Language', 'codes', 'ISO', '639-3', 'peh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Eastern', 'Region', 'Gansu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'yuy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ogea', 'Total', 'speakers', '700']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Ogea', 'Region', 'Madang', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Ogea', 'Language', 'codes', 'ISO', '639-3', 'eri']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Waray', 'Total', 'speakers', '4']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Waray', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Waray', 'Language', 'codes', 'ISO', '639-3', 'wrz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southwestern', 'Spoken', 'in', 'Sichuan', ',', 'Yunnan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Southwestern', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Southwestern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Southwestern', 'Language', 'codes', 'ISO', '639-3', 'cmn-xin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kiga', 'Standard', 'forms', 'Kitara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kiga', 'Spoken', 'in', 'Uganda', ',', 'Rwanda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Kiga', 'Language', 'codes', 'ISO', '639-3', 'cgg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bagri', 'Spoken', 'in', 'Rajasthan', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Bagri', 'Total', 'speakers', '5', 'million', '(', 'approx', '.', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', '\\\\(', 'approx', '.', '\\\)']], ['NUM', '5']], 9, True)

['Bagri', 'Language', 'codes', 'ISO', '639-3', 'bgq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mewati', 'Spoken', 'in', 'India', '(', 'Mewat', 'region', 'of', 'Haryana', ',', 'Rajasthan', ')', 'and', 'Pakistan', '(', 'mostly', 'in', 'Punjab', 'and', 'Sindh', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', '\\\\(', 'Mewat', 'region', 'of', 'Haryana', ',', 'Rajasthan', '\\\)'], 'and', 'Pakistan', '\\\\(', 'mostly', 'in', 'Punjab', 'and', 'Sindh', '\\\)']]], 21, True)

['Mewati', 'Language', 'codes', 'ISO', '639-2', 'raj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mewati', 'Language', 'codes', 'ISO', '639-3', 'wtm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shekhawati', 'Language', 'codes', 'ISO', '639-2', 'raj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Shekhawati', 'Language', 'codes', 'ISO', '639-3', 'swv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Domari', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Spoken', 'in', 'Egypt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Spoken', 'in', 'other']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Region', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Region', 'North', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Region', 'Caucasus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Domari', 'Region', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Domari', 'Region', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Domari', 'Language', 'codes', 'ISO', '639-3', 'rmt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Godwar', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Godwar', 'Language', 'codes', 'ISO', '639-2', 'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Godwar', 'Language', 'codes', 'ISO', '639-3', 'gdx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romanesco', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Romanesco', 'Region', 'Lazio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Romanesco', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Romanesco', 'Language', 'codes', 'Linguasphere', '51-AAA-rab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Woiwurrung', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Woiwurrung', 'Spoken', 'in', 'Victoria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Media', 'Spoken', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Media', 'Language', 'codes', 'ISO', '639-3', 'mue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Darkinjung', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Darkinjung', 'Spoken', 'in', 'New', 'South', 'Wales', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Wales', ',', 'Australia']], 8, True)

['Jicarilla', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jicarilla', 'Region', 'New', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jicarilla', 'Language', 'codes', 'ISO', '639-2', 'apa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Jicarilla', 'Language', 'codes', 'ISO', '639-3', 'apj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karaj\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karaj\xc3\xa1', 'Region', 'Araguaia', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karaj\xc3\xa1', 'Dialects', 'Java', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Java',  
'\xc3', '\xa9']]], 5, True)

['Karaj\xc3\xa1', 'Dialects', 'Xambio', '\xc3', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Xambio',  
'\xc3', '\xa1']]], 5, True)

['Karaj\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'kpj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wallisian', 'Spoken', 'in', 'Wallis', 'and', 'Futuna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Wallisian', 'Language', 'codes', 'ISO', '639-3', 'wls']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['West', 'Spoken', 'in', 'New', 'Caledonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'uve']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Futunan', 'Language', 'codes', 'ISO', '639-3', 'fud']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Niuafu'ou", 'Spoken', 'in', 'Tonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Niuafu'ou", 'Language', 'codes', 'ISO', '639-3', 'num']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sallaans', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Sallaans', 'Pronunciation', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sallaans', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sallaans', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sallaans', 'Language', 'codes', 'ISO', '639-3', 'sdz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mutsun', '(', 'San', '&']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mutsun', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', '\\(', 'Type: ANY', '\\)']], 8, True)

['Mutsun', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mutsun', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Mutsun', 'Language', 'codes', 'ISO', '639-3', 'css']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Toba', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



['Toba', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Toba', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Toba', 'Language', 'codes', 'ISO', '639-3', 'tob']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyah', 'Dialects', 'Chaiyaphum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyah', 'Dialects', 'Petchabun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyah', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nyah', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nyah', 'Language', 'codes', 'ISO', '639-3', 'cbn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['West', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['West', 'Total', 'speakers', '1500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['West', 'Dialects', 'Aniwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'fut']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zuruah\xc3\xa1', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']], 8, True)

['Zuruah\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'swx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kadiw\xc3\xa9u', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kadiw\xc3\xa9u', 'Total', 'speakers', '1800']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kadiw\xc3\xa9u', 'Language', 'codes', 'ISO', '639-3', 'kbc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mocov\xc3\xad', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mocov\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'moc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuna', 'Spoken', 'in', 'Panama', ',', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Kuna', 'Region', 'San', 'Blas', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kuna', 'Region', 'Panama', ';', 'north', 'coastal', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Panama', ';', 'north', 'coastal', 'region']], 7, True)

['Kuna', 'Region', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kuna', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuna', 'Language', 'codes', 'ISO', '639-3', 'kvn', '\xe2', '\x80',  
'\x93', 'Border', 'Kuna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'kvn', '\xe2', '\x80', '\x93', 'Border', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Ngurmbur', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Ngurmbur', 'Total', 'speakers', '0', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '?']], ['NUM', '0']], 5, True)

['Ngurmbur', 'Language', 'codes', 'ISO', '639-3', 'nrx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ubugarla', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Ubugarla', 'Total', 'speakers', '0', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '?']], ['NUM', '0']], 5, True)

['Ubugarla', 'Language', 'codes', 'ISO', '639-3', 'umr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chontal', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chontal', 'Language', 'codes', 'ISO', '639-3', 'chf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gaagudju', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

```

['Gaagudju', 'Language', 'codes', 'ISO', '639-3', 'gbu']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minkin', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)

['Minkin', 'Total', 'speakers', '0']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Minkin', 'Language', 'codes', 'ISO', '639-3', 'None']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laragiya', 'Spoken', 'in', 'Near', 'Darwin']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Laragiya', 'Spoken', 'in', 'Northern', 'Territory']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Laragiya', 'Spoken', 'in', 'Australia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Laragiya', 'Language', 'codes', 'ISO', '639-3', 'lrg']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Enindhilyagwa', 'Spoken', 'in', 'Northern', 'Territory']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Enindhilyagwa', 'Spoken', 'in', 'Australia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Enindhilyagwa', 'Language', 'codes', 'ISO', '639-3', 'aoi']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Garawa', 'Language', 'codes', 'ISO', '639-3', 'gbc']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kott', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kott', 'Language', 'codes', 'ISO', '639-3', 'zko']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tataviam', 'Spoken', 'in', 'Southern', 'California']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tataviam', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Tataviam', 'Language', 'codes', 'ISO', '639-3', 'qc5']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka', 'Spoken', 'in', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aka', 'Language', 'codes', 'ISO', '639-3', 'soh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ersu', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ersu', 'Writing', 'system', 'Ersu', 'Shaba', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Ersu', 'Language', 'codes', 'ISO', '639-3', 'ers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rengao', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rengao', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Rengao', 'Language', 'codes', 'ISO', '639-3', 'ren']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Fe'fe'', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Fe'fe'', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Fe'fe'', 'Language', 'codes', 'ISO', '639-3', 'fmp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Zyphe', 'Spoken', 'in', 'Burma', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Zyphe', 'Language', 'codes', 'ISO', '639-3', 'zyp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Zumbun', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zumbun', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zumbun', 'Language', 'codes', 'ISO', '639-3', 'jmb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Barupu', 'Language', 'codes', 'ISO', '639-3', 'wra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Koch', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Koch', 'Language', 'codes', 'ISO', '639-3', 'kdq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saka', 'Spoken', 'in', 'Kingdom', 'of', 'Khotan', 'and', 'Tumxuk', '(',  
'today', 'Xinjiang', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Kingdom', 'of', 'Khotan', 'and', 'Tumxuk', '\\(', 'today', 'Xinjiang', '\\  
)']]], 12, True)

['Saka', 'Dialects', 'Khotanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Saka', 'Dialects', 'Tumshuqese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Saka', 'Region', 'Central', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saka', 'Writing', 'system', 'Br', '\\xc4', '\\x81', 'hm', '\\xc4', '\\xab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Br', '\\xc4', '\\x81', 'hm', '\\xc4', '\\xab']]], 9, True)

['Saka', 'Language', 'codes', 'ISO', '639-2', 'kho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Saka', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saka', 'Language', 'codes', 'ISO', '639-3', ':', 'kho', '\\xe2', '\\x80',  
'\\x93', 'Khotanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'kho', '\\xe2', '\\x80', '\\x93', 'Khotanese']]], ['NUM',  
'639-3']], 11, True)

['Saka', 'Language', 'codes', 'ISO', '639-3', 'xtq', '\\xe2', '\\x80',

'\x93', 'Tumshuqese']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'xtq', '\xe2', '\x80', '\x93', 'Tumshuqese']], ['NUM', '639-3']], 10, True)

['Powhatan', 'Spoken', 'in', 'Eastern', 'Virginia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Powhatan', 'Spoken', 'in', 'Southern', 'Maryland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Powhatan', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Powhatan', 'Language', 'codes', 'ISO', '639-3', 'pim']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cotabato', 'Spoken', 'in', 'Sultan', 'Kudarat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Cotabato', 'Spoken', 'in', 'Mindanao']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cotabato', 'Spoken', 'in', 'the', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Cotabato', 'Language', 'codes', 'ISO', '639-2', 'mno']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cotabato', 'Language', 'codes', 'ISO', '639-3', 'mta']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kabiy\xc3\xa8', 'Spoken', 'in', 'Togo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kabiy\xc3\xa8', 'Spoken', 'in', 'Benin']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kabiy\xc3\xa8', 'Spoken', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kabiy\xc3\xa8', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kabiy\xc3\xa8', 'Language', 'codes', 'ISO', '639-3', 'kbp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Walmajarri', 'Spoken', 'in', 'Western', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Walmajarri', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Walmajarri', 'Language', 'codes', 'ISO', '639-3', 'wmt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Djabugay', 'Spoken', 'in', 'Queensland', ',', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Djabugay', 'Language', 'codes', 'ISO', '639-3', 'dyy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inor', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Inor', 'Language', 'codes', 'ISO', '639-3', 'ior']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'io']], ['NUM', '639-3']], 6, True)

['Persian', 'Signed', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Persian', 'Total', 'signers', 'no', 'estimate', 'available']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'no',
'estimate', 'available']]]], 6, True)
```

```
['North', 'Spoken', 'in', 'Picenum']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]]], 4, True)
```

```
['North', 'Region', 'Marche', ',', 'Italy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['North', 'Writing', 'system', 'Picene', 'alphabet', 'variants']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 6, True)
```

```
['North', 'Language', 'codes', 'ISO', '639-3', 'nrp', 'Approximate',
'distribution', 'of', 'languages', 'in', 'Iron', 'Age', 'Italy', 'during',
'the', 'sixth', 'century', 'BC']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'nrp', 'Approximate', 'distribution', 'of', 'languages',
'in', 'Iron', 'Age', 'Italy', 'during', 'the', 'sixth', 'century', 'BC']]],
['NUM', '639-3']], 19, True)
```

```
['Warlmanpa', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']]]], 7, True)
```

```
['Warlmanpa', 'Language', 'codes', 'ISO', '639-3', 'wrl']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['W\xc3\xa8', 'Region', 'Dix-Huit', 'Montagnes', ',', 'Moyen-Cavally']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Dix-Huit',
'Montagnes', ',', 'Moyen-Cavally']]]], 6, True)
```

```
['W\xc3\xa8', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['W\xc3\xa8', 'Language', 'codes', 'ISO', '639-3', ':', 'gxx', '\xe2',
'\x80', '\x93', 'W', '\xc3', '\xa8', 'Southern']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'gxx', '\xe2', '\x80', '\x93', 'W', '\xc3', '\xa8',
'Type: ANY']]], ['NUM', '639-3']], 14, True)
```

['W\\xc3\\xa8', 'Language', 'codes', 'ISO', '639-3', 'wec', '\\xe2', '\\x80', '\\x93', 'W', '\\xc3', '\\xa8', 'Western']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wec', '\\xe2', '\\x80', '\\x93', 'W', '\\xc3', '\\xa8', 'Type: ANY']], ['NUM', '639-3']], 13, True)

['W\\xc3\\xa8', 'Language', 'codes', 'ISO', '639-3', 'wob', '\\xe2', '\\x80', '\\x93', 'W', '\\xc3', '\\xa8', 'Northern']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wob', '\\xe2', '\\x80', '\\x93', 'W', '\\xc3', '\\xa8', 'Type: ANY']], ['NUM', '639-3']], 13, True)

['Nukuoro', 'Spoken', 'in', 'Micronesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nukuoro', 'Region', 'Western', 'Nukuoro', 'Island']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Nukuoro', 'Language', 'codes', 'ISO', '639-3', 'nkr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kunwinjku', 'Total', 'speakers', '1410']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kunwinjku', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kunwinjku', 'Region', 'Northern', 'Territory']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kunwinjku', 'Language', 'codes', 'ISO', '639-3', 'gup']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sakau', 'Total', 'speakers', '4000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakau', 'Spoken', 'in', 'Vanuatu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sakau', 'Language', 'codes', 'ISO', '639-3', 'sku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Doukhobor', 'Spoken', 'in', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Doukhobor', 'Spoken', 'in', 'Saskatchewan', 'and', 'Alberta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Doukhobor', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Doukhobor', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Doukhobor', 'Language', 'codes', 'ISO', '639-2', 'rus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Doukhobor', 'Language', 'codes', 'ISO', '639-3', 'rus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Doukhobor', 'Language', 'codes', 'ISO', '639-1', 'ru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Valencian', 'Region', 'Valencian', 'Community']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Valencian', 'Signed', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Valencian', 'Total', 'signers', '(', 'unknown', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\\\(', 'unknown', '\\\)']]], 6, True)

```
['Vlach', 'Spoken', 'in', 'Serbia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Vlach', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Blin', 'Spoken', 'in', 'central', 'Eritrea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Blin', 'Language', 'codes', 'ISO', '639-2', 'byn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Blin', 'Language', 'codes', 'ISO', '639-3', 'byn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['North', 'Spoken', 'in', 'Vancouver', 'Island']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['North', 'Spoken', 'in', 'British', 'Columbia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['North', 'Spoken', 'in', 'Canada', ';', 'Washington']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',  
';', 'Washington']]], 6, True)
```

```
['North', 'Spoken', 'in', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['North', 'Language', 'codes', 'ISO', '639-3', 'str']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lummi', 'Spoken', 'in', 'Pacific', 'Northwest']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Lummi', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Lummi', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jabo', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Jabo', 'Pronunciation', '\xe2', '\x80', '\x89', '\xc9', '\x9f', '\xca', '\xb1', '\xc9', '\x91', 'bo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xe2', '\x80', '\x89', '\xc9', '\x9f', '\xca', '\xb1', '\xc9', '\x91', 'bo']], 12, True)

['Jabo', 'Spoken', 'in', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jabo', 'Language', 'codes', 'ISO', '639-3', 'grj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ixil', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ixil', 'Region', 'Quich', '\xc3', '\xa9', '(', 'department', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Quich', '\xc3', '\xa9', '\\(', 'department', '\\)']], 8, True)

['Ixil', 'Language', 'codes', 'ISO', '639-3', 'ixl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Achi', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Achi', 'Region', 'Baja', 'Verapaz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Achi', 'Language', 'codes', 'ISO', '639-3', 'acr']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Awakatek', 'Spoken', 'in', 'Guatemala']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Awakatek', 'Region', 'Huehuetenango']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Awakatek', 'Language', 'codes', 'ISO', '639-3', 'agu']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tojolabal', 'Spoken', 'in', 'Mexico']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tojolabal', 'Region', 'Southeast', 'Chiapas']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tojolabal', 'Language', 'codes', 'ISO', '639-3', 'toj']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Itza'", 'Spoken', 'in', 'Guatemala']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
["Itza'", 'Region', 'Pet', '\xc3', '\xa9', 'n']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Pet', '\xc3', '\xa9', 'n']], 6, True)
```

```
["Itza'", 'Language', 'codes', 'ISO', '639-3', 'itz']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mopan', 'Spoken', 'in', 'Belize', ',', 'Guatemala']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)
```

```
['Mopan', 'Language', 'codes', 'ISO', '639-2', 'myn']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```

['Mopan', 'Language', 'codes', 'ISO', '639-3', 'mop']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uspanteko', 'Spoken', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Uspanteko', 'Region', 'Quich', '\xc3', '\xa9', '(', 'department', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Quich',
'\xc3', '\xa9', '\\(', 'department', '\\)']]], 8, True)

['Uspanteko', 'Language', 'codes', 'ISO', '639-3', 'usp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Tz'utujil", 'Spoken', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Tz'utujil", 'Region', 'Western', 'Highlands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Tz'utujil", 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

["Tz'utujil", 'Regulated', 'by', 'Academia', 'de', 'Lenguas', 'Mayas',
'de', 'Guatemala', '(', 'ALMG', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Academia', 'de', 'Lenguas', 'Mayas', 'de', 'Guatemala', '\\(', 'ALMG', '\\
)']]], 12, True)

["Tz'utujil", 'Recognised', 'minority', 'language', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'Type: ANY']]], 6, True)

["Tz'utujil", 'Language', 'codes', 'ISO', '639-3', 'tzj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tektitek', 'Spoken', 'in', 'Guatemala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```



['Tektitek', 'Region', 'Western', 'Highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tektitek', 'Language', 'codes', 'ISO', '639-3', 'ttc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Meru', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Meru', 'Language', 'codes', 'ISO', '639-3', 'mer']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mwotlap', 'Total', 'speakers', '2100']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mwotlap', 'Pronunciation', '\xc5', '\x8b', '\xcd', '\xa1', 'm', '\xca', '\xb7', '\xc9', '\x94', 't', '\xcb', '\x88', 'lap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc5', '\x8b', '\xcd', '\xa1', 'm', '\xca', '\xb7', '\xc9', '\x94', 't', '\xcb', '\x88', 'lap']], 15, True)

['Mwotlap', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mwotlap', 'Language', 'codes', 'ISO', '639-3', 'mlv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Stockholmska', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Stockholmska', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Stockholmska', 'Region', 'Northern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Stockholmska', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Luwati', 'Spoken', 'in', 'Oman', '(', 'Muttrah', 'walled', 'quarter', ', ', 'facing', 'the', 'old', 'harbor', ';', 'Muscat', 'and', 'other', 'cities', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Oman', '\\(', 'Muttrah', 'walled', 'quarter', ', ', 'facing', 'the', 'old', 'harbor', ';', 'Muscat', 'and', 'other', 'cities', '\\)']], 19, True)

['Luwati', 'Total', 'speakers', '5000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Luwati', 'Language', 'codes', 'ISO', '639-3', 'luv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Torwali', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Torwali', 'Writing', 'system', 'Arabic', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Torwali', 'Language', 'codes', 'ISO', '639-3', 'trw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ushoji', 'Spoken', 'in', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ushoji', 'Language', 'codes', 'ISO', '639-3', 'ush']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Munduruku', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Munduruku', 'Language', 'codes', 'ISO', '639-3', 'myu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tharawal', 'Spoken', 'in', 'New', 'South', 'Wales', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Wales', ',', 'Australia']]], 8, True)

['Tharawal', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Tharawal', 'Language', 'codes', 'ISO', '639-3', 'tbh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Muruwari', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Muruwari', 'Language', 'codes', 'ISO', '639-3', 'zmu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saisiyat', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Saisiyat', 'Language', 'codes', 'ISO', '639-3', 'xsy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sop', 'Region', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sop', 'Language', 'codes', 'ISO', '639-3', 'urw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chicomuceltec', 'Spoken', 'in', 'Mexico', '(', 'Chiapas', 'state', ')', ';', 'Guatemala', '(', 'Huehuetenango', 'Dept', '.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', '\\(', 'Chiapas', 'state', '\\)', ';', 'Guatemala', '\\(', 'Huehuetenango', 'Dept', '.', '\\)']]], 15, True)

['Chicomuceltec', 'Pronunciation', 't', '\xca', '\x83', 'ikomusel', '\xcb',

```

'\x88', 'teko']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't',
'\xca', '\x83', 'ikomusel', '\xcb', '\x88', 'teko']]], 9, True)

['Chicomuceltec', 'Language', 'codes', 'ISO', '639-2', 'myn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chicomuceltec', 'Language', 'codes', 'ISO', '639-3', 'cob']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Soso', 'Total', 'speakers', 'over', '1', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'over', 'NUM', 'million']], ['NUM', '1']], 6, True)

['Soso', 'Spoken', 'in', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Soso', 'Spoken', 'in', 'Sierra', 'Leone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Soso', 'Spoken', 'in', 'Guinea', 'Bissau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Soso', 'Region', 'Coastal', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Soso', 'Language', 'codes', 'ISO', '639-2', 'sus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Soso', 'Language', 'codes', 'ISO', '639-3', 'sus', 'Where', 'it', 'is',
'mainly', 'spoken']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'sus', 'Where', 'it', 'is', 'mainly', 'spoken']], ['NUM',
'639-3']], 11, True)

['Kalami', 'Spoken', 'in', 'Pakistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

```
['Kalami', 'Writing', 'system', 'Arabic', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kalami', 'Language', 'codes', 'ISO', '639-3', 'gwc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kalami', 'Language', 'codes', 'Linguasphere', '59-AAC-c']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Huichol', 'Spoken', 'in', 'Mexico', ':', 'parts', 'of', 'the', 'states',  
'of', 'Jalisco']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  

```

```
['Huichol', 'Spoken', 'in', 'Nayarit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Huichol', 'Spoken', 'in', 'Durango']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Huichol', 'Spoken', 'in', 'Zacatecas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Huichol', 'Official', 'language', 'in', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Huichol', 'Regulated', 'by', 'Instituto', 'Nacional', 'de', 'Lenguas',  
'Ind', '\xc3', '\xad', 'genas', '(', 'INALI', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Instituto', 'Nacional', 'de', 'Lenguas', 'Ind', '\xc3', '\xad', 'genas',  
'\xc3', '(', 'INALI', '\xc3', ')']]], 14, True)
```

```
['Huichol', 'Language', 'codes', 'ISO', '639-3', 'hch']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sempan', 'Language', 'codes', 'ISO', '639-3', 'xse']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Allar', 'Spoken', 'in', 'Palakkad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Allar', 'Spoken', 'in', 'Malappuram', 'districts', 'of', 'Kerala', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Malappuram', 'districts', 'of', 'Kerala', 'State']], 8, True)

['Allar', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Allar', 'Total', 'speakers', '350', '(', 'SIL', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\(', 'SIL', '\\)']], ['NUM', '350']], 7, True)

['Allar', 'Language', 'codes', 'ISO', '639-3', 'all']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chocho,', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chocho,', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Chocho,', 'Language', 'codes', 'ISO', '639-3', 'coz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Worimi', 'Spoken', 'in', 'New', 'South', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Wales']], 6, True)

['Worimi', 'Language', 'codes', 'ISO', '639-3', 'kda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gadang', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Gadang', 'Region', 'southwest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Gadang', 'Language', 'codes', 'ISO', '639-3', 'gdk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Khandeshi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Khandeshi', 'Region', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Khandeshi', 'Region', 'Gujarat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Khandeshi', 'Region', 'Madhya', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Khandeshi', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Khandeshi', 'Language', 'codes', 'ISO', '639-3', ':', 'khn', '\xe2',  
'\x80', '\x93', 'Khandeshi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'khn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Khandeshi', 'Language', 'codes', 'ISO', '639-3', 'ahr', '\xe2', '\x80',  
'\x93', 'Ahirani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ahr', '\xe2', '\x80', '\x93', 'Ahirani']], ['NUM',  
'639-3']], 10, True)
```

```
['Khandeshi', 'Language', 'codes', 'ISO', '639-3', 'dhn', '\xe2', '\x80',  
'\x93', 'Dhanki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'dhn', '\xe2', '\x80', '\x93', 'Dhanki']], ['NUM', '639-3']],  
10, True)
```

```
['Susquehannock', 'Language', 'extinction', '1763']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Susquehannock', 'Language', 'codes', 'ISO', '639-3', 'sqn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pitjantjatjara', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Pitjantjatjara', 'Language', 'codes', 'ISO', '639-3', 'pjt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Derung', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Derung', 'Region', 'Yunnan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Derung', 'Language', 'codes', 'ISO', '639-3', 'duu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Grebo', 'Spoken', 'in', 'Liberia', ',', 'Ivory', 'Coast']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Liberia', ',', 'Ivory', 'Type: ANY']]], 7, True)
```

```
['Grebo', 'Language', 'codes', 'ISO', '639-3', 'grb', '\xe2', '\x80',
'\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'grb', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)
```

```
['Grebo', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Grebo', 'Language', 'codes', 'ISO', '639-3', 'gec', '\xe2', '\x80',
'\x93', 'Gboloo', 'Grebo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'gec', '\xe2', '\x80', '\x93', 'Gboloo', 'Type: ANY']],
```



['NUM', '639-3']], 11, True)

['Grebo', 'Language', 'codes', 'ISO', '639-3', 'gry', '\xe2', '\x80', '\x93', 'Barclayville', 'Grebo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gry', '\xe2', '\x80', '\x93', 'Barclayville', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Adnyamathana', 'Spoken', 'in', 'South', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Adnyamathana', 'Language', 'codes', 'ISO', '639-3', 'adt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngaanyatjarra', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ngaanyatjarra', 'Total', 'speakers', '1200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngaanyatjarra', 'Language', 'codes', 'ISO', '639-3', 'ntj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yalarnnga', 'Language', 'extinction', '1980']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yalarnnga', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yalarnnga', 'Language', 'codes', 'ISO', '639-3', 'ylr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pitta', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pitta', 'Language', 'codes', 'ISO', '639-3', 'pit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eggon', 'Spoken', 'in', 'Nasarawa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Eggon', 'Spoken', 'in', 'Kaduna']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Eggon', 'Spoken', 'in', 'Benue', 'state', 'and', 'FCT', 'abuja', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Benue', 'state', 'and', 'FCT', 'abuja', 'in', 'Nigeria']], 10, True)

['Eggon', 'Language', 'codes', 'ISO', '639-3', 'ego']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Severn', 'Ojibwa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']], 7, True)

['Severn', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Severn', 'Region', 'Ontario', ',', 'Manitoba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Severn', 'Language', 'codes', 'ISO', '639-3', 'ojs']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Manam', 'Language', 'codes', 'ISO', '639-3', 'mva']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nalik', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Nalik', 'Language', 'codes', 'ISO', '639-3', 'nal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Garadjari', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Garadjari', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Garadjari', 'Language', 'codes', 'ISO', '639-3', 'gbd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shubi', 'Spoken', 'in', 'Kagera']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Shubi', 'Region', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Shubi', 'Language', 'codes', 'ISO', '639-3', 'suj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Coatl\xc3\xa1n', 'Spoken', 'in', 'Oaxaca', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Oaxaca',  
'in', 'Mexico']]], 6, True)

['Coatl\xc3\xa1n', 'Total', 'speakers', '500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Coatl\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'zps']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yurumangu\xc3\xad', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yurumangu\xc3\xad', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Yurumangu\&#x3\&#xad', 'Language', 'codes', 'ISO', '639-3', 'qic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Salinan', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Salinan', 'Region', 'central', 'coast', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Salinan', 'Language', 'codes', 'ISO', '639-3', 'sln']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shuswap', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shuswap', 'Language', 'codes', 'ISO', '639-3', 'shs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sahaptin', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sahaptin', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Sahaptin', 'Region', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sahaptin', 'Region', 'Oregon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sahaptin', 'Region', 'and', 'Idaho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sahaptin', 'Language', 'codes', 'ISO', '639-3', 'qot', '\&#xe2', '\&#x80',

```

'\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'qot', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)

['Sahaptin', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Sahaptin', 'Language', 'codes', 'ISO', '639-3', ':', 'uma', '\xe2',
'\x80', '\x93', 'Umatilla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'uma', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 11, True)

['Sahaptin', 'Language', 'codes', 'ISO', '639-3', 'waa', '\xe2', '\x80',
'\x93', 'Walla', 'Walla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'waa', '\xe2', '\x80', '\x93', 'Walla', 'Walla']], ['NUM',
'639-3']], 11, True)

['Sahaptin', 'Language', 'codes', 'ISO', '639-3', 'yak', '\xe2', '\x80',
'\x93', 'Yakima']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'yak', '\xe2', '\x80', '\x93', 'Yakima']], ['NUM', '639-3']],
10, True)

['Sahaptin', 'Language', 'codes', 'ISO', '639-3', 'tqn', '\xe2', '\x80',
'\x93', 'Tenino']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'tqn', '\xe2', '\x80', '\x93', 'Tenino']], ['NUM', '639-3']],
10, True)

['Rekhta', 'Spoken', 'in', 'South', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Rekhta', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Rekhta', 'Writing', 'system', 'Nasta', '\xca', '\xbf', 'l', '\xc4',
'\xab', 'q']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Nasta', '\xca', '\xbf', 'l', '\xc4', '\xab', 'q']]], 10, True)

```

['Rekhta', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Rekhta', 'Language', 'codes', 'ISO', '639-3', 'hin-rek']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kata', 'Region', 'One', 'village', 'in', 'the', 'northern', 'part', 'of', 'the', 'island']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'One', 'village', 'in', 'the', 'northern', 'part', 'of', 'the', 'island']]], 11, True)

['Kata', 'Signed', 'in', 'Bali']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ban', 'Signed', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ban', 'Total', 'signers', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Saa', 'Spoken', 'in', 'Pentecost', 'Island', ',', 'Vanuatu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Pentecost', 'Island', ',', 'Vanuatu']]], 7, True)

['Saa', 'Dialects', 'Ponorwal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Saa', 'Dialects', 'Lolatavola']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Saa', 'Dialects', 'Ninebulo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Saa', 'Language', 'codes', 'ISO', '639-3', 'sax']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maninka', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maninka', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maninka', 'Spoken', 'in', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maninka', 'Spoken', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maninka', 'Spoken', 'in', 'Sierra', 'Leone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Maninka', 'Writing', 'system', 'N'Ko', ',', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Maninka', 'Official', 'language', 'in', 'Guinea', ',', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', ',', 'Type: ANY']], 7, True)

['Maninka', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Maninka', 'Language', 'codes', 'ISO', '639-2', 'man']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Maninka', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maninka', 'Language', 'codes', 'ISO', '639-3', 'emk', '\xe2', '\x80', '\x93', 'Maninkaka', ',', 'Eastern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'emk', '\xe2', '\x80', '\x93', 'Maninkaka', ',', 'Type:

ANY']], ['NUM', '639-3']], 12, True)

['Maninka', 'Language', 'codes', 'ISO', '639-3', 'mzj', '\xe2', '\x80', '\x93', 'Manya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mzj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Otuho', 'Spoken', 'in', 'Southern', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Otuho', 'Language', 'codes', 'ISO', '639-3', 'lot']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yimas', 'Total', 'speakers', '300']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yimas', 'Language', 'codes', 'ISO', '639-3', 'yee']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Barrow', 'Spoken', 'in', 'Queensland', ',', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Barrow', 'Language', 'codes', 'ISO', '639-3', 'bpt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dhundhari', 'Spoken', 'in', 'India', '(', 'Dhundhar', 'region', 'of', 'Rajasthan', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', '\\(', 'Dhundhar', 'region', 'of', 'Rajasthan', '\\)']], 10, True)

['Dhundhari', 'Language', 'codes', 'ISO', '639-2', 'raj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dhundhari', 'Language', 'codes', 'ISO', '639-3', 'dhd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngan\xe2\x80\x99gityemerri', 'Spoken', 'in', 'Daly', 'River', 'Mission']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Daly', 'River', 'Mission']]], 6, True)

['Ngan\xe2\x80\x99gityemerri', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ngan\xe2\x80\x99gityemerri', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngan\xe2\x80\x99gityemerri', 'Language', 'codes', 'ISO', '639-3', 'nam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Naasioi', 'Spoken', 'in', 'Kieta', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Naasioi', 'Spoken', 'in', 'Bougainville', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Naasioi', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Naasioi', 'Language', 'codes', 'ISO', '639-3', 'nas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Oksapmin', 'Spoken', 'in', 'Telefomin', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Oksapmin', 'Spoken', 'in', 'Sandaun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oksapmin', 'Dialects', 'Upper', 'Oksapmin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oksapmin', 'Dialects', 'Lower', 'Oksapmin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oksapmin', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Oksapmin', 'Language', 'codes', 'ISO', '639-3', 'opm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mizo', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mizo', 'Spoken', 'in', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mizo', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mizo', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mizo', 'Region', 'Mizoram']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mizo', 'Region', 'Tripura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mizo', 'Region', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mizo', 'Region', 'Manipur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mizo', 'Region', 'Meghalaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mizo', 'Region', 'Nagaland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mizo', 'Official', 'language', 'in', 'Mizoram', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\\\(', 'India', '\\\\)']]], 8, True)

['Mizo', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Mizo', 'Language', 'codes', 'ISO', '639-2', 'lus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Mizo', 'Language', 'codes', 'ISO', '639-3', 'lus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Garó', 'Spoken', 'in', 'India', 'and', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Garó', 'Region', 'Meghalaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Garó', 'Region', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Garó', 'Region', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Garó', 'Official', 'language', 'in', 'Meghalaya', '(', 'India', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', '\\\\(', 'India', '\\\\)']]], 8, True)

['Garó', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Garó', 'Language', 'codes', 'ISO', '639-3', 'grt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kuku-Thaypan', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kuku-Thaypan', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kuku-Thaypan', 'Language', 'codes', 'ISO', '639-3', 'typ']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Meriam', 'Spoken', 'in', 'Murray', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Meriam', 'Spoken', 'in', 'Torres', 'Strait']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Meriam', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Meriam', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Meriam', 'Language', 'codes', 'ISO', '639-3', 'ulk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mer', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Mer', 'Language', 'codes', 'ISO', '639-3', 'mnu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dadibi', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Dadibi', 'Language', 'codes', 'ISO', '639-3', 'mps']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zumaya', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zumaya', 'Language', 'codes', 'ISO', '639-3', 'zuy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warembori', 'Pronunciation', '\xcb', '\x88', 'wa', '\xc9', '\xbe', '\xc9', '\x9b', 'mboi', '\xce', '\xb2', 'o', '\xc9', '\xbe', 'o']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'wa', '\xc9', '\xbe', '\xc9', '\x9b', 'mboi', '\xce', '\xb2', 'o', '\xc9', '\xbe', 'o']], 16, True)

['Warembori', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Warembori', 'Language', 'codes', 'ISO', '639-3', 'wsa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Babuza', 'Spoken', 'in', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Babuza', 'Dialects', 'Poavosa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Babuza', 'Dialects', 'Taokas', '(', 'extinct', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Babuza', 'Language', 'codes', 'ISO', '639-3', 'bzg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dimasa', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dimasa', 'Region', 'Assam', ',', 'Nagaland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dimasa', 'Language', 'codes', 'ISO', '639-3', 'dis']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lavukaleve', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Lavukaleve', 'Region', 'Russell', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lavukaleve', 'Language', 'codes', 'ISO', '639-3', 'lvk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Venture\xc3\xb1o', 'Language', 'codes', 'ISO', '639-3', 'veo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Purari', 'Spoken', 'in', 'Purari', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Purari', 'Spoken', 'in', 'Gulf', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Purari', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Purari', 'Language', 'codes', 'ISO', '639-3', 'iar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mailu', 'Language', 'codes', 'ISO', '639-3', 'mgu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sentani', 'Spoken', 'in', 'Lake', 'Sentani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sentani', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sentani', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sentani', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Sentani', 'Language', 'codes', 'ISO', '639-3', 'set']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Golin', 'Spoken', 'in', 'Gumine', 'District', ',', 'Simbu', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Gumine', 'District', ',', 'Simbu', 'Province']], 8, True)

['Golin', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Golin', 'Language', 'codes', 'ISO', '639-3', 'gvf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wambon', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Wambon', 'Language', 'codes', 'ISO', '639-3', 'wms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maisin', 'Language', 'codes', 'ISO', '639-3', 'mbq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ghari', 'Spoken', 'in', 'Guadalcanal', ',', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Solomon', 'Islands']], 7, True)

['Ghari', 'Language', 'codes', 'ISO', '639-3', 'gri']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oroha', 'Language', 'codes', 'ISO', '639-3', 'ora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Namakura', 'Spoken', 'in', 'North', 'Efate']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Namakura', 'Spoken', 'in', 'Tongoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Namakura', 'Spoken', 'in', 'Tongariki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Namakura', 'Language', 'codes', 'ISO', '639-3', 'nmk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['East', 'Spoken', 'in', 'Ambae', ',', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['East', 'Language', 'codes', 'ISO', '639-3', 'omb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aulua', 'Language', 'codes', 'ISO', '639-3', 'aul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['North', 'Language', 'codes', 'ISO', '639-2', 'map']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'llp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Hadiyya', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hadiyya', 'Region', 'Gurage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Hadiyya', 'Region', 'Hadiya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Hadiyya', 'Region', 'Kembata', 'regions']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hadiyya', 'Region', 'area', 'around', 'Hosaena']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Hadiyya', 'Writing', 'system', 'Ethiopic', '/', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Hadiyya', 'Language', 'codes', 'ISO', '639-3', 'hdy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aghu', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Aghu', 'Language', 'codes', 'ISO', '639-3', 'ahh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tunggare', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Tunggare', 'Language', 'codes', 'ISO', '639-3', 'trt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mixtecan', 'Spoken', 'in', 'Mexico', ';', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  
';', 'USA']]], 6, True)

['Mixtecan', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mixtecan', 'Region', 'Puebla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mixtecan', 'Region', 'Guerrero', ';', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Mixtecan', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Mixtecan', 'Official', 'language', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Mixtecan', 'Regulated', 'by', 'Academy', 'of', 'the', 'Mixtec',  
'Language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Academy', 'of', 'the', 'Mixtec', 'Type: ANY']]], 8, True)

['Mixtecan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Mixtecan', 'Language', 'codes', 'ISO', '639-3', 'Indigenous',  
'languages', 'of', 'Mexico', 'with', 'more', 'than', '100', ',', '000',  
'speakers', '.', 'Mixtec', 'is', 'marked', 'on', 'the', 'map', 'by',  
'dark', 'green']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Indigenous', 'languages', 'of', 'Mexico', 'with', 'more',  
'than', 'NUM', ',', 'NUM', 'speakers', '.', 'Mixtec', 'is', 'marked', 'on',  
'the', 'map', 'by', 'dark', 'green']], ['NUM', '639-3'], ['NUM', '100'],  
['NUM', '000']], 26, True)

['Aghu', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```

['Aghu', 'Language', 'codes', 'ISO', '639-3', 'ggr']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vietnamese', 'Signed', 'in', 'Vietnam']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Vietnamese', 'Total', 'signers', '?']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vietnamese', 'Language', 'codes', 'ISO', '639-3', 'haf', '\xe2', '\x80',
'\x93', 'Haiphong', 'Sign', 'Language']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'haf', '\xe2', '\x80', '\x93', 'Haiphong', 'Sign', 'Type:
ANY']], ['NUM', '639-3']], 12, True)

['Isthmus', 'Spoken', 'in', 'Oaxaca', ',', 'Mexico']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)

['Isthmus', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

['Isthmus', 'Pronunciation', 'd', '\xc3', '\xac', 'd', '\xca', '\x92',
'\xc3', '\xa0', '\xcb', '\x88', 'z', '\xc3', '\xa0']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'd',
'\xc3', '\xac', 'd', '\xca', '\x92', '\xc3', '\xa0', '\xcb', '\x88', 'z',
'\xc3', '\xa0']], 15, True)

['Isthmus', 'Language', 'codes', 'ISO', '639-2', 'zap']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Isthmus', 'Language', 'codes', 'ISO', '639-3', 'zai']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Central-Java']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'East-Java']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'Madura']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Spoken', 'in', 'and', 'Bali']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Standard', 'forms', 'Kawi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'kaw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngkoth', 'Language', 'extinction', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngkoth', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bulu', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bulu', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Bulu', 'Language', 'codes', 'ISO', '639-3', 'bum']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uradhi', 'Language', 'extinction', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uradhi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Uradhi', 'Dialects', 'Atampaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Uradhi', 'Dialects', 'Gudang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Uradhi', 'Dialects', 'Yinwum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Uradhi', 'Dialects', 'Wuthati']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Uradhi', 'Language', 'codes', 'ISO', '639-3', 'urf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wik-Ngatharr', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wik-Ngatharr', 'Language', 'codes', 'ISO', '639-3', 'wik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wurrugu', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', ',', 'Australia']]], 7, True)

['Wurrugu', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Wurrugu', 'Language', 'codes', 'ISO', '639-3', 'wur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ngardi', 'Total', 'speakers', '80', '\xe2', '\x80', '\x93', '100', 'with', 'varying', 'degrees', 'of', 'fluency']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'with', 'varying', 'degrees', 'of', 'fluency']], ['NUM', '80'], ['NUM', '100']], 13, True)

['Goldogrin', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of', 'Tolkien']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of', 'Tolkien']]], 10, True)

['Goldogrin', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Goldogrin', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Goldogrin', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Goldogrin', 'Category', '(', 'purpose', ')', 'Goldogrin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Goldogrin', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Goldogrin', 'Setting', 'and', 'usage', 'The', 'fictional', 'world', 'of', 'Middle-earth']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'The', 'fictional', 'world', 'of', 'Middle-earth']]], 9, True)

['Goldogrin', 'Language', 'codes', 'ISO', '639-2', 'art']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Goldogrin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Jarai', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jarai', 'Spoken', 'in', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jarai', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Jarai', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jarai', 'Official', 'language', 'in', 'none', ',', 'recognised', 'as',  
'a', 'minority', 'language', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'none', ',', 'recognised', 'as', 'a', 'minority', 'language', 'in',  
'Vietnam']]], 13, True)

['Jarai', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Jarai', 'Language', 'codes', 'ISO', '639-3', 'jra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gudang', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gudang', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gudang', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gudang', 'Language', 'codes', 'ISO', '639-3', 'urf', '\xe2', '\x80',  
'\x93', 'Uradhi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'urf', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Wuthati', 'Language', 'extinction', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Wuthati', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wuthati', 'Language', 'codes', 'ISO', '639-3', 'urf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luthigh', 'Language', 'extinction', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Luthigh', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Luthigh', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Luthigh', 'Language', 'codes', 'ISO', '639-3', '081']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yinwum', 'Language', 'extinction', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yinwum', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yinwum', 'Language', 'codes', 'ISO', '639-3', 'urf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mpakwithi', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mpakwithi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



'Type: ANY']]], 4, True)

['Mpakwithi', 'Total', 'speakers', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mpakwithi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Mpakwithi', 'Language', 'codes', 'ISO', '639-3', '0gw-mpa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Atampaya', 'Language', 'extinction', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Atampaya', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Atampaya', 'Language', 'codes', 'ISO', '639-3', 'amz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Linngithigh', 'Language', 'extinction', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Linngithigh', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Linngithigh', 'Language', 'codes', 'ISO', '639-3', 'lnj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alngith', 'Language', 'extinction', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Alngith', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Alngith', 'Language', 'codes', 'ISO', '639-3', 'aid']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anguthimri', 'Language', 'extinction', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Anguthimri', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Anguthimri', 'Dialects', 'Alngith']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Anguthimri', 'Dialects', 'Linngithigh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Anguthimri', 'Dialects', 'Mpakwithi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Anguthimri', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anguthimri', 'Language', 'codes', 'ISO', '639-3', ':', 'aid', '\xe2',  
'\x80', '\x93', 'Alngith']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'aid', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Anguthimri', 'Language', 'codes', 'ISO', '639-3', 'lnj', '\xe2', '\x80',  
'\x93', 'Linngithigh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lnj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Bardi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Bardi', 'Total', 'speakers', '20']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bardi', 'Dialects', 'Jawi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Bardi', 'Language', 'codes', 'ISO', '639-3', 'bcj']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yir-Yoront', 'Spoken', 'in', 'Australia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Yir-Yoront', 'Language', 'codes', 'ISO', '639-3', 'yiy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Diyari', 'Language', 'extinction', '?']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Diyari', 'Spoken', 'in', 'South', 'Australia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Diyari', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Diyari', 'Language', 'codes', 'ISO', '639-3', 'dif']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Northern', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Northern', 'Region', 'Nevada']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Northern', 'Region', 'California']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Northern', 'Region', 'Oregon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Northern', 'Region', 'Idaho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'pao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kalkatungu', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalkatungu', 'Spoken', 'in', 'Mount', 'Isa', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mount', 'Isa', 'area']]], 6, True)

['Kalkatungu', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalkatungu', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalkatungu', 'Language', 'codes', 'ISO', '639-3', 'ktg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tangoa', 'Language', 'codes', 'ISO', '639-2', 'map']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tangoa', 'Language', 'codes', 'ISO', '639-3', 'tgp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Standard', 'Tibetan', 'Total', 'speakers', 'between', '5', 'and', '10', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Tibetan', 'Total', 'speakers', 'between', 'NUM', 'and', 'NUM', 'million']]], ['NUM', '5'],

['NUM', '10']], 9, True)

['Standard', 'Region', 'Tibet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Standard', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Standard', 'Spoken', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Standard', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Standard', 'Writing', 'system', 'Tibetan', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Standard', 'Official', 'language', 'in', 'Tibet', 'Autonomous', 'Region']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Tibet', 'Autonomous', 'Region']], 7, True)

['Standard', 'Regulated', 'by', 'Committee', 'for', 'the', 'Standardisation', 'of', 'the', 'Tibetan', 'Language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Committee', 'for', 'the', 'Standardisation', 'of', 'the', 'Tibetan', 'Type: ANY']], 11, True)

['Standard', 'Language', 'codes', 'ISO', '639-2', 'tib', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tib', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Standard', 'Language', 'codes', 'ISO', '639-2', 'bod', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bod', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Standard', 'Language', 'codes', 'ISO', '639-3', 'bod']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Standard', 'Language', 'codes', 'ISO', '639-1', 'bo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Bundjalung', 'Spoken', 'in', 'New', 'South', 'Wales', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Wales', ',', 'Australia']]], 8, True)

['Bundjalung', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bundjalung', 'Language', 'codes', 'ISO', '639-3', 'bdy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngarinyin', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ngarinyin', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ngarinyin', 'Language', 'codes', 'ISO', '639-3', 'ung']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wunambal', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Wunambal', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Wunambal', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wunambal', 'Language', 'codes', 'ISO', '639-3', ':', 'wub', '\xe2', '\x80', '\x93', 'Wunambal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'wub', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM', '639-3']], 11, True)

['Wunambal', 'Language', 'codes', 'ISO', '639-3', 'wil', '\xe2', '\x80',

```

'\x93', 'Wilawila']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'wil', '\xe2', '\x80', '\x93', 'Wilawila']], ['NUM',
'639-3']], 10, True)

['Wunambal', 'Language', 'codes', 'ISO', '639-3', 'gma', '\xe2', '\x80',
'\x93', 'Gamberre']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'gma', '\xe2', '\x80', '\x93', 'Gamberre']], ['NUM',
'639-3']], 10, True)

['Wunambal', 'Language', 'codes', 'ISO', '639-3', 'gww', '\xe2', '\x80',
'\x93', 'Kwini']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'gww', '\xe2', '\x80', '\x93', 'Kwini']], ['NUM', '639-3']],
10, True)

['Wunambal', 'Language', 'codes', 'ISO', '639-3', 'vmi', '\xe2', '\x80',
'\x93', 'Miwa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'vmi', '\xe2', '\x80', '\x93', 'Miwa']], ['NUM', '639-3']],
10, True)

['Kavalan', 'Spoken', 'in', 'Taiwan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kavalan', 'Language', 'codes', 'ISO', '639-3', 'ckv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngarrindjeri', 'Language', 'extinction', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Ngarrindjeri', 'Spoken', 'in', 'South', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Ngarrindjeri', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Ngarrindjeri', 'Language', 'codes', 'ISO', '639-3', 'nay']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Bahing', 'Spoken', 'in', 'Okhaldhunga', 'district', ',', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Okhaldhunga', 'district', ',', 'Type: ANY']]], 7, True)

['Bahing', 'Official', 'language', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Bahing', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Bahing', 'Language', 'codes', 'ISO', '639-3', 'bhj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kariti\xc3\xa2na', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kariti\xc3\xa2na', 'Language', 'codes', 'ISO', '639-3', 'ktn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyikina', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nyikina', 'Language', 'codes', 'ISO', '639-3', 'nyh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsez', 'Region', 'Southern', 'Dagestan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsez', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsez', 'Pronunciation', 't', '\xcd', '\xa1', 's', '\xc9', '\x9b', 'z', '\xcb', '\x88', 'jas', 'm', '\xc9', '\x9b', 't', '\xcd', '\xa1', 's']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't', '\xcd', '\xa1', 's', '\xc9', '\x9b', 'z', '\xcb', '\x88', 'jas', 'm', '\xc9', '\x9b', 't', '\xcd', '\xa1', 's']]], 19, True)



```

['Tsez', 'Language', 'codes', 'ISO', '639-3', 'ddo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tregami', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tregami', 'Region', 'Nurestan', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tregami', 'Language', 'codes', 'ISO', '639-3', 'trm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Waigali', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Waigali', 'Region', 'Kunar', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Waigali', 'Language', 'codes', 'ISO', '639-3', 'wbk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Waigali', 'Language', 'codes', 'Linguasphere', '58-ACC-a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Kata-vari', 'Total', 'speakers', '18', ',', '700', '(', 'Ethnologue',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '18'], ['NUM',
'700']], 9, True)

['Kata-vari', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kata-vari', 'Region', 'Kunar', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

```

['Kata-vari', 'Language', 'codes', 'ISO', '639-3', 'bsh']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vasi-vari', 'Spoken', 'in', 'Afghanistan']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Vasi-vari', 'Region', 'Prasun', 'Valley']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Vasi-vari', 'Language', 'codes', 'ISO', '639-3', 'prn']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vasi-vari', 'Language', 'codes', 'Linguasphere', '58-ACB-b']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Woi', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Woi', 'Language', 'codes', 'ISO', '639-3', 'wbw']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Matngele', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)

['Matngele', 'Writing', 'system', 'Latin', 'alphabet']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Matngele', 'Language', 'codes', 'ISO', '639-3', 'zml']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warluwara', 'Spoken', 'in', 'Queensland']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Warluwara', 'Writing', 'system', 'Latin', 'alphabet']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Warluwara', 'Language', 'codes', 'ISO', '639-3', 'wrb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Domaaki', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Domaaki', 'Region', 'Nager', 'and', 'Hunza', 'Valleys']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nager', 'and', 'Hunza', 'Valleys']]], 6, True)

['Domaaki', 'Language', 'codes', 'ISO', '639-3', 'dmk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Pashayi', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pashayi', 'Total', 'speakers', '216', ',', '842', '(', 'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']]], ['NUM', '216'], ['NUM', '842']], 9, True)

['Pashayi', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Pashayi', 'Language', 'codes', 'ISO', '639-3', ':', 'aee', '\\xe2', '\\x80', '\\x93', 'Northeastern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'aee', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM', '639-3']], 11, True)

['Pashayi', 'Language', 'codes', 'ISO', '639-3', 'glh', '\\xe2', '\\x80', '\\x93', 'Northwestern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'glh', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)

['Pashayi', 'Language', 'codes', 'ISO', '639-3', 'psi', '\\xe2', '\\x80',

```

'\x93', 'Southeastern']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'psi', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Pashayi', 'Language', 'codes', 'ISO', '639-3', 'psh', '\xe2', '\x80',
'\x93', 'Southwestern']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'psh', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Pashayi', 'Language', 'codes', 'Linguasphere', '59-AAA-a']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Eonavian', 'Region', 'Asturias']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Eonavian', 'Spoken', 'in', 'Spain']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Eonavian', 'Writing', 'system', 'Latin', '(', 'Asturian', 'variant',
')']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Eonavian', 'Pronunciation', '\xc9', '\xa1', 'a', '\xcb', '\x88', '\xca',
'\x8e', 'e', '\xc9', '\xa3', 'o', '.', 'astu', '\xc9', '\xbe', 'i', '\xcb',
'\x88', 'ano', ',', 'e', '.', 'ona', '\xcb', '\x88', '\xce', '\xb2', 'je',
'\xc9', '\xa3', 'o']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',
'\xa1', 'a', '\xcb', '\x88', '\xca', '\x8e', 'e', '\xc9', '\xa3', 'o', '.',
'astu', '\xc9', '\xbe', 'i', '\xcb', '\x88', 'ano', ',', 'e', '.', 'ona',
'\xcb', '\x88', '\xce', '\xb2', 'je', '\xc9', '\xa3', 'o']]], 33, True)

['Eonavian', 'Official', 'language', 'in', 'None']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Eonavian', 'Regulated', 'by', 'Navia', 'Secretary', 'Linguistics']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Navia', 'Secretary', 'Linguistics']]], 6, True)

['Eonavian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']]:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Shumashti', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shumashti', 'Region', 'Kunar', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shumashti', 'Language', 'codes', 'ISO', '639-3', 'sts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Itonama', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Itonama', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Itonama', 'Language', 'codes', 'ISO', '639-3', 'ito']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dan', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dan', 'Spoken', 'in', 'and', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Dan', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Dan', 'Language', 'codes', 'ISO', '639-3', 'daf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ikwerre', 'Dialects', 'Ndele']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ikwerre', 'Dialects', '\xe1', '\xbb', '\x8c', 'gbakiri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xe1',  
'\xbb', '\x8c', 'gbakiri']]], 6, True)

['Ikwerre', 'Dialects', '\xe1', '\xbb', '\x8c', 'b', '\xe1', '\xbb',  
'\x8b', '\xe1', '\xbb', '\x8d']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xe1',  
'\xbb', '\x8c', 'b', '\xe1', '\xbb', '\x8b', '\xe1', '\xbb', '\x8d']]], 12,  
True)

['Ikwerre', 'Dialects', 'Al', '\xe1', '\xbb', '\xa5', '\xe1', '\xbb',  
'\xa5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Al', '\xe1',  
'\xbb', '\xa5', '\xe1', '\xbb', '\xa5']]], 9, True)

['Ikwerre', 'Dialects', '\xe1', '\xbb', '\x8a', 'baa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '\xe1',  
'\xbb', '\x8a', 'baa']]], 6, True)

['Ikwerre', 'Dialects', 'Elele']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ikwerre', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ikwerre', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ikwerre', 'Pronunciation', '\xc3', '\xac', 'k', '\xca', '\xb7', '\xc3',  
'\xa9', 'r', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc3',  
'\xac', 'k', '\xca', '\xb7', '\xc3', '\xa9', 'r', '\xc3', '\xa9']]], 12,  
True)

['Ikwerre', 'Language', 'codes', 'ISO', '639-3', 'ikw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gogo', 'Region', 'and', 'Manyoni', '(', 'Singida', ')', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Manyoni', '\\(', 'Singida', '\\)'], ',', 'Tanzania']]], 9, True)

```
['Gogo', 'Spoken', 'in', 'Dodoma']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Gogo', 'Language', 'codes', 'ISO', '639-2', 'bnt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Gogo', 'Language', 'codes', 'ISO', '639-3', 'gog']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pongu', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pongu', 'Region', 'Niger', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pongu', 'Language', 'codes', 'ISO', '639-3', 'png']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gondi', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Gondi', 'Total', 'speakers', '2', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '2']], 5, True)
```

```
['Gondi', 'Writing', 'system', 'Devanagari', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Gondi', 'Writing', 'system', 'Telugu', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Gondi', 'Writing', 'system', 'Gondi', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Gondi', 'Language', 'codes', 'ISO', '639-2', 'gon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gondi', 'Language', 'codes', 'ISO', '639-3', 'gon', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gon', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Gondi', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Woleaian', 'Total', 'speakers', '1700']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Woleaian', 'Spoken', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

['Woleaian', 'Region', 'Woleai']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Woleaian', 'Official', 'language', 'in', 'Federated', 'States', 'of', 'Micronesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Federated', 'States', 'of', 'Micronesia']], 8, True)

['Woleaian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Woleaian', 'Language', 'codes', 'ISO', '639-3', 'woe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['monastic', 'Signed', 'in', 'Europe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['monastic', 'Total', 'signers', '(', 'not', 'a', 'primary', 'language', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\(', 'not', 'a', 'primary', 'language', '\\)']], 9, True)



```
['monastic', 'Dialects', 'Anglo-Saxon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['monastic', 'Dialects', 'Augustinian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['monastic', 'Dialects', 'Benedictine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['monastic', 'Dialects', 'Cistercian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['monastic', 'Dialects', 'Trappist']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Communicationssprache', 'Total', 'speakers', 'Unknown', ',', 'probably',  
'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Unknown', ',', 'probably', 'none']]], 7, True)
```

```
['Communicationssprache', 'Date', 'founded', '1839']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Communicationssprache', 'Category', '(', 'purpose', ')', 'constructed',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)
```

```
['Communicationssprache', 'Category', '(', 'purpose', ')',  
'Communicationssprache']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]], 6, True)
```

```
['Communicationssprache', 'Created', 'by', 'Joseph', 'Schipfer']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Communicationssprache', 'Category', '(', 'sources', ')', 'Most', 'of',  
'the', 'vocabulary', 'and', 'grammar', 'from', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',
```

'sources', '\\)', 'Most', 'of', 'the', 'vocabulary', 'and', 'grammar',  
'from', 'Type: ANY']]]], 13, True)

['Communicationssprache', 'Category', '(', 'sources', ')', 'with', 'some',  
'influences', 'from', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'with', 'some', 'influences', 'from', 'Type: ANY']]]], 10,  
True)

['Communicationssprache', 'Language', 'codes', 'ISO', '639-2', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Communicationssprache', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Communicationssprache', 'Language', 'codes', 'ISO', '639-1', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Kolami', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Kolami', 'Language', 'codes', 'ISO', '639-2', 'dra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kolami', 'Language', 'codes', 'ISO', '639-3', 'kfb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kituba', 'Spoken', 'in', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Kituba', 'Total', 'speakers', '5-15', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '5-15']], 5, True)

['Kituba', 'Official', 'language', 'in', 'Democratic', 'Republic', 'of',  
'the', 'Congo', ',', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Democratic', 'Republic', 'of', 'the', 'Congo', ',', 'Republic',  
'of', 'the', 'Congo']]]], 14, True)

['Kituba', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Kituba', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mozabite', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mozabite', 'Language', 'codes', 'ISO', '639-3', 'mzb', 'Berber-speaking', 'areas', 'of', 'the', 'Mzab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mzb', 'Berber-speaking', 'areas', 'of', 'the', 'Mzab']], ['NUM', '639-3']], 11, True)

['Mozabite', 'Language', 'codes', 'ISO', '639-3', 'Ouargla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Arb\\xc3\\xabresh', 'Region', 'Apulia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arb\\xc3\\xabresh', 'Region', 'Basilicata']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arb\\xc3\\xabresh', 'Region', 'Calabria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arb\\xc3\\xabresh', 'Region', 'Molise']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arb\\xc3\\xabresh', 'Region', 'Sicily']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Arb\\xc3\\xabresh', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Arb\xc3\xabresh', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Arb\xc3\xabresh', 'Pronunciation', '\xcb', '\x8c', '\xc3', '\xa6',  
'\xc9', '\xbe', 'b', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xbe', 'i',  
'\xca', '\x83', 't']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x8c', '\xc3', '\xa6', '\xc9', '\xbe', 'b', '\xc9', '\x99', '\xcb',  
'\x88', '\xc9', '\xbe', 'i', '\xca', '\x83', 't']]], 19, True)

['Arb\xc3\xabresh', 'Language', 'codes', 'ISO', '639-3', 'aae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ishkashimi', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ishkashimi', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ishkashimi', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ishkashimi', 'Total', 'speakers', 'Not', 'known']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Not', 'known']]], 5, True)

['Ishkashimi', 'Writing', 'system', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ishkashimi', 'Language', 'codes', 'ISO', '639-3', 'sgl', '\xe2', '\x80',  
'\x93', 'Sanglechi', '\xe2', '\x80', '\x93', 'Ishkashimi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sgl', '\xe2', '\x80', '\x93', 'Sanglechi', '\xe2', '\x80',  
'\x93', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Ishkashimi', 'Language', 'codes', 'Linguasphere', '58-ABD-db']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

```
['Kano\xc3\xa9', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kano\xc3\xa9', 'Total', 'speakers', '7']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kano\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'kxo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Emerillon', 'Spoken', 'in', 'French', 'Guiana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Emerillon', 'Total', 'speakers', '400']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Emerillon', 'Language', 'codes', 'ISO', '639-3', 'eme']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arrernte', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)
```

```
['Arrernte', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Arrernte', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arrernte', 'Language', 'codes', 'ISO', '639-3', ':', 'amx', '\xe2',
'\x80', '\x93', 'Anmatjirra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'amx', '\xe2', '\x80', '\x93', 'Anmatjirra']], ['NUM',
'639-3']], 11, True)
```

```
['Arrernte', 'Language', 'codes', 'ISO', '639-3', 'aly', '\xe2', '\x80',
'\x93', 'Alyawarr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'aly', '\xe2', '\x80', '\x93', 'Alyawarr']], ['NUM',
```

'639-3']], 10, True)

['Arrernte', 'Language', 'codes', 'ISO', '639-3', 'adg', '\xe2', '\x80', '\x93', 'Antekerrepenhe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'adg', '\xe2', '\x80', '\x93', 'Antekerrepenhe']], ['NUM', '639-3']], 10, True)

['Upper', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Upper', 'Region', 'Alaska', '(', 'middle', 'Yukon', 'River', ',', 'Koyukuk', 'River', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\(', 'middle', 'Yukon', 'River', ',', 'Koyukuk', 'River', '\)']], 11, True)

['Upper', 'Language', 'codes', 'ISO', '639-3', 'tau']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Borujerdi', 'Spoken', 'in', 'Iran', ',', 'Iraq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Borujerdi', 'Region', 'Central', 'Zagros']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tera', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tera', 'Language', 'codes', 'ISO', '639-3', 'ttr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gunai', 'Spoken', 'in', 'Gippsland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gunai', 'Spoken', 'in', 'Victoria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Gunai', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Gunai', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wolaytta', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Wolaytta', 'Writing', 'system', 'Ethiopic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Wolaytta', 'Language', 'codes', 'ISO', '639-2', 'wal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Wolaytta', 'Language', 'codes', 'ISO', '639-3', 'wal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kamkata-viri', 'Total', 'speakers', '18', ',', '700', '(', 'Ethnologue',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '18'], ['NUM',
'700']], 9, True)
```

```
['Kamkata-viri', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kamkata-viri', 'Region', 'Kunar', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kamkata-viri', 'Dialects', 'Kata-vari']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Kamkata-viri', 'Dialects', 'Kamviri']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```

['Kamkata-viri', 'Dialects', 'Mumviri']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kamkata-viri', 'Language', 'codes', 'ISO', '639-3', 'bsh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamkata-viri', 'Language', 'codes', 'Linguasphere', '58-ACB-a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Bakhtiari', 'Spoken', 'in', 'as', 'native', 'language', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'as', 'native', 'language', 'in', 'Iran']]], 8, True)

['Bakhtiari', 'Writing', 'system', 'Persian', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bakhtiari', 'Language', 'codes', 'ISO', '639-3', 'bqi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chinese', 'Signed', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Signed', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Signed', 'in', 'Taiwan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Total', 'signers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Dialects', 'Southern', 'CSL']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chinese', 'Dialects', 'Northern', 'CSL']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```



'Type: ANY']]], 4, True)

['Nagamese', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nagamese', 'Language', 'codes', 'ISO', '639-3', 'nag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Malak-Malak', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Malak-Malak', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Malak-Malak', 'Language', 'codes', 'ISO', '639-3', 'mpb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tobelo', 'Language', 'codes', 'ISO', '639-3', 'tlb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'Nusa', 'Tenggara', 'Timur', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nusa',  
'Tenggara', 'Timur', ',', 'Indonesia']]], 8, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'lev']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['(Lower)', 'Region', 'Alaska', '(', 'middle', 'Yukon', 'River', ',',  
'Koyukuk', 'River', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\\n(\n', 'middle', 'Yukon', 'River', ',', 'Koyukuk', 'River', '\\\n)']]], 11,  
True)

['(Lower)', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['(Lower)', 'Language', 'codes', 'ISO', '639-3', 'taa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koyukon', 'Total', 'speakers', '300']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Koyukon', 'Region', 'Alaska', '(', 'middle', 'Yukon', 'River', ',', 'Koyukuk', 'River', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\(', 'middle', 'Yukon', 'River', ',', 'Koyukuk', 'River', '\\)']], 11, True)

['Koyukon', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Koyukon', 'Language', 'codes', 'ISO', '639-3', 'koy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ogh-Undjan', 'Total', 'speakers', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ogh-Undjan', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ogh-Undjan', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Ogh-Undjan', 'Language', 'codes', 'ISO', '639-3', 'kjn-ogh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nedebang', 'Total', 'speakers', '200']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nedebang', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Nedebang', 'Region', 'Pantar', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nedebang', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Nedebang', 'Language', 'codes', 'ISO', '639-3', 'nec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ekari', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Ekari', 'Language', 'codes', 'ISO', '639-3', 'ekg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Petjo', 'Spoken', 'in', 'Indonesia', ',', 'the', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Indonesia', ',', 'the', 'Netherlands']]], 7, True)
```

```
['Petjo', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Petjo', 'Language', 'codes', 'ISO', '639-3', 'pey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yol\xc5\x8bu', 'Spoken', 'in', 'Northern', 'Territory', ',',  
'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)
```

```
['Yol\xc5\x8bu', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Yol\xc5\x8bu', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'variously']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', ':', 'dji', '\xe2', '\x80', '\x93', 'Djinang']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'dji', '\xe2', '\x80', '\x93', 'Djinang']], ['NUM', '639-3']], 11, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'dhg', '\xe2', '\x80', '\x93', 'Dhangu']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'dhg', '\xe2', '\x80', '\x93', 'Dhangu']], ['NUM', '639-3']], 10, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'duj', '\xe2', '\x80', '\x93', 'Dhuwal']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'duj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'djb', '\xe2', '\x80', '\x93', 'Djinba']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'djb', '\xe2', '\x80', '\x93', 'Djinba']], ['NUM', '639-3']], 10, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'rit', '\xe2', '\x80', '\x93', 'Ritharngu']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'rit', '\xe2', '\x80', '\x93', 'Ritharngu']], ['NUM', '639-3']], 10, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'jay', '\xe2', '\x80', '\x93', 'Yan-nhangu']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'jay', '\xe2', '\x80', '\x93', 'Yan-nhangu']], ['NUM', '639-3']], 10, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'dax', '\xe2', '\x80', '\x93', 'Dayi']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'dax', '\xe2', '\x80', '\x93', 'Dayi']], ['NUM', '639-3']], 10, True)
```

```
['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'djr', '\xe2',
```

```

'\x80', '\x93', 'Djambarrpuyngu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'djr', '\xe2', '\x80', '\x93', 'Djambarrpuyngu']], ['NUM',
'639-3']], 10, True)

['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'gnn', '\xe2',
'\x80', '\x93', 'Gumatj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'gnn', '\xe2', '\x80', '\x93', 'Gumatj']], ['NUM', '639-3']],
10, True)

['Yol\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'guf', '\xe2',
'\x80', '\x93', 'Gupapuyngu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'guf', '\xe2', '\x80', '\x93', 'Gupapuyngu']], ['NUM',
'639-3']], 10, True)

['Mumviri', 'Total', 'speakers', '1', ',', '500', '(', 'Richard', 'F', '.',
'Strand', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Richard', 'F', '.', 'Strand', '\\)']], ['NUM',
'1'], ['NUM', '500']], 12, True)

['Mumviri', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mumviri', 'Region', 'basin', 'of', 'eastern', 'Nuristan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'basin', 'of',
'eastern', 'Nuristan']]], 6, True)

['Mumviri', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Mumviri', 'Language', 'codes', 'ISO', '639-3', 'bsh-mum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mumviri', 'Language', 'codes', 'Linguasphere', '58-ACB-ac']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Takalak', 'Language', 'extinction', '?']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Takalak', 'Spoken', 'in', 'Queensland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Takalak', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Takalak', 'Language', 'codes', 'ISO', '639-3', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Betta', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Betta', 'Language', 'codes', 'ISO', '639-3', 'xub']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Samaritan', 'Writing', 'system', 'Samaritan', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Samaritan', 'Pronunciation', '\xc9', '\x91', 'r', '\xc9', '\x91', 'm', '\xc9', '\x91', 'j', '\xc9', '\x91', ',', '\xc9', '\x94', 'r', '\xc9', '\x94', 'm', '\xc9', '\x94', 'j', '\xc9', '\x94']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9', '\x91', 'r', '\xc9', '\x91', 'm', '\xc9', '\x91', 'j', '\xc9', '\x91', ',', '\xc9', '\x94', 'r', '\xc9', '\x94', 'm', '\xc9', '\x94', 'j', '\xc9', '\x94']]], 25, True)

['Samaritan', 'Pronunciation', 'ar', '\xc9', '\x91', 'mi', '\xce', '\xb8']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ar', '\xc9', '\x91', 'mi', '\xce', '\xb8']]], 8, True)

['Samaritan', 'Pronunciation', 'ar', '\xc9', '\x91', 'mit']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ar', '\xc9', '\x91', 'mit']]], 6, True)

['Samaritan', 'Language', 'codes', 'ISO', '639-2', 'sam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```

['Samaritan', 'Language', 'codes', 'ISO', '639-3', 'sam']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gwere', 'Spoken', 'in', 'Uganda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gwere', 'Language', 'codes', 'ISO', '639-3', 'gwr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nese', 'Spoken', 'in', 'Matanvat', 'area']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Nese', 'Spoken', 'in', 'northwest', 'Malakula']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Nese', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nese', 'Total', 'speakers', '20']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nese', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hlai', 'Region', 'Hainan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Hlai', 'Language', 'codes', 'ISO', '639-3', 'lic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huli', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Huli', 'Language', 'codes', 'ISO', '639-3', 'hui']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kom', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kom', 'Region', 'North-West', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kom', 'Language', 'codes', 'ISO', '639-2', 'bnt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kom', 'Language', 'codes', 'ISO', '639-3', 'bkm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sansiboli', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sansiboli', 'Region', 'Rajasthan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Sansiboli', 'Language', 'codes', 'ISO', '639-3', 'ssi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uruava', 'Language', 'extinction', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uruava', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Uruava', 'Language', 'codes', 'ISO', '639-3', 'urv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Usku', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)



```
['Usku', 'Total', 'speakers', '20']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Usku', 'Language', 'codes', 'ISO', '639-3', 'ulf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Melpa', 'Spoken', 'in', 'Mount', 'Hagen', 'District']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mount',
'Hagen', 'District']]], 6, True)
```

```
['Melpa', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Melpa', 'Language', 'codes', 'ISO', '639-3', 'med']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Arop-Lokep', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Arop-Lokep', 'Language', 'codes', 'ISO', '639-3', 'apr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Seereer-Siin', 'Spoken', 'in', 'Senegal', 'and', 'the', 'Gambia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Senegal', 'and', 'the', 'Gambia']]], 7, True)
```

```
['Seereer-Siin', 'Language', 'codes', 'ISO', '639-2', 'niv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Seereer-Siin', 'Language', 'codes', 'ISO', '639-3', 'srr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tamambo', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tamambo', 'Region', 'Malo', 'Island', ',', 'Espiritu', 'Santo']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Malo', 'Island', ',', 'Espiritu', 'Santo']]], 7, True)

['Tamambo', 'Language', 'codes', 'ISO', '639-3', 'mla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sissano', 'Spoken', 'in', 'Aitape', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sissano', 'Spoken', 'in', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sissano', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sissano', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sissano', 'Language', 'codes', 'ISO', '639-3', 'sso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amaimon', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Amaimon', 'Language', 'codes', 'ISO', '639-3', 'ali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\xe2\x80\x99Auhelawa', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['\xe2\x80\x99Auhelawa', 'Language', 'codes', 'ISO', '639-3', 'kud']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adzera', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

```

['Adzera', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adzera', 'Language', 'codes', 'ISO', '639-3', ':', 'adz', '\xe2', '\x80',
'\x93', 'Adzera']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'adz', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 11, True)

['Adzera', 'Language', 'codes', 'ISO', '639-3', 'zsu', '\xe2', '\x80',
'\x93', 'Sukurum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zsu', '\xe2', '\x80', '\x93', 'Sukurum']], ['NUM',
'639-3']], 10, True)

['Adzera', 'Language', 'codes', 'ISO', '639-3', 'zsa', '\xe2', '\x80',
'\x93', 'Sarasira']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zsa', '\xe2', '\x80', '\x93', 'Sarasira']], ['NUM',
'639-3']], 10, True)

['Kanembu', 'Spoken', 'in', 'Chad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kanembu', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Kanembu', 'Language', 'codes', 'ISO', '639-3', 'kbl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Elseng', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Elseng', 'Total', 'speakers', '300']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Elseng', 'Language', 'codes', 'ISO', '639-3', 'mrf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Waris', 'Spoken', 'in', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Waris', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Waris', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Waris', 'Language', 'codes', 'ISO', '639-3', 'wrs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Amanab', 'Spoken', 'in', 'Amanab', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Amanab', 'Spoken', 'in', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Amanab', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Amanab', 'Language', 'codes', 'ISO', '639-3', 'amn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Haryanvi', 'Total', 'speakers', 'ca', '.', '30', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'c',  
'.', 'NUM', 'million']]], ['NUM', '30']], 7, True)

['Haryanvi', 'Region', 'Haryana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Haryanvi', 'Region', 'Northern', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Haryanvi', 'Region', 'Pakistan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Haryanvi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Haryanvi', 'Writing', 'system', 'Devanagari', 'script', ',', 'Nagari', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Haryanvi', 'Language', 'codes', 'ISO', '639-3', 'bgc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kannauji', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kannauji', 'Total', 'speakers', '6', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '6']], 5, True)

['Kannauji', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kannauji', 'Language', 'codes', 'ISO', '639-3', 'bjj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Afshar', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afshar', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afshar', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Afshar', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Afshar', 'Region', 'Anatolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Afshar', 'Region', 'Mesopotamia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Afshar', 'Region', 'Khorasan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Afshar', 'Region', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Afshar', 'Region', 'Kerman', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Afshar', 'Region', 'Kabul', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Afshar', 'Writing', 'system', 'Perso-Arabic', 'script', ',', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Khorasani', 'Turkic', 'Total', 'speakers', '400', ',', '000', '(', 'estimate', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Turkic', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'estimate', '\\\\)']]], ['NUM', '400'], ['NUM', '000']]], 10, True)

['Khorasani', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Khorasani', 'Region', 'Khorasan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Khorasani', 'Language', 'codes', 'ISO', '639-3', 'kmz']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cupe\\xc3\\xb1o', 'Language', 'extinction', '1987']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cupe\\xc3\\xb1o', 'Writing', 'system', 'Latin', 'alphabet']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Cupe\\xc3\\xb1o', 'Language', 'codes', 'ISO', '639-3', 'cup']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baeggu', 'Total', 'speakers', '6000']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Baeggu', 'Spoken', 'in', 'Solomon', 'Islands']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Baeggu', 'Region', 'Malaita', 'Island']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Baeggu', 'Language', 'codes', 'ISO', '639-3', 'bvd']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hrangkhoh', 'Spoken', 'in', 'India', 'and', 'Myanmar']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Hrangkhoh', 'Region', 'Southeast', 'Manipur', ',', 'parts', 'of', 'Assam']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Southeast', 'Manipur', ',', 'parts', 'of', 'Assam']], 8, True)

['Hrangkhoh', 'Language', 'codes', 'ISO', '639-3', 'hra']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['\\xc3\\x84iwoo', 'Spoken', 'in', 'Solomon', 'Islands']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['\xc3\x84iwoo', 'Region', 'Santa', 'Cruz', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['\xc3\x84iwoo', 'Region', 'eastern', 'Solomons']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['\xc3\x84iwoo', 'Region', 'Temotu', 'Province', '.', 'Coordinates']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Temotu',  
'Province', '.', 'Coordinates']]], 6, True)

['\xc3\x84iwoo', 'Language', 'codes', 'ISO', '639-3', 'nfl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luwian', 'Spoken', 'in', 'Anatolia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Luwian', 'Language', 'codes', 'ISO', '639-3', 'hlu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fiji', 'Spoken', 'in', 'Fiji']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fiji', 'Spoken', 'in', 'with', 'significant', 'minorities', 'in',  
'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'with',  
'significant', 'minorities', 'in', 'Australia']]], 8, True)

['Fiji', 'Spoken', 'in', 'New', 'Zealand']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Fiji', 'Spoken', 'in', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fiji', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



```

['Fiji', 'Language', 'codes', 'ISO', '639-3', 'hif']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pular', 'Total', 'speakers', '10', '\xe2', '\x80', '\x93', '16',
'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '10'], ['NUM',
'16']], 9, True)

['Pular', 'Region', 'Fouta', 'Jalon', '(', 'in', 'Guinea', ')', ',', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Fouta',
'Jalon', '\\(', 'in', 'Guinea', '\\)'], ',', 'Africa']]], 10, True)

['Pular', 'Spoken', 'in', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pular', 'Spoken', 'in', 'Guinea', 'Bissau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Pular', 'Spoken', 'in', 'Sierra', 'Leone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Pular', 'Spoken', 'in', 'Mali']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Pular', 'Writing', 'system', 'Fula', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Pular', 'Language', 'codes', 'ISO', '639-2', 'ful']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pular', 'Language', 'codes', 'ISO', '639-3', 'fuf', '\xe2', '\x80',
'\x93', ']']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'fuf', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

```

['Pular', 'Language', 'codes', 'ISO', '639-1', 'ff']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Italian', 'Signed', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Italian', 'Total', 'signers', 'tens', 'of', 'thousands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',  
'tens', 'of', 'thousands']], 6, True)

['Broome', 'Total', 'speakers', '40', '(', 'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '40']], 7, True)

['Broome', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Broome', 'Language', 'codes', 'ISO', '639-3', 'bpl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Otomi', 'Spoken', 'in', 'Mexico', ':', 'M', '\\xc3', '\\xa9', 'xico', '(',  
'state', ')', ',', 'Puebla', ',', 'Veracruz', ',', 'Hidalgo', ',',  
'Guanajuato', ',', 'Quer', '\\xc3', '\\xa9', 'taro', ',', 'Tlaxcala', ',',  
'Michoac', '\\xc3', '\\xa1', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  
'\\xa9', 'taro', ',', 'Tlaxcala', ',', 'Michoac', '\\xc3', '\\xa1', 'n']]],  
32, True)

['Otomi', 'Regulated', 'by', 'Instituto', 'Nacional', 'de', 'Lenguas',  
'Ind', '\\xc3', '\\xad', 'genas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Instituto', 'Nacional', 'de', 'Lenguas', 'Ind', '\\xc3', '\\xad',  
'genas']]], 11, True)

['Otomi', 'Language', 'codes', 'ISO', '639-2', 'oto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Otomi', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Otomí', 'Language', 'codes', 'ISO', '639-3', 'otl', '\xe2', '\x80', '\x93', 'Tilapa', 'Otomí']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'otl', '\xe2', '\x80', '\x93', 'Tilapa', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Otomí', 'Language', 'codes', 'ISO', '639-3', 'otn', '\xe2', '\x80', '\x93', 'Tenango', 'Otomí']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'otn', '\xe2', '\x80', '\x93', 'Tenango', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Otomí', 'Language', 'codes', 'ISO', '639-3', 'otq', '\xe2', '\x80', '\x93', 'Quer', '\xc3', '\xa9', 'taro', 'Otomí']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'otq', '\xe2', '\x80', '\x93', 'Quer', '\xc3', '\xa9', 'taro', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Otomí', 'Language', 'codes', 'ISO', '639-3', 'ots', '\xe2', '\x80', '\x93', 'Estado', 'de', 'M', '\xc3', '\xa9', 'xico', 'Otomí']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ots', '\xe2', '\x80', '\x93', 'Estado', 'de', 'M', '\xc3', '\xa9', 'xico', 'Type: ANY']], ['NUM', '639-3']], 16, True)

['Otomí', 'Language', 'codes', 'ISO', '639-3', 'otx', '\xe2', '\x80', '\x93', 'Texcatepec', 'Otomí']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'otx', '\xe2', '\x80', '\x93', 'Texcatepec', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Otomí', 'Language', 'codes', 'ISO', '639-3', 'otz', '\xe2', '\x80', '\x93', 'Ixtenco', 'Otomí']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'otz', '\xe2', '\x80', '\x93', 'Ixtenco', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Cypriot', 'Region', 'Cyprus', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cypriot', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

```
['Cypriot', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Cypriot', 'Language', 'codes', 'ISO', '639-2', 'tur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Cypriot', 'Language', 'codes', 'ISO', '639-3', 'tur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Cypriot', 'Language', 'codes', 'ISO', '639-1', 'tr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Japanese', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Japanese', 'Total', 'speakers', '\xe2', '\x80', '\x94']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\x94']]], 6, True)
```

```
['Japanese', 'Language', 'codes', 'ISO', '639-2', 'cpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Japanese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)
```

```
['Nyulnyul', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Nyulnyul', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nyulnyul', 'Language', 'codes', 'ISO', '639-3', 'nyv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Doolboong', 'Language', 'extinction', 'Extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Doolboong', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Doolboong', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Doolboong', 'Language', 'codes', 'ISO', '639-3', '08j']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Najdi', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Najdi', 'Spoken', 'in', 'Jordan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Najdi', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Najdi', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Najdi', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Najdi', 'Language', 'codes', 'ISO', '639-3', 'ars']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hejazi', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Hejazi', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hejazi', 'Language', 'codes', 'ISO', '639-3', 'acw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Onhan', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Onhan', 'Region', 'Western', 'Visayas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Onhan', 'Language', 'codes', 'ISO', '639-3', 'loc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Welsh', 'Spoken', 'in', 'Wales', '(', 'United', 'Kingdom', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Wales',  
'\\(', 'United', 'Kingdom', '\\)']], 8, True)

['Welsh', 'Total', 'speakers', 'Probably', 'extinct', 'as', 'a', 'first',  
'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Probably', 'extinct', 'as', 'a', 'first', 'language']], 9, True)

['Welsh', 'Language', 'codes', 'ISO', '639-2', 'rom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Welsh', 'Language', 'codes', 'ISO', '639-3', 'rmw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Berti', 'Spoken', 'in', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Berti', 'Region', 'Darfur', 'and', 'Kordofan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Berti', 'Language', 'codes', 'ISO', '639-3', 'byn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Bidayat', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Bidayat', 'Total', 'speakers', '3000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Bidayat', 'Language', 'codes', 'ISO', '639-3', 'zag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cora', 'Spoken', 'in', 'Mexico', ':', 'Nayarit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  

```

```
['Cora', 'Spoken', 'in', 'Jalisco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Cora', 'Spoken', 'in', 'Durango']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Cora', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)
```

```
['Cora', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de',  
'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P',  
'\xc3', '\xba', 'blica']], 16, True)
```

```
['Cora', 'Language', 'codes', 'ISO', '639-3', 'crn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Central', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Central', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```

['Central', 'Language', 'codes', 'ISO', '639-3', 'czo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Central', 'Language', 'codes', 'ISO', '639-1', 'zh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Berom', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Berom', 'Region', 'Plateau', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Berom', 'Language', 'codes', 'ISO', '639-3', 'bom']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Airoran', 'Spoken', 'in', 'Papua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Airoran', 'Total', 'speakers', '1000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Airoran', 'Language', 'codes', 'ISO', '639-3', 'air']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nete', 'Language', 'codes', 'ISO', '639-3', 'net']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oring', 'Total', 'speakers', '75000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Oring', 'Spoken', 'in', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2',
'\x80', '\x93']], 6, True)

['Oring', 'Region', '\xe2', '\x80', '\x93']:

```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Oring', 'Language', 'codes', 'ISO', '639-3', 'org']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baloi', 'Spoken', 'in', 'Equateur', 'Province', 'of', 'the', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Equateur', 'Province', 'of', 'the', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 12, True)

['Baloi', 'Total', 'speakers', '20000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Baloi', 'Language', 'codes', 'ISO', '639-3', 'biz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cita', 'Spoken', 'in', 'Mappi', 'Regency']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Cita', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cita', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cita', 'Language', 'codes', 'ISO', '639-3', 'txt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marba', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Marba', 'Language', 'codes', 'ISO', '639-3', 'mpg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Logo', 'Spoken', 'in', 'Orientale', 'Province', ',', 'Democratic',

'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Orientale', 'Province', ',', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 11, True)

['Logo', 'Language', 'codes', 'ISO', '639-3', 'log']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Naxi', 'Region', 'Yunnan', 'and', 'Tibet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Naxi', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Naxi', 'Writing', 'system', 'Dongba', 'script', 'and', 'Geba', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Naxi', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Naxi', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Naxi', 'Language', 'codes', 'ISO', '639-3', 'nbf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Avikam', 'Total', 'speakers', '21000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Avikam', 'Language', 'codes', 'ISO', '639-3', 'avi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shughni', 'Writing', 'system', 'Arabic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shughni', 'Writing', 'system', 'Cyrillic']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shughni', 'Writing', 'system', 'Latin']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shughni', 'Language', 'codes', 'ISO', '639-3', 'sgn']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vanji', 'Spoken', 'in', 'Bukharan', 'Emirate']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Vanji', 'Writing', 'system', 'None']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vanji', 'Language', 'codes', 'ISO', '639-3', 'None']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mnong', 'Spoken', 'in', 'Vietnam', ',', 'Cambodia', 'and', 'USA']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Vietnam', ',', 'Cambodia', 'and', 'USA']]], 8, True)

['Mnong', 'Region', 'throughout', 'T', '\xc3', '\xa2', 'y', 'Nguy', '\xc3', '\xaa', 'n', 'region']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'throughout', 'T', '\xc3', '\xa2', 'y', 'Nguy', '\xc3', '\xaa', 'n', 'region']]], 12, True)

['Mnong', 'Region', 'especially', 'in', '\xc4', '\x90', '\xe1', '\xba', '\xaf', 'k', 'L', '\xe1', '\xba', '\xaf', 'k']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'especially', 'in', '\xc4', '\x90', '\xe1', '\xba', '\xaf', 'k', 'L', '\xe1', '\xba', '\xaf', 'k']]], 15, True)

['Mnong', 'Region', 'L', '\xc3', '\xa2', 'm', '\xc4', '\x90', '\xe1', '\xbb', '\x93', 'ng']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'L', '\xc3', '\xa2', 'm', '\xc4', '\x90', '\xe1', '\xbb', '\x93', 'ng']]], 12, True)

['Mnong', 'Language', 'codes', 'ISO', '639-3', 'variously']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tindi', 'Spoken', 'in', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Tindi', 'Language', 'codes', 'ISO', '639-3', 'tin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['African', 'Spoken', 'in', 'North', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['African', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['African', 'Language', 'codes', 'ISO', '639-3', 'lat-afr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Coast', 'Spoken', 'in', 'Canada', ',', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'S']]], 7, True)

['Coast', 'Language', 'codes', 'ISO', '639-2', 'tsi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Coast', 'Language', 'codes', 'ISO', '639-3', 'tsi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sarikoli', 'Writing', 'system', 'None', ',', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Sarikoli', 'Language', 'codes', 'ISO', '639-3', 'srh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sarikoli', 'Language', 'codes', 'Linguasphere', '58-ABD-eb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'Linguasphere', 'Type: ANY']]], 5, True)

['Amis', 'Spoken', 'in', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Amis', 'Language', 'codes', 'ISO', '639-3', 'ami']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsou', 'Spoken', 'in', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tsou', 'Region', 'Alishan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tsou', 'Language', 'codes', 'ISO', '639-3', 'tsu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Puyuma', 'Spoken', 'in', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Puyuma', 'Language', 'codes', 'ISO', '639-3', 'pyu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rukai', 'Spoken', 'in', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rukai', 'Language', 'codes', 'ISO', '639-3', 'dru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Thao', 'Spoken', 'in', 'Taiwan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Thao', 'Language', 'codes', 'ISO', '639-3', 'ssf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pazeh', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pazeh', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pazeh', 'Language', 'codes', 'ISO', '639-3', 'uun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kanakanabu', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kanakanabu', 'Language', 'codes', 'ISO', '639-3', 'xnb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saaroa', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Saaroa', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Saaroa', 'Region', 'west', 'central', 'Mountains', 'of', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'west', 'central', 'Mountains', 'of', 'Taiwan']]], 7, True)

['Saaroa', 'Region', 'south', 'and', 'southeast', 'of', 'Minchuan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'south', 'and', 'southeast', 'of', 'Minchuan']]], 7, True)

['Saaroa', 'Region', 'along', 'the', 'Laonung', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'along', 'the', 'Laonung', 'River']]], 6, True)

['Saaroa', 'Language', 'codes', 'ISO', '639-3', 'sxr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laurentian', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Laurentian', 'Language', 'codes', 'ISO', '639-3', 'lre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rejang', 'Spoken', 'in', 'Southwestern', 'Sumatra', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Rejang', 'Total', 'speakers', 'Up', 'to', '1', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Up', 'to', 'NUM', 'million']], ['NUM', '1']], 7, True)

['Rejang', 'Language', 'codes', 'ISO', '639-3', 'rej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zhang-Zhung', 'Language', 'codes', 'ISO', '639-3', 'xzh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vestinian', 'Spoken', 'in', 'East-central', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Vestinian', 'Language', 'codes', 'ISO', '639-3', 'xvs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sebat', 'Bet', 'Gurage', 'Total', 'speakers', '2', ',', '320', ',', '000',  
'(', 'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Bet', 'Gurage', 'Total',  
'speakers', 'NUM', ',', 'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']],  
['NUM', '2'], ['NUM', '320'], ['NUM', '000']], 13, True)

['Sebat', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sebat', 'Region', 'West', 'Gurage', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sebat', 'Writing', 'system', 'Ethiopic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Sebat', 'Language', 'codes', 'ISO', '639-3', 'sgw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northeastern', 'Mandarin', 'Total', 'speakers', '82', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', 'NUM', 'million']], ['NUM', '82']], 6, True)

['Northeastern', 'Region', 'Manchuria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Northeastern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Northeastern', 'Language', 'codes', 'ISO', '639-3', 'cmn-nem']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bzyb', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Bzyb', 'Spoken', 'in', 'Turkey', ',', 'Abkhazia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Bzyb', 'Region', 'Bzyb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bzyb', 'Language', 'codes', 'ISO', '639-3', 'abk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Quechua', 'Total', 'speakers', '5', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', 'NUM', 'million']], ['NUM', '5']], 6, True)

['Southern', 'Region', 'Andes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:



ANY']]], 3, True)

['Southern', 'Spoken', 'in', 'Peru', ',', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Southern', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Southern', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Southern', 'Language', 'codes', 'ISO', '639-1', 'qu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Cusco', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cusco', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cusco', 'Language', 'codes', 'ISO', '639-3', 'quz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'odt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Ji-Lu', 'Spoken', 'in', 'Hebei', ',', 'Shandong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Ji-Lu', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Ji-Lu', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Ji-Lu', 'Language', 'codes', 'ISO', '639-3', 'cmn-jil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ancient', 'Spoken', 'in', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ancient', 'Writing', 'system', 'South', 'Arabian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Ancient', 'Language', 'codes', 'ISO', '639-3', 'xna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngong', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ngong', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngong', 'Language', 'codes', 'ISO', '639-3', ':', 'nnx', '\xe2', '\x80',  
'\x93', 'Ngong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'nnx', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Ngong', 'Language', 'codes', 'ISO', '639-3', 'ngv', '\xe2', '\x80',  
'\x93', 'Nagumi']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ng', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Central', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Central', 'Region', 'Bicol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Central', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']], 7, True)

['Central', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Central', 'Language', 'codes', 'ISO', '639-2', 'bik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'bcl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Austrian', 'Signed', 'in', 'Austria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Austrian', 'Total', 'signers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\xe2', '\x80', '\x93']], 6, True)

['Namwanga', 'Spoken', 'in', 'Zambia', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Namwanga', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

```
['Namwanga', 'Language', 'codes', 'ISO', '639-3', 'mwn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngulu', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ngulu', 'Region', 'East-central', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ngulu', 'Language', 'codes', 'ISO', '639-3', 'ngp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wambaya', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Wambaya', 'Language', 'codes', 'ISO', '639-3', 'wmb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Capiznon', 'Spoken', 'in', 'Capiz', ',', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Capiznon', 'Region', 'Visayas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Capiznon', 'Language', 'codes', 'ISO', '639-3', 'cps']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kazukuru', 'Total', 'speakers', 'extinct']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kazukuru', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Kazukuru', 'Region', 'New', 'Georgia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Kazukuru', 'Language', 'codes', 'ISO', '639-3', 'kzk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dororo', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dororo', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Dororo', 'Region', 'New', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dororo', 'Language', 'codes', 'ISO', '639-3', 'drr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guliguli', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Guliguli', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Guliguli', 'Region', 'New', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Guliguli', 'Language', 'codes', 'ISO', '639-3', 'gli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Qingdao', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Qingdao', 'Region', 'Qingdao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Qingdao', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Archi', 'Spoken', 'in', 'Archib']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Archi', 'Spoken', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Archi', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Archi', 'Language', 'codes', 'ISO', '639-3', 'aqc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nai', 'Language', 'codes', 'ISO', '639-3', 'bio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ona', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Ona', 'Language', 'codes', 'ISO', '639-3', 'ona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kunigami', 'Total', 'speakers', '5000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kunigami', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kunigami', 'Region', 'Northern', 'Okinawa', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Kunigami', 'Language', 'codes', 'ISO', '639-3', 'xug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wanka', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]]], 6, True)

['Wanka', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Wanka', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wanka', 'Language', 'codes', 'ISO', '639-3', ':', 'qvw', '\xe2', '\x80',  
'\x93', 'Waylla', 'Wanka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'qvw', '\xe2', '\x80', '\x93', 'Waylla', 'Type: ANY']],  
['NUM', '639-3']], 12, True)

['Wanka', 'Language', 'codes', 'ISO', '639-3', 'qwx', '\xe2', '\x80',  
'\x93', 'Shawsha', 'Wanka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qwx', '\xe2', '\x80', '\x93', 'Shawsha', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Gooniyandi', 'Spoken', 'in', 'Western', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Gooniyandi', 'Total', 'speakers', '100']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Gooniyandi', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Gooniyandi', 'Language', 'codes', 'ISO', '639-3', 'gni']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Myene', 'Total', 'speakers', '46', ',', '743', '(', 'Ethnologue', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']]], ['NUM', '46'], ['NUM',  
'743']], 9, True)

['Myene', 'Spoken', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Myene', 'Language', 'codes', 'ISO', '639-3', 'mye']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Belhare', 'Language', 'codes', 'ISO', '639-3', 'byw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["P'urh\xc3\xa9pecha,", 'Pronunciation', 'p', '\xca', '\xb0', 'u', '\xcb', '\x88', '\xc9', '\xbd', 'epet', '\xcd', '\xa1', '\xca', '\x83', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'p', '\xca', '\xb0', 'u', '\xcb', '\x88', '\xc9', '\xbd', 'epet', '\xcd', '\xa1', '\xca', '\x83', 'a']], 16, True)

["P'urh\xc3\xa9pecha,", 'Spoken', 'in', 'Michoac', '\xc3', '\xa1', 'n', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Michoac', '\xc3', '\xa1', 'n', ',', 'Mexico']], 9, True)

["P'urh\xc3\xa9pecha,", 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["P'urh\xc3\xa9pecha,", 'Language', 'codes', 'ISO', '639-3', ':', 'tsz', '\xe2', '\x80', '\x93', 'Eastern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'tsz', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

["P'urh\xc3\xa9pecha,", 'Language', 'codes', 'ISO', '639-3', 'pua', '\xe2', '\x80', '\x93', 'Western']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pua', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Tetela', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Tetela', 'Language', 'codes', 'ISO', '639-3', 'tll']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ancash', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Ancash', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', ':', 'qwa', '\xe2', '\x80',  
'\x93', 'Corongo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'qwa', '\xe2', '\x80', '\x93', 'Corongo']], ['NUM',  
'639-3']], 11, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', 'qwh', '\xe2', '\x80',  
'\x93', 'Huaylas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qwh', '\xe2', '\x80', '\x93', 'Huaylas']], ['NUM',  
'639-3']], 10, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', 'qxn', '\xe2', '\x80',  
'\x93', 'Norte', 'de', 'Conchucos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qxn', '\xe2', '\x80', '\x93', 'Norte', 'de', 'Conchucos']],  
['NUM', '639-3']], 12, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', 'qws', '\xe2', '\x80',  
'\x93', 'Sihuas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qws', '\xe2', '\x80', '\x93', 'Sihuas']], ['NUM', '639-3']],  
10, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', 'qxo', '\xe2', '\x80',  
'\x93', 'Sur', 'de', 'Conchucos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qxo', '\xe2', '\x80', '\x93', 'Sur', 'de', 'Conchucos']],  
['NUM', '639-3']], 12, True)

['Ancash', 'Language', 'codes', 'ISO', '639-3', 'qvn', '\xe2', '\x80',  
'\x93', 'Huamal', '\xc3', '\xad', 'es', 'y', 'Norte', 'de', 'Dos', 'de',

```

'Mayo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'qvn', '\xe2', '\x80', '\x93', 'Huamal', '\xc3', '\xad',
'es', 'y', 'Norte', 'de', 'Dos', 'de', 'Type: ANY']], ['NUM', '639-3']],
19, True)

['Bura-Pabir', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bura-Pabir', 'Region', 'Borno', 'State', ',', 'Adamawa', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Borno',
'State', ',', 'Adamawa', 'State']]], 7, True)

['Bura-Pabir', 'Language', 'codes', 'ISO', '639-3', 'bwr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nara', 'Spoken', 'in', 'Eritrea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nara', 'Language', 'codes', 'ISO', '639-3', 'nrb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kunama', 'Spoken', 'in', 'Eritrea', ',', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Kunama', 'Region', 'western', 'Eritrea', ',', 'northern', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western',
'Eritrea', ',', 'northern', 'Ethiopia']]], 7, True)

['Kunama', 'Writing', 'system', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kunama', 'Language', 'codes', 'ISO', '639-3', 'kun']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bade', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

['Bade', 'Region', 'Yobe', 'State', ',', 'Jigawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Yobe',  
'State', ',', 'Jigawa', 'State']]], 7, True)

['Bade', 'Language', 'codes', 'ISO', '639-3', 'bde']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saraveca', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saraveca', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saraveca', 'Region', 'Eastern', 'jungle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saraveca', 'Language', 'codes', 'ISO', '639-3', 'sar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pumpokol', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Pumpokol', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Pumpokol', 'Language', 'codes', 'ISO', '639-3', 'xpm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyanga-li', 'Spoken', 'in', 'Orientale', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Nyanga-li', 'Language', 'codes', 'ISO', '639-3', 'nyc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yami', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Yami', 'Language', 'codes', 'ISO', '639-3', 'tao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Western', 'Region', 'Primarily', 'south-east', 'Arizona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'apa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'apw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['War', 'Spoken', 'in', 'Bangladesh', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['War', 'Language', 'codes', 'ISO', '639-3', 'aml']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kiput', 'Spoken', 'in', 'Northern', 'Sarawak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Kiput', 'Spoken', 'in', 'Borneo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kiput', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kiput', 'Total', 'speakers', '450']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```

['Kiput', 'Language', 'codes', 'ISO', '639-3', 'kyi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vitu', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Vitu', 'Language', 'codes', 'ISO', '639-3', 'wiv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bube', 'Spoken', 'in', 'Equatorial', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Bube', 'Spoken', 'in', 'Gabon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bube', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bube', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Bube', 'Language', 'codes', 'ISO', '639-3', 'bvb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Holikachuk', 'Total', 'speakers', '12']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Holikachuk', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Holikachuk', 'Region', 'Alaska', '(', 'lower', 'Yukon', 'River', ',',
'Innoko', 'River', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alaska', '\\
(', 'lower', 'Yukon', 'River', ',', 'Innoko', 'River', '\\)']]], 11, True)

['Holikachuk', 'Language', 'codes', 'ISO', '639-3', 'hoi']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Waimoa', 'Language', 'codes', 'ISO', '639-3', 'wmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mankanya', 'Spoken', 'in', 'Senegal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mankanya', 'Spoken', 'in', 'Guinea-Bissau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mankanya', 'Spoken', 'in', 'and', 'the', 'Gambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'the', 'Gambia']], 6, True)

['Mankanya', 'Region', 'Southwest', 'Senegal', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Mankanya', 'Language', 'codes', 'ISO', '639-3', 'knf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gajirrabeng', 'Total', 'speakers', '3', 'or', '4', 'fluent']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'or', 'NUM', 'fluent']], ['NUM', '3'], ['NUM', '4']], 7, True)

['Gajirrabeng', 'Language', 'codes', 'ISO', '639-3', 'gdh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Miriwoong', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Miriwoong', 'Language', 'codes', 'ISO', '639-3', 'mep']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jukun', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Jukun', 'Total', 'speakers', 'No', 'fluent', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'No', 'fluent', 'speakers']]], 6, True)

['Jukun', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Jukun', 'Language', 'codes', 'ISO', '639-3', 'dyd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bai', 'Spoken', 'in', 'Yunnan', ',', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Bai', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lungga', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Lungga', 'Language', 'codes', 'ISO', '639-3', 'lga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kija', 'Total', 'speakers', '100']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kija', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Kija', 'Language', 'codes', 'ISO', '639-3', 'gia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ili', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ili', 'Region', 'Xinjiang']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Ili', 'Language', 'codes', 'ISO', '639-3', 'ili']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tofa', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tofa', 'Region', 'Irkutsk', 'Oblast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Tofa', 'Language', 'codes', 'ISO', '639-3', 'kim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Palawano', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Palawano', 'Region', 'Western', 'Visayas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Palawano', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Agutaynen', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Agutaynen', 'Region', 'MIMAROPA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Agutaynen', 'Language', 'codes', 'ISO', '639-3', 'agn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kagayanen', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```



```

['Kagayanen', 'Region', 'MIMAROPA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Kagayanen', 'Language', 'codes', 'ISO', '639-3', 'cgc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Porohanon', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Porohanon', 'Region', 'Central', 'Visayas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Porohanon', 'Language', 'codes', 'ISO', '639-3', 'prh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Assan', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Assan', 'Language', 'codes', 'ISO', '639-3', 'xss']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baghdad', 'Spoken', 'in', 'Israel', ',', 'Iraq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Baghdad', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Baghdad', 'Writing', 'system', 'Hebrew', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Baghdad', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Buhid', 'Spoken', 'in', 'Philippines']:

```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Buhid', 'Region', 'MIMAROPA']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Buhid', 'Writing', 'system', 'Buhid']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Buhid', 'Language', 'codes', 'ISO', '639-3', 'bku']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tadyawan', 'Spoken', 'in', 'Philippines']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tadyawan', 'Region', 'MIMAROPA']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Tadyawan', 'Language', 'codes', 'ISO', '639-3', 'tdy']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Hanun\xc3\xb3'o", 'Spoken', 'in', 'Philippines']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
["Hanun\xc3\xb3'o", 'Region', 'MIMAROPA']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
["Hanun\xc3\xb3'o", 'Language', 'codes', 'ISO', '639-3', 'hnn']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tawbuid', 'Spoken', 'in', 'Philippines']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tawbuid', 'Region', 'MIMAROPA']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Tawbuid', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Alangan', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Alangan', 'Region', 'MIMAROPA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Alangan', 'Language', 'codes', 'ISO', '639-3', 'alj']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Iraya', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Iraya', 'Region', 'MIMAROPA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Iraya', 'Language', 'codes', 'ISO', '639-3', 'iry']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ixcatec', 'Spoken', 'in', 'Oaxaca', ',', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Ixcatec', 'Total', 'speakers', '8']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Ixcatec', 'Language', 'codes', 'ISO', '639-3', 'ixc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tlapanec', 'Spoken', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Tlapanec', 'Region', 'Guerrero', ',', 'Morelos']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tlapanec', 'Language', 'codes', 'ISO', '639-3', 'tcf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tlapanec', 'Language', 'codes', 'ISO', '639-3', 'tpc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tlapanec', 'Language', 'codes', 'ISO', '639-3', 'tpl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tlapanec', 'Language', 'codes', 'ISO', '639-3', 'tpx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hani', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hani', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hani', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hani', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hani', 'Region', 'Far', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hani', 'Language', 'codes', 'ISO', '639-3', 'hni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dano', 'Spoken', 'in', 'Goroka', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Dano', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Dano', 'Language', 'codes', 'ISO', '639-3', 'aso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hunzib', 'Total', 'speakers', '1840']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hunzib', 'Region', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hunzib', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hunzib', 'Pronunciation', '\xcb', '\x88', 'hont', '\xcd', '\xa1', '\xc9',  
'\xac', '\xca', '\xbc', 'os', 'm', '\xc9', '\xa8', 't', '\xcd', '\xa1',  
's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'hont', '\xcd', '\xa1', '\xc9', '\xac', '\xca', '\xbc', 'os', 'm',  
'\xc9', '\xa8', 't', '\xcd', '\xa1', 's']]], 19, True)

['Hunzib', 'Language', 'codes', 'ISO', '639-3', 'huz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Murrinh-patha', 'Spoken', 'in', 'Wadeye']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Murrinh-patha', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Murrinh-patha', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Murrinh-patha', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Type: ANY', 'Type: ANY']]], 5, True)

['Murrinh-patha', 'Language', 'codes', 'ISO', '639-3', 'mwf']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nunggubuyu', 'Language', 'codes', 'ISO', '639-3', 'nuy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Trique', 'Spoken', 'in', 'Mexico', ';', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  
';', 'USA']]], 6, True)

['Trique', 'Region', 'Oaxaca', ';', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Trique', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Trique', 'Official', 'language', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Trique', 'Regulated', 'by']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Trique', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Trique', 'Language', 'codes', 'ISO', '639-3', ':', 'trc', '\xe2', '\x80',  
'\x93', 'Copala']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'trc', '\xe2', '\x80', '\x93', 'Copala']], ['NUM',  
'639-3']], 11, True)

['Trique', 'Language', 'codes', 'ISO', '639-3', 'trs', '\xe2', '\x80',  
'\x93', 'Chichahuaxtla']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'trs', '\xe2', '\x80', '\x93', 'Chichahuaxtla']], ['NUM',  
'639-3']], 10, True)

['Subtiaba', 'Spoken', 'in', 'Nicaragua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Subtiaba', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Subtiaba', 'Language', 'codes', 'ISO', '639-3', 'sut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cuicatec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cuicatec', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Cuicatec', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Cuicatec', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cuicatec', 'Language', 'codes', 'ISO', '639-3', ':', 'cux', '\xe2',  
'\x80', '\x93', 'Tepeuxila']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'cux', '\xe2', '\x80', '\x93', 'Tepeuxila']], ['NUM',  
'639-3']], 11, True)

['Cuicatec', 'Language', 'codes', 'ISO', '639-3', 'cut', '\xe2', '\x80',  
'\x93', 'Teutila']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cut', '\xe2', '\x80', '\x93', 'Teutila']], ['NUM',  
'639-3']], 10, True)

['Big', 'Spoken', 'in', 'Northwest', 'Malekula', ',', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Northwest', 'Malekula', ',', 'Vanuatu']]], 7, True)

['Big', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Type: ANY', 'Type: ANY']]], 5, True)

['Big', 'Pronunciation', '\xcb', '\x88', '\xce', '\xb8', '\xcc', '\xbc', '\xc9', '\x99', 'n', '\xc9', '\x9b', 'n', 'taut', '\xca', '\xb0']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', '\xce', '\xb8', '\xcc', '\xbc', '\xc9', '\x99', 'n', '\xc9', '\x9b', 'n', 'taut', '\xca', '\xb0']]], 17, True)

['Big', 'Language', 'codes', 'ISO', '639-2', 'map']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Big', 'Language', 'codes', 'ISO', '639-3', 'nmb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mixe', 'Spoken', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mixe', 'Region', 'Oaxaca']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mixe', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mixe', 'Language', 'codes', 'ISO', '639-3', ':', 'mco', '\xe2', '\x80', '\x93', 'Coatl', '\xc3', '\xa1', 'n', 'Mixe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'mco', '\xe2', '\x80', '\x93', 'Coatl', '\xc3', '\xa1', 'n', 'Type: ANY']], ['NUM', '639-3']], 15, True)

['Mixe', 'Language', 'codes', 'ISO', '639-3', 'mzl', '\xe2', '\x80', '\x93', 'Mazatl', '\xc3', '\xa1', 'n', 'Mixe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mzl', '\xe2', '\x80', '\x93', 'Mazatl', '\xc3', '\xa1', 'n', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Zoque', 'Spoken', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zoque', 'Region', 'Chiapas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)



ANY''''], 3, True)

['Zoque', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Zoque', 'Region', 'Tabasco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', ':', 'zoc', '\xe2', '\x80', '\x93', 'Copainal', '\xc3', '\xa1', 'Zoque']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'zoc', '\xe2', '\x80', '\x93', 'Copainal', '\xc3', '\xa1', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', 'zoh', '\xe2', '\x80', '\x93', 'Chimalapa', 'Zoque']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zoh', '\xe2', '\x80', '\x93', 'Chimalapa', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', 'zor', '\xe2', '\x80', '\x93', 'Ray', '\xc3', '\xb3', 'n', 'Zoque']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zor', '\xe2', '\x80', '\x93', 'Ray', '\xc3', '\xb3', 'n', 'Type: ANY']], ['NUM', '639-3']], 14, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', 'zos', '\xe2', '\x80', '\x93', 'Francisco', 'Le', '\xc3', '\xb3', 'n', 'Zoque']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zos', '\xe2', '\x80', '\x93', 'Francisco', 'Le', '\xc3', '\xb3', 'n', 'Type: ANY']], ['NUM', '639-3']], 15, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', 'zoq', '\xe2', '\x80', '\x93', 'Tabasco', 'Zoque']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zoq', '\xe2', '\x80', '\x93', 'Tabasco', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Zoque', 'Language', 'codes', 'ISO', '639-3', 'xep', '\xe2', '\x80', '\x93', '?', 'Epi-0lmec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'xep', '\xe2', '\x80', '\x93', '?', 'Epi-Olmec']], ['NUM', '639-3']], 11, True)

['Nanai', 'Spoken', 'in', 'Russia', ',', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Nanai', 'Region', 'Siberia', ',', 'Heilongjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nanai', 'Language', 'codes', 'ISO', '639-3', 'gld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Guruntum', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Guruntum', 'Language', 'codes', 'ISO', '639-3', 'grd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Apurin\xc3\xa3', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Apurin\xc3\xa3', 'Region', 'Amazonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Apurin\xc3\xa3', 'Language', 'codes', 'ISO', '639-3', 'apu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yokohama', 'Spoken', 'in', 'Yokohama', ',', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Yokohama', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Yokohama', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Yokohama', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Kera', 'Spoken', 'in', 'Chad', ',', 'Cameroon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Kera', 'Language', 'codes', 'ISO', '639-3', 'ker']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sivandi', 'Spoken', 'in', 'Iran']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sivandi', 'Region', 'Fars', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sivandi', 'Language', 'codes', 'ISO', '639-3', 'siy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jambi', 'Spoken', 'in', 'Jambi', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Jambi', 'Language', 'codes', 'ISO', '639-2', 'may', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'may', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Jambi', 'Language', 'codes', 'ISO', '639-2', 'msa', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'msa', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Jambi', 'Language', 'codes', 'ISO', '639-3', 'jax']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jambi', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Maricopa', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maricopa', 'Region', 'Arizona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maricopa', 'Language', 'codes', 'ISO', '639-3', 'mrc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wardaman', 'Total', 'speakers', 'as', 'cited', 'in', 'the', 'Ethnologue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'as', 'cited', 'in', 'the', 'Ethnologue']]], 8, True)

['Wardaman', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wardaman', 'Region', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wardaman', 'Language', 'codes', 'ISO', '639-3', 'wrr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['South', 'Total', 'speakers', '4', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '4']], 5, True)

['South', 'Official', 'language', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['South', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['South', 'Language', 'codes', 'ISO', '639-2', 'umb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'umb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gurindji', 'Spoken', 'in', 'Kalkaringi', 'and', 'Dagaragu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Gurindji', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Gurindji', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gurindji', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gurindji', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Gurindji', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Gurindji', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Pipil', 'Language', 'codes', 'ISO', '639-3', 'ppl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wadikali', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Sre', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sre', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sre', 'Language', 'codes', 'ISO', '639-3', 'kpm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Xaasongaxango', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Xaasongaxango', 'Language', 'codes', 'ISO', '639-3', 'kao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Drehet', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Drehet', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Drehet', 'Language', 'codes', 'ISO', '639-3', 'tlx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Jicaque', 'Spoken', 'in', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jicaque', 'Language', 'codes', 'ISO', '639-3', 'jic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Huave', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Huave', 'Region', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Huave', 'Language', 'codes', 'ISO', '639-3', 'variously']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sierra', 'Spoken', 'in', 'Mexico']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Sierra', 'Region', 'Vera', 'Cruz']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Sierra', 'Language', 'codes', 'ISO', '639-3', 'poi']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Oluta', 'Spoken', 'in', 'Mexico']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Oluta', 'Region', 'Vera', 'Cruz']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Oluta', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Oluta', 'Language', 'codes', 'ISO', '639-3', 'plo']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chatino', 'Spoken', 'in', 'Mexico']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Chatino', 'Region', 'Oaxaca']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Chatino', 'Language', 'codes', 'ISO', '639-3', 'ctp']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sayula', 'Spoken', 'in', 'Mexico']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

['Sayula', 'Region', 'Vera', 'Cruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sayula', 'Language', 'codes', 'ISO', '639-3', 'pos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Texistepec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Texistepec', 'Region', 'Vera', 'Cruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Texistepec', 'Language', 'codes', 'ISO', '639-3', 'poq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luoyang', 'dialect', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Luoyang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Luoyang', 'Region', 'Luoyang', ',', 'Henan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Luoyang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Modern', 'Indo-European', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Modern', 'Date', 'founded', '2006']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Modern', 'Category', '(', 'purpose', ')', 'constructed', 'language']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Modern', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Modern', 'Created', 'by', 'Carlos', 'Quiles', ',', 'Mar', '\\xc3', '\\xad',  
'a', 'Teresa', 'Batalla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
'Carlos', 'Quiles', ',', 'Mar', '\\xc3', '\\xad', 'a', 'Teresa',  
'Batalla']]], 12, True)

['Modern', 'Category', '(', 'sources', ')', 'Indo-European', 'based']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)'], 'Type: ANY', 'Type: ANY']]], 7, True)

['Modern', 'Setting', 'and', 'usage', 'international', 'auxiliary',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Modern', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Betawi', 'Spoken', 'in', 'Jakarta', '(', 'Indonesia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Betawi', 'Language', 'codes', 'ISO', '639-3', 'bew']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mats\\xc3\\xa9s', 'Spoken', 'in', 'Per', '\\xc3', '\\xba', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\\xc3', '\\xba', ',', 'Brazil']]], 8, True)

['Mats\\xc3\\xa9s', 'Language', 'codes', 'ISO', '639-3', 'mcf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baraba', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Baraba', 'Region', 'Siberia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Baraba', 'Language', 'codes', 'ISO', '639-2', 'tat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Baraba', 'Language', 'codes', 'ISO', '639-3', 'tat']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baraba', 'Language', 'codes', 'ISO', '639-1', 'tt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Syrian', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Syrian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Syrian', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Syrian', 'Language', 'codes', 'ISO', '639-2', 'ara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Syrian', 'Language', 'codes', 'ISO', '639-3', 'apc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Isnag', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Isnag', 'Region', 'Luzon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Isnag', 'Language', 'codes', 'ISO', '639-3', 'isd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anyin', 'Spoken', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Anyin', 'Language', 'codes', 'ISO', '639-3', 'any']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'skd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aghul', 'Spoken', 'in', 'Russia', ',', 'also', 'spoken', 'in', 'Azerbaijan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Russia', ',', 'also', 'spoken', 'in', 'Azerbaijan']], 9, True)

['Aghul', 'Region', 'Southeastern', 'Dagestan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aghul', 'Official', 'language', 'in', 'Dagestan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Aghul', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Aghul', 'Language', 'codes', 'ISO', '639-3', 'agx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kitara', 'Spoken', 'in', 'Uganda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kitara', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Kitara', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Nauo', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nauo', 'Language', 'codes', 'ISO', '639-3', 'wiw', '\xe2', '\x80',  
\x93', 'Wirangu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'wiw', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)

['Siculo-Arabic', 'Spoken', 'in', 'Emirate', 'of', 'Sicily']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Emirate', 'of', 'Sicily']]], 6, True)

['Siculo-Arabic', 'Writing', 'system', 'Arabic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Siculo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'sqr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tarahumara', 'Spoken', 'in', 'Chihuahua', ',', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Tarahumara', 'Official', 'language', 'in', 'One', 'of', '63', 'national',  
'languages', 'of', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'One', 'of', 'NUM', 'national', 'languages', 'of', 'Mexico']]],  
['NUM', '63']], 11, True)

['Tarahumara', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de',  
'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P',  
'\xc3', '\xba', 'blica']]], 16, True)

['Tarahumara', 'Language', 'codes', 'ISO', '639-3', 'tar']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Thiin', 'Spoken', 'in', 'Western', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

```
['Thiin', 'Total', 'speakers', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Thiin', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Bar\xc3\xad', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Bar\xc3\xad', 'Total', 'speakers', '1', ',', '700', '(', '1980', ',',  
'2000', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'Type: ANY', 'NUM', ',', 'NUM', 'Type: ANY']]], ['NUM',  
'1'], ['NUM', '700'], ['NUM', '1980'], ['NUM', '2000']], 11, True)
```

```
['Bar\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'mot']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Yamphu', 'Spoken', 'in', 'Nepal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Yamphu', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Yamphu', 'Language', 'codes', 'ISO', '639-3', 'ybi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Aguano', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aguano', 'Spoken', 'in', 'Peru', 'Ethnicity']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Aguano', 'Spoken', 'in', 'Aguano', 'people']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

['Aguano', 'Language', 'codes', 'ISO', '639-3', 'aga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Camsa', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Camsa', 'Language', 'codes', 'ISO', '639-3', 'kbh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pijao', 'Language', 'extinction', '1950s']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pijao', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pijao', 'Region', 'Tolima', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pijao', 'Region', 'Pijao', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pijao', 'Language', 'codes', 'ISO', '639-3', 'pij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chong', 'Spoken', 'in', 'Thailand', '/', 'Cambodia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Thailand', '/', 'Cambodia']], 6, True)

['Chong', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chong', 'Writing', 'system', 'Thai', ',', 'Khmer']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Chong', 'Language', 'codes', 'ISO', '639-3', 'cog']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jawi', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Jawi', 'Total', 'speakers', 'None', 'fully', 'fluent']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'None', 'fully', 'fluent']], 6, True)

['Jawi', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Jawi', 'Language', 'codes', 'ISO', '639-3', 'djm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Korean', 'Total', 'signers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Bolinao', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bolinao', 'Region', 'Bolinao', ',', 'Anda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Bolinao', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Bolinao', 'Language', 'codes', 'ISO', '639-3', 'smk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bari', 'Spoken', 'in', 'Southern', 'Sudan', 'and', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

['Bari', 'Language', 'codes', 'ISO', '639-3', 'bfa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Writing', 'system', 'Hanja']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'oko']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Silla', 'Spoken', 'in', 'Silla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Silla', 'Region', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Silla', 'Language', 'codes', 'ISO', '639-3', 'oko']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gorgani', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gorgani', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gorgani', 'Region', 'NW', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gorgani', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Ewondo', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ewondo', 'Total', 'speakers', '\xe2', '\x80', '\x93']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Ewondo', 'Language', 'codes', 'ISO', '639-3', 'ewo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mambwe', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mambwe', 'Dialects', 'Mambwe', '(', 'Ichimambwe', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Mambwe', 'Dialects', 'Lungu', '(', 'Ichirungu', ',', 'Adong', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Lungu', '\\(', 'Ichirungu', ',', 'Adong', '\\)']], 8, True)

['Mambwe', 'Dialects', 'Fipa-Mambwe', '(', 'Kifipa', 'cha', 'kimambwe', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Fipa-Mambwe', '\\(', 'Kifipa', 'cha', 'kimambwe', '\\)']], 8, True)

['Mambwe', 'Language', 'codes', 'ISO', '639-3', 'mgr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["S\x3\xb6l'ring", 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["S\x3\xb6l'ring", 'Region', 'Sylt', ',', 'Nordfriesland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

["S\x3\xb6l'ring", 'Language', 'codes', 'ISO', '639-2', 'frr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

["S\x3\xb6l'ring", 'Language', 'codes', 'ISO', '639-3', 'frr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["S\x3\xb6l'ring", 'Language', 'codes', 'Linguasphere', '52-ABB-da']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'Linguasphere', 'Type: ANY']]], 5, True)

['Tepehu\xc3\xa1n', 'Dialects', 'Tepecano']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tepehu\xc3\xa1n', 'Region', 'Chihuahua', ',', 'Durango']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tepehu\xc3\xa1n', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tepehu\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tepehu\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'tep', '\xe2', '\x80', '\x93', 'Tepecano']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'tep', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Pima', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pima', 'Region', 'Chihuahua', ',', 'Sonora']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Pima', 'Language', 'codes', 'ISO', '639-3', 'pia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mexicanero', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mexicanero', 'Language', 'codes', 'ISO', '639-2', 'nah']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mexicanero', 'Language', 'codes', 'ISO', '639-3', 'nln']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zande', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Zande', 'Spoken', 'in', 'Central', 'African', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', 'Republic']]], 6, True)

['Zande', 'Spoken', 'in', 'Southern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Zande', 'Language', 'codes', 'ISO', '639-3', 'zne']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lenca', 'Spoken', 'in', 'El', 'Salvador', ',', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Lenca', 'Language', 'codes', 'ISO', '639-3', 'len']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marka', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Marka', 'Language', 'codes', 'ISO', '639-3', 'rkm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Manya', 'Spoken', 'in', 'Liberia', ',', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Manya', 'Language', 'codes', 'ISO', '639-3', 'mzj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rushani', 'Spoken', 'in', 'Afghanistan', ',', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Rushani', 'Total', 'speakers', 'Not', 'known']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Not', 'known']]], 5, True)

['Rushani', 'Writing', 'system', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rushani', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Ngaatjatjarra', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ngaatjatjarra', 'Total', 'speakers', '500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ngaatjatjarra', 'Language', 'codes', 'ISO', '639-3', 'ntj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Awaswas', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Awaswas', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Awaswas', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Awaswas', 'Language', 'codes', 'ISO', '639-3', 'cst']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Omaha\xe2\x80\x93Ponca', 'Total', 'speakers', '85']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Omaha\xe2\x80\x93Ponca', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Omaha\xe2\x80\x93Ponca', 'Region', 'Nebraska', 'and', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Omaha\xe2\x80\x93Ponca', 'Language', 'codes', 'ISO', '639-3', 'oma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yora', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yora', 'Region', 'Manu', 'Park', ',', 'Panagua', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Manu', 'Park',  
,', 'Panagua', 'River']]], 7, True)

['Yora', 'Language', 'codes', 'ISO', '639-3', 'mts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chorotega', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chorotega', 'Spoken', 'in', 'Mangue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chorotega', 'Spoken', 'in', 'Chorotega']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chorotega', 'Spoken', 'in', 'Monimbo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chorotega', 'Language', 'codes', 'ISO', '639-3', 'mom', 'Linguist',  
'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mom', 'Linguist', 'List']], ['NUM', '639-3']], 8, True)

['Chorotega', 'Language', 'codes', 'ISO', '639-3', 'mom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Chiapanec', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Chiapanec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Chiapanec', 'Region', 'Chiapas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Chiapanec', 'Language', 'codes', 'ISO', '639-3', 'cip']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pame', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Pame', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chichimeca', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Chichimeca', 'Region', 'Guanajuato', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Chichimeca', 'Language', 'codes', 'ISO', '639-3', 'pei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tamyen', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', '\\\\(', 'Type: ANY', '\\\\)']], 8, True)
```

```
['Tamyen', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Tamyen', 'Writing', 'system', 'Latin', 'alphabet']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tamyen', 'Language', 'codes', 'ISO', '639-3', 'cst']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chochenyo', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'States', '\\\\(', 'Type: ANY', '\\\\)']]], 8, True)

['Chochenyo', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chochenyo', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chochenyo', 'Language', 'codes', 'ISO', '639-3', 'cst']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Tupiniquim', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tupiniquim', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tupiniquim', 'Region', 'Espirito', 'Santo', ',', 'Bahia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Espirito', 'Santo', ',', 'Bahia']]], 6, True)

['Tupiniquim', 'Language', 'codes', 'ISO', '639-3', 'tpx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chico', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chico', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```

['Chico', 'Region', 'California']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Chico', 'Language', 'codes', 'ISO', '639-3', 'vmv']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Imbo', 'Language', 'codes', 'ISO', '639-3', 'imo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lango', 'Spoken', 'in', 'Uganda']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Lango', 'Language', 'codes', 'ISO', '639-3', 'laj']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lango', 'Spoken', 'in', 'Sudan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Lango', 'Language', 'codes', 'ISO', '639-3', 'lno']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nisenan', 'Total', 'speakers', '1']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nisenan', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Nisenan', 'Region', 'California']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Nisenan', 'Language', 'codes', 'ISO', '639-3', 'nsz']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rangpuri', 'Spoken', 'in', 'India', ',', 'Bangladesh']:

```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Rangpuri', 'Total', 'speakers', '15', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '15']], 5, True)

['Rangpuri', 'Language', 'codes', 'ISO', '639-3', 'rkt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koryo-mar', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koryo-mar', 'Spoken', 'in', 'Kazakhstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koryo-mar', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koryo-mar', 'Spoken', 'in', 'Kyrgyzstan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koryo-mar', 'Writing', 'system', 'Hangul', ';', 'Cyrillic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Koryo-mar', 'Language', 'codes', 'ISO', '639-3', 'kor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huambisa', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']]], 6, True)

['Huambisa', 'Language', 'codes', 'ISO', '639-3', 'hub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bahamas', 'Spoken', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3, True)

['Bahamas', 'Language', 'codes', 'ISO', '639-3', 'bah']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Bahamas', 'Language', 'codes', 'Linguasphere', '52-ABB-an']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY' ] ] ], 5, True)

['Wik-Mungknh', 'Spoken', 'in', 'Australia']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Wik-Mungknh', 'Language', 'codes', 'ISO', '639-3', 'wim']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Wik-Ngathan', 'Spoken', 'in', 'Australia']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Wik-Ngathan', 'Language', 'codes', 'ISO', '639-3', 'wig']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Pochutec', 'Spoken', 'in', 'Pochutla']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Pochutec', 'Spoken', 'in', 'Oaxaca']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Pochutec', 'Spoken', 'in', 'Mexico']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY' ] ] ], 4, True)

['Pochutec', 'Writing', 'system', 'Latin', 'alphabet']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY' ] ] ], 5, True)

['Pochutec', 'Language', 'codes', 'ISO', '639-3', 'xpo']:  
 ([[ 'Start', [ 'Phrase' ] ], [ 'Phrase', [ 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY' ] ], [ 'NUM', '639-3' ] ], 6, True)

['Canela', 'Spoken', 'in', 'Brazil']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Canela', 'Region', 'Maranh', '\xc3', '\xa3', 'o', ',', 'southeastern', 'Par', '\xc3', '\xa1']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maranh', '\xc3', '\xa3', 'o', ',', 'southeastern', 'Par', '\xc3', '\xa1']]], 11, True)
```

```
['Canela', 'Language', 'codes', 'ISO', '639-3', 'ram']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Matlatzinca', 'Spoken', 'in', 'Mexico']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Matlatzinca', 'Region', 'State', 'of', 'Mexico']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Matlatzinca', 'Language', 'codes', 'ISO', '639-3', 'either']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Matlatzinca', 'Language', 'codes', 'ISO', '639-3', ':', 'ocu', '\xe2', '\x80', '\x93', 'Ocuiltec', '/', 'Tlahuica']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ocu', '\xe2', '\x80', '\x93', 'Ocuiltec', '/', 'Tlahuica']]], ['NUM', '639-3']], 13, True)
```

```
['Matlatzinca', 'Language', 'codes', 'ISO', '639-3', 'mat', '\xe2', '\x80', '\x93', 'Matlatzinca']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mat', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)
```

```
['Animere', 'Spoken', 'in', 'Ghana']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Animere', 'Language', 'codes', 'ISO', '639-3', 'anf']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tiriy\xc3\xb3', 'Region', 'Tiriy', '\xc3', '\xb3', 'people']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Tiriy',  
'\xc3', '\xb3', 'people']]], 6, True)

['Tiriy\xc3\xb3', 'Spoken', 'in', 'Brazil', ',', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Tiriy\xc3\xb3', 'Pronunciation', 'ta', '\xc9', '\xbd', '\xc9', '\x99',  
'\xcb', '\x90', 'no', 'ijo', '\xcb', '\x90', 'mi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ta',  
'\xc9', '\xbd', '\xc9', '\x99', '\xcb', '\x90', 'no', 'ijo', '\xcb',  
'\x90', 'mi']]], 14, True)

['Tiriy\xc3\xb3', 'Language', 'codes', 'ISO', '639-3', 'tri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sangsari', 'Region', 'Elburz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Sangsari', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sangsari', 'Language', 'codes', 'ISO', '639-3', 'sgr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Iwal', 'Region', 'Morobe', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Iwal', 'Language', 'codes', 'ISO', '639-3', 'kbm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yorta', 'Spoken', 'in', 'Victoria', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Yorta', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Yorta', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',

```

'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Yorta', 'Language', 'codes', 'ISO', '639-3', '0hm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Boro', 'Spoken', 'in', 'Ghana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Boro', 'Region', 'central', 'eastern', 'Ghana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)

['Boro', 'Language', 'codes', 'ISO', '639-3', 'xxb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Santiago', 'Spoken', 'in', 'Argentina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Santiago', 'Language', 'codes', 'ISO', '639-2', 'qus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Santiago', 'Language', 'codes', 'ISO', '639-3', 'qus']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Batak', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Batak', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Batak',
'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Batak', 'Language', 'codes', 'ISO', '639-3', 'bbc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

```
['Sora', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sora', 'Language', 'codes', 'ISO', '639-3', 'srb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Korku', 'Language', 'codes', 'ISO', '639-3', 'kfq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Hangzhou', 'dialect', 'Total', 'speakers', 'approx', '.', '3-4',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '3-4']], 8, True)
```

```
['Hangzhou', 'Region', 'Hangzhou', ',', 'Zhejiang', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
',', 'Zhejiang', 'province']]], 6, True)
```

```
['Hangzhou', 'Pronunciation', 'han', '.', 'tsei', '.', 'wu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'han',  
'.', 'tsei', '.', 'wu']]], 7, True)
```

```
['Hangzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Hangzhou', 'Language', 'codes', 'ISO', '639-3', 'wuu-haz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kugu-Uwanh', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kugu-Uwanh', 'Language', 'codes', 'ISO', '639-3', 'uwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Total', 'speakers', '1', ',', '600', ',', '000', '(', 'est',  
'.', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
```

['NUM', ',', 'NUM', ',', 'NUM', '\\(', 'est', '.', '\\)']], ['NUM', '1'],  
['NUM', '600'], ['NUM', '000']], 12, True)

['Southern', 'Spoken', 'in', 'Netherlands', ',', 'Belgium', 'and',  
'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Netherlands', ',', 'Belgium', 'and', 'Germany']]]], 8, True)

['Southern', 'Region', 'Limburg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['Southern', 'Region', 'North', 'Rhine-Westphalia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Southern', 'Region', 'Province', 'of', 'Liege']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Palpa', 'Language', 'codes', 'ISO', '639-3', 'plp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wayuu', 'Region', 'Wayuu', 'people']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Wayuu', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)

['Wayuu', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Wayuu', 'Pronunciation', 'wajunaiki']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['Wayuu', 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',

'in', 'Type: ANY']]], 5, True)

['Wayuu', 'Regulated', 'by', 'Centro', 'Etnoeducativo', 'Kamusuchiwo',  
'\xe2', '\x80', '\x99', 'u']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Centro', 'Etnoeducativo', 'Kamusuchiwo', '\xe2', '\x80', '\x99', 'u']]],  
10, True)

['Wayuu', 'Language', 'codes', 'ISO', '639-3', 'guc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Argentine', 'Signed', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Argentine', 'Total', 'signers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Central', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Central', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Central', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Central', 'Language', 'codes', 'ISO', '639-2', 'ayr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'ayr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chirip\xc3\xa1', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chirip\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',



'Type: ANY']]], 4, True)

['Chirip\xc3\xa1', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chirip\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'nhd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eklenhui', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eklenhui', 'Total', 'speakers', '800']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eklenhui', 'Language', 'codes', 'ISO', '639-3', 'crt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Iyo'wujwa", 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Iyo'wujwa", 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Iyo'wujwa", 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Iyo'wujwa", 'Language', 'codes', 'ISO', '639-3', 'crq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mby\xc3\xa1', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mby\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Mby\xc3\xa1', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mby\xc3\xa1', 'Language', 'codes', 'ISO', '639-2', 'grn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['Mby\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'gun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kai\xc3\xa1', 'Spoken', 'in', 'Argentina', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Kai\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'kgk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Nivac\xc3\xa1', 'Spoken', 'in', 'Argentina', ',', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Nivac\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'cag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Pilag\xc3\xa1', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pilag\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'plg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Puelche', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Puelche', 'Total', 'speakers', '5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Puelche', 'Language', 'codes', 'ISO', '639-3', 'pue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tapiet\xc3\xa9', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tapiet\xc3\xa9', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tapiet\xc3\xa9', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tapiet\xc3\xa9', 'Total', 'speakers', '203']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tapiet\xc3\xa9', 'Language', 'codes', 'ISO', '639-2', 'grn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tapiet\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'tpj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tehuelche', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tehuelche', 'Total', 'speakers', '4']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tehuelche', 'Language', 'codes', 'ISO', '639-3', 'teh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vilela', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Vilela', 'Total', 'speakers', '20']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Vilela', 'Language', 'codes', 'ISO', '639-3', 'vil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wich\xc3\xad', 'Spoken', 'in', 'Argentina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Wich\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'mzh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wich\xc3\xad', 'Spoken', 'in', 'Argentina', ',', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Wich\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'mtp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wich\xc3\xad', 'Spoken', 'in', 'Argentina', ',', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Wich\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'wlv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chilean', 'Signed', 'in', 'Chile']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Chilean', 'Total', 'signers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Chilean', 'Spoken', 'in', 'Chile']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Chilean', 'Total', 'speakers', '8000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Chilean', 'Language', 'codes', 'ISO', '639-2', 'que']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chilean', 'Language', 'codes', 'ISO', '639-3', 'cqu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Changsha', 'dialect', 'Total', 'speakers', 'approx', '.', '6',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '6']], 8, True)

['Changsha', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Changsha', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Changsha', 'Language', 'codes', 'ISO', '639-3', 'hsn-cha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kunza', 'Spoken', 'in', 'Atacama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kunza', 'Language', 'codes', 'ISO', '639-3', 'kuz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Haitian', 'Spoken', 'in', 'Haiti', ',', 'Haitian', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY', 'diaspora']], 7, True)

['Haitian', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Haitian', 'Language', 'codes', 'ISO', '639-2', 'cpf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Haitian', 'Language', 'codes', 'ISO', '639-3', 'hvc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koromfe', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Koromfe', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Koromfe', 'Language', 'codes', 'ISO', '639-3', 'kfz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shuangfeng', 'dialect', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shuangfeng', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shuangfeng', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Shuangfeng', 'Language', 'codes', 'ISO', '639-3', 'hsn-luo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wenzhou', 'Total', 'speakers', '5', 'million', 'est']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'million', 'est']], ['NUM', '5']], 6, True)

['Wenzhou', 'Spoken', 'in', 'Wenzhou', 'Prefecture']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Wenzhou', 'Spoken', 'in', 'Zhejiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wenzhou', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wenzhou', 'Region', 'Southeastern', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wenzhou', 'Region', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wenzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Wenzhou', 'Language', 'codes', 'ISO', '639-3', 'wuu-ouj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ovambo', 'Standard', 'forms', 'Kwanyama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ovambo', 'Standard', 'forms', 'Ndonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ovambo', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ovambo', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-2', 'kua', '\xe2', '\x80', '\x93', 'Kwanyama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kua', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-2', 'ndo', '\xe2', '\x80', '\x93', 'Ndonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'ndo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-2']], 10, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-3', ':', 'kua', '\xe2', '\x80', '\x93', 'Kwanyama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'kua', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-3', 'ndo', '\xe2', '\x80', '\x93', 'Ndonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ndo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-3', 'The', 'distribution', 'of', 'Oshiwambo', 'languages', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'The', 'distribution', 'of', 'Oshiwambo', 'languages', 'in', 'Namibia']], ['NUM', '639-3']], 12, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-1', 'kj', '\xe2', '\x80', '\x93', 'Kwanyama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'kj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-1']], 10, True)

['Ovambo', 'Language', 'codes', 'ISO', '639-1', 'ng', '\xe2', '\x80', '\x93', 'Ndonga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ng', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-1']], 10, True)

['Pohnpeian', 'Spoken', 'in', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pohnpeian', 'Region', 'Pohnpei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Pohnpeian', 'Writing', 'system', 'Latin', 'alphabet']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Pohnpeian', 'Language', 'codes', 'ISO', '639-2', 'pon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Pohnpeian', 'Language', 'codes', 'ISO', '639-3', 'pon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Pingelapese', 'Spoken', 'in', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pingelapese', 'Region', 'Pingelap']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Pingelapese', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Pingelapese', 'Language', 'codes', 'ISO', '639-2', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Pingelapese', 'Language', 'codes', 'ISO', '639-3', 'pif']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mokilese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Mokilese', 'Spoken', 'in', 'Micronesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mokilese', 'Region', 'Mokil', 'atoll']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mokilese', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

```

['Mokilese', 'Language', 'codes', 'ISO', '639-2', 'none']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mokilese', 'Language', 'codes', 'ISO', '639-3', 'mkj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Araona', 'Spoken', 'in', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Araona', 'Total', 'speakers', '81']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Araona', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

['Araona', 'Language', 'codes', 'ISO', '639-3', 'aro']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bolivian', 'Signed', 'in', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Cavine\xc3\xb1a', 'Spoken', 'in', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Cavine\xc3\xb1a', 'Official', 'language', 'in', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']], 5, True)

['Cavine\xc3\xb1a', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']], 6, True)

['Cavine\xc3\xb1a', 'Language', 'codes', 'ISO', '639-3', 'cav']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Spoken', 'in', 'Bolivia']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Spoken', 'in', 'Argentina']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Spoken', 'in', 'Paraguay']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Language', 'codes', 'ISO', '639-2', 'grn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'gui']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Macushi', 'Spoken', 'in', 'Macushi', 'people']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Macushi', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macushi', 'Spoken', 'in', 'Guyana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macushi', 'Spoken', 'in', 'Venezuela', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Macushi', 'Language', 'codes', 'ISO', '639-3', 'mbc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ter\xc3\xaana', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ter\xc3\xaana', 'Language', 'codes', 'ISO', '639-3', 'ter']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Avokaya', 'Language', 'codes', 'ISO', '639-3', 'avu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eskayan', 'Total', 'speakers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Eskayan', 'Region', 'Bohol']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Eskayan', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Eskayan', 'Writing', 'system', 'Eskayan', 'script', '(', 'syllabary',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Eskayan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

["Jian'ou", 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)

["Jian'ou", 'Spoken', 'in', 'Southern', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

["Jian'ou", 'Region', "Jian'ou", ',', 'Fujian']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

["Jian'ou", 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

["Jian'ou", 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

```

```
["Jian'ou", 'Language', 'codes', 'ISO', '639-3', 'mnp']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Jian'ou", 'Language', 'codes', 'ISO', '639-1', 'zh']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Leizhou', 'Spoken', 'in', 'China']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Leizhou', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',
'States', '\\(', 'Type: ANY', '\\)']], 8, True)
```

```
['Leizhou', 'Spoken', 'in', 'Taiwan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Leizhou', 'Spoken', 'in', 'Malaysia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Leizhou', 'Spoken', 'in', 'Singapore']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Leizhou', 'Spoken', 'in', 'Hong', 'Kong']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Leizhou', 'Region', 'Leizhou']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Leizhou', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',
'Linguist', 'List']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)
```

```
['Leizhou', 'Language', 'codes', 'ISO', '639-3', 'nan-lei']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Poqomchi'", 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Poqomchi'", 'Region', 'Baja', 'Verapaz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Poqomchi'", 'Language', 'codes', 'ISO', '639-3', 'poh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Poqomam', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Poqomam', 'Region', 'Alta', 'Verapaz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Poqomam', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Poqomam', 'Regulated', 'by', 'Academia', 'de', 'Lenguas', 'Mayas', 'de',  
'Guatemala', '(', 'ALMG', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Academia', 'de', 'Lenguas', 'Mayas', 'de', 'Guatemala', '\\(', 'ALMG', '\\  
)']]], 12, True)
```

```
['Poqomam', 'Recognised', 'minority', 'language', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)
```

```
['Poqomam', 'Language', 'codes', 'ISO', '639-3', 'poc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sakapultek', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sakapultek', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Sakapultek', 'Regulated', 'by', 'Academia', 'de', 'Lenguas', 'Mayas',  
'de', 'Guatemala', '(', 'ALMG', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Academia', 'de', 'Lenguas', 'Mayas', 'de', 'Guatemala', '\\(', 'ALMG', '\\  
)']]], 12, True)
```

```
['Sakapultek', 'Recognised', 'minority', 'language', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Type: ANY']]], 6, True)
```

```
['Sakapultek', 'Language', 'codes', 'ISO', '639-3', 'quv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sipakapense', 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sipakapense', 'Region', 'Sipacapa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sipakapense', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Sipakapense', 'Regulated', 'by', 'ALMG']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sipakapense', 'Language', 'codes', 'ISO', '639-3', 'qum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Q'anjo'bal", 'Spoken', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Q'anjo'bal", 'Region', 'Huehuetenango']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
["Q'anjo'bal", 'Language', 'codes', 'ISO', '639-3', 'kjb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Mocho\xe2\x80\x99', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mocho\xe2\x80\x99', 'Language', 'codes', 'ISO', '639-3', 'mhc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Writing', 'system', 'Armenian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Middle', 'Language', 'codes', 'ISO', '639-3', 'axm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyanga', 'Spoken', 'in', 'Kivu', 'Province', ',', 'Walikale',  
'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Kivu',  
'Province', ',', 'Walikale', 'Territory']]], 8, True)

['Nyanga', 'Language', 'codes', 'ISO', '639-3', 'nyj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ancient', 'Spoken', 'in', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ancient', 'Writing', 'system', 'South', 'Arabian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Ancient', 'Language', 'codes', 'ISO', '639-3', 'xna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hamer-Banna', 'Total', 'speakers', '47', ',', '531', 'of', 'Hamer',  
'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'of', 'Hamer', 'dialect']], ['NUM', '47'], ['NUM',  
'531']], 9, True)

['Hamer-Banna', 'Total', 'speakers', '26', ',', '869', 'of', 'Banna',  
'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',



['NUM', ',', 'NUM', 'of', 'Banna', 'dialect']], ['NUM', '26'], ['NUM', '869']], 9, True)

['Hamer-Banna', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hamer-Banna', 'Region', 'South', 'Omo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hamer-Banna', 'Language', 'codes', 'ISO', '639-3', 'amf']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ancient', 'Spoken', 'in', 'Arabia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ancient', 'Writing', 'system', 'South', 'Arabian', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Ancient', 'Language', 'codes', 'ISO', '639-3', 'xna']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tidore', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tidore', 'Region', 'North', 'Maluku']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tidore', 'Region', 'islands', 'of', 'Tidore']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tidore', 'Region', 'Maitara']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tidore', 'Region', 'Mare']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tidore', 'Region', 'northern', 'half', 'of', 'Moti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Tidore', 'Region', 'and', 'some', 'areas', 'of', 'west', 'coast', 'of',  
'Halmahera', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'some',  
'areas', 'of', 'west', 'coast', 'of', 'Halmahera', '.', 'Coordinates']]],  
12, True)

['Tidore', 'Language', 'codes', 'ISO', '639-3', 'tvo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Udege', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Udege', 'Region', 'Russian', 'Far', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Udege', 'Language', 'codes', 'ISO', '639-3', 'ude']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ancient', 'Spoken', 'in', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ancient', 'Writing', 'system', 'South', 'Arabian', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Ancient', 'Language', 'codes', 'ISO', '639-3', 'xna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nabataean', 'Writing', 'system', 'Nabataean', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nabataean', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Nabataean', 'Language', 'codes', 'ISO', '639-3', 'qhy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nupe', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Nupe', 'Region', 'Niger', 'State', ',', 'Federal', 'Capital', 'Territory']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Niger', 'State', ',', 'Federal', 'Capital', 'Territory']], 8, True)
```

```
['Nupe', 'Language', 'codes', 'ISO', '639-3', 'nup']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['English', 'Total', 'speakers', 'First', 'language', ':', '309', '\xe2', '\x80', '\x93', '400', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'First', 'language', ':', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']], ['NUM', '309'], ['NUM', '400']], 12, True)
```

```
['English', 'Spoken', 'in', '(', 'see', 'below', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\\(', 'see', 'below', '\\)']], 7, True)
```

```
['English', 'Writing', 'system', 'Latin', '(', 'English', 'variant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)
```

```
['English', 'Official', 'language', 'in', '53', 'countries']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['English', 'Official', 'language', 'in', 'United', 'Nations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['English', 'Official', 'language', 'in', 'European', 'Union']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']], 6, True)
```

```
['English', 'Official', 'language', 'in', 'Commonwealth', 'of', 'Nations']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Commonwealth', 'of', 'Nations']]]], 7, True)
```

```
['English', 'Official', 'language', 'in', 'CoE']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Official', 'language', 'in', 'NATO']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Official', 'language', 'in', 'NAFTA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Official', 'language', 'in', 'OAS']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Official', 'language', 'in', 'OIC']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Official', 'language', 'in', 'PIF']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Official', 'language', 'in', 'UKUSA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)
```

```
['English', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]]], 6, True)
```

```
['English', 'Language', 'codes', 'ISO', '639-2', 'eng']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)
```

```
['English', 'Language', 'codes', 'ISO', '639-3', 'eng']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['English', 'Language', 'codes', 'Linguasphere', '52-ABA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'Linguasphere', 'Type: ANY']]], 5, True)

['English', 'Language', 'codes', 'ISO', '639-1', 'en']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sentinel', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sentinel', 'Region', 'North', 'Sentinel', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sentinel', 'Language', 'codes', 'ISO', '639-2', 'mis']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sentinel', 'Language', 'codes', 'ISO', '639-3', 'std']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paipai', 'Spoken', 'in', 'Mexico', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Paipai', 'Spoken', 'in', 'Paipai', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Paipai', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Paipai', 'Language', 'codes', 'ISO', '639-3', 'ppi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsakhur', 'Spoken', 'in', 'Azerbaijan', ',', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Tsakhur', 'Language', 'codes', 'ISO', '639-3', 'tkr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Havasupai\xe2\x80\x93Hualapai', 'Spoken', 'in', 'Arizona', ',', 'USA', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Arizona', ',', 'USA', 'Ethnicity']]], 7, True)

['Havasupai\xe2\x80\x93Hualapai', 'Spoken', 'in', '565', 'Havasupai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'NUM', 'Type: ANY']], ['NUM', '565']], 5, True)

['Havasupai\xe2\x80\x93Hualapai', 'Spoken', 'in', '1']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Havasupai\xe2\x80\x93Hualapai', 'Spoken', 'in', '872', 'Walapai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'NUM', 'Type: ANY']], ['NUM', '872']], 5, True)

['Havasupai\xe2\x80\x93Hualapai', 'Total', 'speakers', '530', 'Havasupai', ',', '1', ',', '000', 'Walapai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'Havasupai', ',', 'NUM', ',', 'NUM', 'Walapai']]], ['NUM', '530'], ['NUM', '1'], ['NUM', '000']], 10, True)

['Havasupai\xe2\x80\x93Hualapai', 'Dialects', 'Havasupai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Havasupai\xe2\x80\x93Hualapai', 'Dialects', 'Hualapai']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Havasupai\xe2\x80\x93Hualapai', 'Language', 'codes', 'ISO', '639-3', 'yuf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lahu', 'Spoken', 'in', 'Yunnan', ',', 'China', ';', 'Thailand', ';', 'Laos', ';', 'Myanmar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Yunnan', ',', 'China', ';', 'Thailand', ';', 'Laos', ';', 'Type: ANY']]], 12, True)

['Lahu', 'Writing', 'system', 'Latin', '(', 'Lahu', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Lahu', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Lahu', 'Language', 'codes', 'ISO', '639-3', 'lhu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xibe', 'Region', 'Xinjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Xibe', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Xibe', 'Writing', 'system', 'Manchu', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Xibe', 'Language', 'codes', 'ISO', '639-3', 'sjo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Negidal', 'Total', 'speakers', '100-175']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Negidal', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Negidal', 'Region', 'Russian', 'Far', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Negidal', 'Language', 'codes', 'ISO', '639-3', 'neg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Apma', 'Total', 'speakers', '7800']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Apma', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Apma', 'Region', 'Pentecost', 'island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Apma', 'Language', 'codes', 'ISO', '639-3', 'app']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sowa', 'Total', 'speakers', 'approximately', '10']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'approx', '.']], 5, True)

['Sowa', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sowa', 'Region', 'Pentecost', 'island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sowa', 'Language', 'codes', 'ISO', '639-3', 'sww']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngala', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Ngala', 'Language', 'codes', 'ISO', '639-3', 'bxg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ipai', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ipai', 'Language', 'codes', 'ISO', '639-3', 'dih']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kumeyaay', 'Spoken', 'in', 'USA', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Kumeyaay', 'Region', 'California', ',', 'Baja', 'California']:



```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'California',  
,',', 'Baja', 'California']]], 6, True)
```

```
['Kumeyaay', 'Language', 'codes', 'ISO', '639-3', 'dih', '\xe2', '\x80',  
\x93', 'Campo']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'dih', '\xe2', '\x80', '\x93', 'Campo']], ['NUM', '639-3']],  
10, True)
```

```
['Tiipai', 'Spoken', 'in', 'USA', ',,', 'Mexico']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',,', 'Type: ANY']]], 6, True)
```

```
['Tiipai', 'Region', 'California', ',,', 'Baja', 'California']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'California',  
,',', 'Baja', 'California']]], 6, True)
```

```
['Tiipai', 'Language', 'codes', 'ISO', '639-3', 'dih', '\xe2', '\x80',  
\x93', 'Campo']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'dih', '\xe2', '\x80', '\x93', 'Campo']], ['NUM', '639-3']],  
10, True)
```

```
['Sui', 'Spoken', 'in', 'China']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sui', 'Spoken', 'in', 'Vietnam']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sui', 'Region', 'Guangxi']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sui', 'Region', 'Yunnan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sui', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Sui', 'Language', 'codes', 'ISO', '639-3', 'swi']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Uru', 'Spoken', 'in', 'Uru', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Uru', 'Total', 'speakers', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Uru', 'Language', 'codes', 'ISO', '639-3', 'ure']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Digo', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Digo', 'Region', 'Mombasa', 'and', 'Kwale', 'districts', 'in', 'Kenya',  
';', 'Muheza', 'and', 'Tanga', 'districts', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mombasa',  
'and', 'Kwale', 'districts', 'in', 'Kenya', ';', 'Muheza', 'and', 'Tanga',  
'districts', 'in', 'Tanzania']]], 15, True)
```

```
['Digo', 'Language', 'codes', 'ISO', '639-3', 'dig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Brianz\xc3\xb6/Brianzoeu', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Brianz\xc3\xb6/Brianzoeu', 'Language', 'codes', 'ISO', '639-3',  
'lmo', '\xe2', '\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)
```

```
['Canz\xc3\xa9s', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Canz\xc3\xa9s', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2',  
'\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 10, True)
```

['Numbami', 'Region', 'coastal', 'village', 'in', 'Morobe', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'coastal',  
'village', 'in', 'Morobe', 'Province']]], 7, True)

['Numbami', 'Language', 'codes', 'ISO', '639-3', 'sij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bor\xc5\xa9ca', 'Spoken', 'in', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Bor\xc5\xa9ca', 'Total', 'speakers', 'few']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bor\xc5\xa9ca', 'Language', 'codes', 'ISO', '639-3', 'brn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chakma', 'Spoken', 'in', 'Bangladesh', 'and', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Chakma', 'Region', 'Chittagong', 'Hill', 'Tracts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Chakma', 'Region', 'Mizoram']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Chakma', 'Region', 'Arunachal', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chakma', 'Region', 'Tripura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Chakma', 'Language', 'codes', 'ISO', '639-3', 'ccp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Scandoromani', 'Spoken', 'in', 'Sweden', ',', 'Norway']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Scandoromani', 'Official', 'language', 'in', 'recognised', 'minority', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'minority', 'language', 'in']]], 8, True)

['Scandoromani', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Scandoromani', 'Language', 'codes', 'ISO', '639-3', 'rmu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Portuguese', 'Signed', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['\xc3\x93pata', 'Language', 'codes', 'ISO', '639-3', 'opt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tepehua', 'Spoken', 'in', 'Mexico', ':', 'Puebla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', ':', 'Puebla']]], 6, True)

['Tepehua', 'Spoken', 'in', 'Veracruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tepehua', 'Spoken', 'in', 'Hidalgo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tepehua', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Tepehua', 'Regulated', 'by', 'INALI']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tepehua', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['French', 'Spoken', 'in', 'French', 'Guiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['French', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['French', 'Language', 'codes', 'ISO', '639-3', 'gcr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Aguaruna', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']]], 6, True)

['Aguaruna', 'Language', 'codes', 'ISO', '639-3', 'agr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Amahuaca', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']]], 8, True)

['Amahuaca', 'Language', 'codes', 'ISO', '639-3', 'amc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Amarakaeri', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']]], 6, True)

['Amarakaeri', 'Language', 'codes', 'ISO', '639-3', 'amr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lampung', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lampung', 'Region', 'Southern', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Lampung', 'Writing', 'system', 'Ulu', 'script', ',', 'Roman',
'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Lampung', 'Language', 'codes', 'ISO', '639-3', 'ljp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Arabela', 'Spoken', 'in', 'Per', '\xc3', '\xba']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',
'\xc3', '\xba']]], 6, True)

['Arabela', 'Official', 'language', 'in', 'Peru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Arabela', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Arabela', 'Language', 'codes', 'ISO', '639-3', 'arl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ash\xc3\xa1ninka', 'Spoken', 'in', 'Per', '\xc3', '\xba']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',
'\xc3', '\xba']]], 6, True)

['Ash\xc3\xa1ninka', 'Language', 'codes', 'ISO', '639-3', 'cni']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Central', 'Franconian', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Central', 'Spoken', 'in', 'Germany']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Central', 'Spoken', 'in', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Central', 'Spoken', 'in', 'Belgium']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Central', 'Spoken', 'in', 'Netherlands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Central', 'Region', 'North', 'Rhine-Westphalia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Central', 'Region', 'Rhineland-Palatinate']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Central', 'Region', 'Saarland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Central', 'Region', 'Lorraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Central', 'Region', 'Li', '\xc3', '\xa8', 'ge']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Li', '\xc3', '\xa8', 'ge']]], 6, True)

['Central', 'Region', 'Limburg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'goh-mid', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'goh-mid', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 9, True)

['Ash\xc3\xa9ninka', 'Spoken', 'in', 'Per', '\xc3', '\xba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']]], 6, True)

['Ash\xc3\xa9ninka', 'Language', 'codes', 'ISO', '639-3', 'cjo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ash\xc3\xa9ninka', 'Spoken', 'in', 'Per', '\xc3', '\xba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Ash\xc3\xa9ninka', 'Language', 'codes', 'ISO', '639-3', 'prq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ash\xc3\xa9ninka', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Ash\xc3\xa9ninka', 'Language', 'codes', 'ISO', '639-3', 'cpu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ash\xc3\xa9ninka', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Ash\xc3\xa9ninka', 'Language', 'codes', 'ISO', '639-3', 'cpy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ash\xc3\xa9ninka', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba', ',', 'Brazil']]], 8, True)

['Ash\xc3\xa9ninka', 'Language', 'codes', 'ISO', '639-3', 'cpb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Awngi', 'Dialects', 'Kunf', '\xc3', '\xa4', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Kunf',  
'\xc3', '\xa4', 'l']]], 6, True)

['Awngi', 'Region', 'Agew', 'Awi', 'Zone', 'Amhara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Agew', 'Awi',  
'Zone', 'Amhara']]], 6, True)

['Awngi', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Awngi', 'Language', 'codes', 'ISO', '639-3', 'awn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Khwarshi', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khwarshi', 'Region', 'Southwestern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Khwarshi', 'Language', 'codes', 'ISO', '639-3', 'khv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Judeo-Moroccan', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Moroccan', 'Spoken', 'in', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Moroccan', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Moroccan', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Judeo-Moroccan', 'Language', 'codes', 'ISO', '639-3', 'aju']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mazahua', 'Spoken', 'in', 'Mexico', ':', 'Estado', 'de', 'M', '\xc3', '\xa9', 'xico', ',', 'Michoac', '\xc3', '\xa1', 'n']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', ':', 'Estado', 'de', 'M', '\xc3', '\xa9', 'xico', ',', 'Michoac', '\xc3', '\xa1', 'n']], 16, True)

['Mazahua', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Mazahua', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',

'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P',  
'\xc3', '\xba', 'blica']]], 16, True)

['Mazahua', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mazahua', 'Language', 'codes', 'ISO', '639-3', ':', 'mmc', '\xe2',  
'\x80', '\x93', 'Michoacan', 'Mazahua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'mmc', '\xe2', '\x80', '\x93', 'Michoacan', 'Type:  
ANY']], ['NUM', '639-3']], 12, True)

['Southern', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Southern', 'Total', 'speakers', 'No', 'estimates', 'available']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'No', 'estimates', 'available']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'aym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'ayc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cahuarano', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']], 6, True)

['Cahuarano', 'Language', 'codes', 'ISO', '639-3', 'cah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Spaniard', 'Region', 'Undetermined', 'central-interior', 'region', 'of',  
'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Undetermined',  
'central-interior', 'region', 'of', 'Spain']], 7, True)

['Spaniard', 'Signed', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Spaniard', 'Language', 'codes', 'ISO', '639-2', 'sgn-ES']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Catalan', 'Region', 'Catalonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Catalan', 'Signed', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Candoshi', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']], 6, True)

['Candoshi', 'Dialects', 'Kandoashi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Candoshi', 'Dialects', 'Chapara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Candoshi', 'Language', 'codes', 'ISO', '639-3', 'cbu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Heiltsuk-Oowekyala', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Heiltsuk-Oowekyala', 'Region', 'al', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Heiltsuk-Oowekyala', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Heiltsuk-Oowekyala', 'Region', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Heiltsuk-Oowekyala', 'Dialects', 'Heiltsuk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

ANY'']], 3, True)

['Heiltsuk-Oowekyala', 'Dialects', 'Oowekyala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Heiltsuk-Oowekyala', 'Language', 'codes', 'ISO', '639-3', 'hei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Capanahua', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']], 6, True)

['Capanahua', 'Language', 'codes', 'ISO', '639-3', 'kaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Makah', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Makah', 'Language', 'codes', 'ISO', '639-3', 'myh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Squamish', 'Region', 'British', 'Columbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Squamish', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Squamish', 'Pronunciation', 'sq', '\xca', '\xb7', '\xcf', '\x87', '\xca', '\xb7', 'u', '\xca', '\x94', 'm', '\xc9', '\x99', '\xca', '\x83', 'snit', '\xcd', '\xa1', '\xca', '\x83', 'im']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'sq', '\xca', '\xb7', '\xcf', '\x87', '\xca', '\xb7', 'u', '\xca', '\x94', 'm', '\xc9', '\x99', '\xca', '\x83', 'snit', '\xcd', '\xa1', '\xca', '\x83', 'im']], 23, True)

['Squamish', 'Language', 'codes', 'ISO', '639-3', 'squ']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Sechelt', 'Total', 'speakers', '40']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Sechelt', 'Region', 'British', 'Columbia', 'Ethnicity']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sechelt', 'Region', '600', 'Sechelt', 'people']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sechelt', 'Spoken', 'in', 'Canada']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Sechelt', 'Pronunciation', '\xca', '\x83', 'a', '\xca', '\x83', 'i',
'\xca', '\x83', 'a', '\xc9', '\xac']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',
'\x83', 'a', '\xca', '\x83', 'i', '\xca', '\x83', 'a', '\xc9', '\xac']]],
13, True)
```

```
['Sechelt', 'Language', 'codes', 'ISO', '639-3', 'sec']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nooksack', 'Language', 'extinction', '1988']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Nooksack', 'Spoken', 'in', 'Whatcom', 'County']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Nooksack', 'Language', 'codes', 'ISO', '639-3', 'nok']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Caquinte', 'Spoken', 'in', 'Per', '\xc3', '\xba']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',
'\xc3', '\xba']]], 6, True)
```

```
['Caquinte', 'Language', 'codes', 'ISO', '639-3', 'cot']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Xipaya', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xipaya', 'Total', 'speakers', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Xipaya', 'Language', 'codes', 'ISO', '639-3', 'xiy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Abnaki', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Abnaki', 'Total', 'speakers', 'extinct']], 5, True)

['Eastern', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Eastern', 'Region', 'Penobscot', ',', 'Maine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'aaq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Klamath', 'Total', 'speakers', '1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Klamath', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Klamath', 'Region', 'Klamath', ',', 'Modoc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Klamath', 'Language', 'codes', 'ISO', '639-3', 'kla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Estonian', 'Region', 'Limited', 'to', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',

'Type: ANY', 'Type: ANY']]], 5, True)

['Estonian', 'Signed', 'in', 'Estonian']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chiquitano', 'Language', 'codes', 'ISO', '639-3', 'cax']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Araki', 'Total', 'speakers', '8']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Araki', 'Spoken', 'in', 'Vanuatu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Araki', 'Language', 'codes', 'ISO', '639-3', 'akr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Are'are', 'Spoken', 'in', 'southern', 'Malaita', ',', 'Solomon', 'Islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'southern', 'Malaita', ',', 'Solomon', 'Islands']]], 8, True)

['Are'are', 'Writing', 'system', 'the', 'writing', 'system', '(', 's', ')', 'used', 'to', 'represent', 'the', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'the', 'writing', 'system', '\\(', 's', '\\)', 'used', 'to', 'represent', 'the', 'language']]], 14, True)

['Are'are', 'Language', 'codes', 'ISO', '639-3', 'alu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cashibo', 'Spoken', 'in', 'Per', '\\xc3', '\\xba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\\xc3', '\\xba']]], 6, True)

['Cashibo', 'Dialects', 'Cashibo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Cashibo', 'Dialects', 'Cacataibo', 'de', 'Mariscal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Cacataibo', 'de', 'Mariscal']]], 5, True)

['Cashibo', 'Dialects', 'Cacataibo', 'de', 'Sinchi', 'Roca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Cacataibo', 'de', 'Sinchi', 'Roca']]], 6, True)

['Cashibo', 'Language', 'codes', 'ISO', '639-3', 'cbr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bafut', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bafut', 'Dialects', 'Bufe', '(', 'Afughe', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 6, True)

['Bafut', 'Language', 'codes', 'ISO', '639-3', 'bfd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Region', 'Lombardy', '(', 'Italy', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Lombardy', '\\\\(', 'Italy', '\\\\)']]], 6, True)

['Banjarese', 'Spoken', 'in', 'Indonesia', ',', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Banjarese', 'Language', 'codes', 'ISO', '639-2', 'msa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Banjarese', 'Language', 'codes', 'ISO', '639-3', 'bjn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chipaya', 'Spoken', 'in', 'Bolivian', 'Andes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)



```

['Chipaya', 'Language', 'codes', 'ISO', '639-3', 'cap']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cashinahua', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',
'\xc3', '\xba', ',', 'Brazil']]], 8, True)

['Cashinahua', 'Language', 'codes', 'ISO', '639-3', 'cbs']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cantabrian', 'Total', 'speakers', 'undefined']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cantabrian', 'Spoken', 'in', 'Spain']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cantabrian', 'Writing', 'system', 'Latin', 'alphabet']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Cantabrian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Nzema', 'Spoken', 'in', 'Ghana']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nzema', 'Language', 'codes', 'ISO', '639-3', 'nzi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Early', 'Spoken', 'in', 'Japan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Early', 'Writing', 'system', 'Hiragana']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

['Early', 'Writing', 'system', 'Katakana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Early', 'Writing', 'system', 'and', 'Han']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Early', 'Language', 'codes', 'ISO', '639-3', 'ojp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Classical', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Classical', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Classical', 'Language', 'codes', 'ISO', '639-3', 'ghc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wagiman', 'Spoken', 'in', 'Pine', 'Creek']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Wagiman', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Wagiman', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wagiman', 'Language', 'codes', 'ISO', '639-3', 'waq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marovo', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Marovo', 'Language', 'codes', 'ISO', '639-3', 'mvo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Susuami', 'Spoken', 'in', 'Morobe', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Susuami', 'Language', 'codes', 'ISO', '639-3', 'ssu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cowlitz', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cowlitz', 'Region', 'Southwestern', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cowlitz', 'Language', 'codes', 'ISO', '639-3', 'cow']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kwaza', 'Spoken', 'in', 'Rond', '\xc3', '\xb4', 'nia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Rond',  
'\xc3', '\xb4', 'nia']]], 7, True)

['Kwaza', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Kwaza', 'Language', 'codes', 'ISO', '639-3', 'xwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Javindo', 'Spoken', 'in', 'Java', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Javindo', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Javindo', 'Language', 'codes', 'ISO', '639-3', 'jvd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Cahto', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cahto', 'Spoken', 'in', 'California', '(', 'Eel', 'River', ')',  
'Ethnicity']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'California', '\\(', 'Eel', 'River', '\\)', 'Ethnicity']]], 9, True)
```

```
['Cahto', 'Spoken', 'in', 'Eel', 'River', 'Athapaskans']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Eel',  
'River', 'Athapaskans']]], 6, True)
```

```
['Cahto', 'Language', 'codes', 'ISO', '639-3', 'ktw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Anejom\\xcc\\x83', 'Spoken', 'in', 'Aneityum', 'Island', ',', 'Vanuatu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Aneityum', 'Island', ',', 'Vanuatu']]], 7, True)
```

```
['Anejom\\xcc\\x83', 'Language', 'codes', 'ISO', '639-3', 'aty']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Pomo', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Southern', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Southern', 'Region', 'Northern', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'peq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Judeo-Yemeni', 'Spoken', 'in', 'Israel', ',', 'Yemen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Judeo-Yemeni', 'Writing', 'system', 'Hebrew', 'alphabet']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Judeo-Yemeni', 'Language', 'codes', 'ISO', '639-3', 'jye']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kpati', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kpati', 'Spoken', 'in', 'Taraba', 'State', ',', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Kpati', 'Language', 'codes', 'ISO', '639-3', 'koc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ambai', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ambai', 'Region', 'Ambai', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ambai', 'Language', 'codes', 'ISO', '639-3', 'amk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ugong', 'Spoken', 'in', 'Western', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ugong', 'Language', 'codes', 'ISO', '639-3', 'ugo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Phuan', 'Spoken', 'in', 'Laos', 'and', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Phuan', 'Language', 'codes', 'ISO', '639-3', 'phu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Chamicuro', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)
```

```
['Chamicuro', 'Language', 'codes', 'ISO', '639-3', 'ccc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Boon', 'Spoken', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Boon', 'Language', 'codes', 'ISO', '639-3', 'bnl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jurruru', 'Language', 'codes', 'ISO', '639-3', 'tju']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Totoli', 'Language', 'codes', 'ISO', '639-3', 'txe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aigon', 'Language', 'codes', 'ISO', '639-3', 'aix']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Miu', 'Language', 'codes', 'ISO', '639-3', 'mpo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kaulong', 'Language', 'codes', 'ISO', '639-3', 'pss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sengseng', 'Language', 'codes', 'ISO', '639-3', 'ssz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Karore', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Karore', 'Language', 'codes', 'ISO', '639-3', 'xkx']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karnai', 'Language', 'codes', 'ISO', '639-3', 'bbv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Malasanga', 'Language', 'codes', 'ISO', '639-3', 'mqz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maleu-Kilenge', 'Spoken', 'in', 'western', 'tip', 'of', 'Talasea', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'western', 'tip', 'of', 'Talasea', 'District']]], 8, True)

['Maleu-Kilenge', 'Language', 'codes', 'ISO', '639-3', 'mgl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amanay\xc3\xa9', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Amanay\xc3\xa9', 'Region', 'S', '\xc3', '\xa3', 'o', 'Domingos', 'do', 'Capim', 'Minicipality']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'S', '\xc3', '\xa3', 'o', 'Domingos', 'do', 'Capim', 'Minicipality']]], 10, True)

['Amanay\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'ama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anamb\xc3\xa9', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Anamb\xc3\xa9', 'Region', 'Par', '\xc3', '\xa1', ',', 'Cairari', 'River', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Par', '\xc3', '\xa1', ',', 'Cairari', 'River', 'Ethnicity']]], 9, True)

['Anamb\xc3\xa9', 'Region', 'Anamb', '\xc3', '\xa9', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Anamb', '\xc3', '\xa9', 'people']]], 6, True)

['Anamb\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'aan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Apiac\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Apiac\xc3\xa1', 'Region', 'northern', 'Mato', 'Grosso', ',', 'upper',  
'Rio', 'Tapajos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northern',  
'Mato', 'Grosso', ',', 'upper', 'Rio', 'Tapajos']], 9, True)

['Apiac\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'api']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hadrami', 'Spoken', 'in', 'Yemen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Hadrami', 'Spoken', 'in', 'Eritrea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Hadrami', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Hadrami', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Hadrami', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Hadrami', 'Language', 'codes', 'ISO', '639-3', 'ayh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Coeur', 'Spoken', 'in', 'United', 'States', 'of', 'America']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']], 7, True)

['Coeur', 'Region', 'northern', 'Idaho']:



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Coeur', 'Language', 'codes', 'ISO', '639-3', 'crd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arikap\xc3\xba', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Arikap\xc3\xba', 'Region', 'Rond', '\xc3', '\xb4', 'nia', ',', 'at', 'the', 'headwaters', 'of', 'the', 'Rio', 'Branco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rond', '\xc3', '\xb4', 'nia', ',', 'at', 'the', 'headwaters', 'of', 'the', 'Rio', 'Branco']]], 14, True)
```

```
['Arikap\xc3\xba', 'Language', 'codes', 'ISO', '639-3', 'ark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aru\xc3\xa1', 'Language', 'extinction', '1877']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Aru\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Aru\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'aru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arutani', 'Spoken', 'in', 'Brazil', ',', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Arutani', 'Region', 'Roraima', '(', 'Brazil', ')', ';']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Roraima', '\\(', 'Brazil', '\\)', ';']]], 7, True)
```

```
['Arutani', 'Region', 'Karum', 'River', 'area', ',', 'Bolivar', 'State', '(', 'Venezuela', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Karum', 'River', 'area', ',', 'Bolivar', 'State', '\\(', 'Venezuela', '\\)']]], 11, True)
```

```

['Arutani', 'Language', 'codes', 'ISO', '639-3', 'atx']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sap\xc3\xa9', 'Total', 'speakers', 'a', 'few']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a',
'few']]], 5, True)

['Sap\xc3\xa9', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sap\xc3\xa9', 'Region', 'Paragua', 'and', 'Karuna', 'rivers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Paragua',
'and', 'Karuna', 'rivers']]], 6, True)

['Sap\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'spc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aur\xc3\xa1', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Aur\xc3\xa1', 'Region', 'Maranh', '\xc3', '\xa3', 'o']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maranh',
'\xc3', '\xa3', 'o']]], 6, True)

['Aur\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'aux']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baure', 'Spoken', 'in', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Baure', 'Total', 'speakers', '13']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Baure', 'Language', 'codes', 'ISO', '639-3', 'brg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'Canada']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Region', 'on', 'the', 'St', '.', 'Lawrence', 'River', 'in', 'Quebec']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'on', 'the', 'St', '.', 'Lawrence', 'River', 'in', 'Type: ANY']]], 10, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'abe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hajong', 'Spoken', 'in', 'India', ',', 'Bangladesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Hajong', 'Region', 'Assam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hajong', 'Region', 'Meghalaya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hajong', 'Writing', 'system', 'Assamese', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hajong', 'Language', 'codes', 'ISO', '639-3', 'haj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Halbi', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Halbi', 'Region', 'Madhya', 'Pradesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Halbi', 'Region', 'Chhattisgarh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Halbi', 'Region', 'Andhra', 'Pradesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Halbi', 'Region', 'Orissa', 'and', 'Maharashtra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Halbi', 'Language', 'codes', 'ISO', '639-3', 'h1b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Halbi', 'Language', 'codes', 'Linguasphere', '59-AAF-tb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Harsusi', 'Spoken', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Harsusi', 'Region', 'Jiddat', 'al-Harasis', ',', 'Dhofar', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jiddat', 'al-  
Harasis', ',', 'Dhofar', 'Province']]], 7, True)

['Harsusi', 'Language', 'codes', 'ISO', '639-3', 'hss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karakhawyana', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karakhawyana', 'Region', 'near', 'Waiwai', ',', 'Amazonas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'near',  
'Waiwai', ',', 'Amazonas']]], 6, True)

['Karakhawyana', 'Language', 'codes', 'ISO', '639-3', 'xkh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karipun\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karipun\xc3\xa1', 'Region', 'Rond', '\xc3', '\xb4', 'nia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rond', '\xc3',  
'\xb4', 'nia']]], 6, True)

```
['Karipun\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'kuq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Katawixi', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Katawixi', 'Region', 'Amazonas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Katawixi', 'Language', 'codes', 'ISO', '639-3', 'xat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kanmar\xc3\xad', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kanmar\xc3\xad', 'Region', 'Amazonas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kanmar\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'knm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Katuk\xc3\xadna', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Katuk\xc3\xadna', 'Region', 'Acre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Katuk\xc3\xadna', 'Language', 'codes', 'ISO', '639-3', 'kav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kinamigin', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kinamigin', 'Region', 'MIMAROPA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY'']], 3, True)

['Kinamigin', 'Language', 'codes', 'ISO', '639-3', 'mkx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kayort', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kayort', 'Region', 'Dakuwa', 'Danga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kayort', 'Language', 'codes', 'ISO', '639-3', 'kyv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kharia', 'Thar', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kharia', 'Spoken', 'in', 'India', ',', 'also', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',  
,', 'also', 'Type: ANY']], 7, True)

['Kharia', 'Language', 'codes', 'ISO', '639-3', 'ksy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guatemalan', 'Signed', 'in', 'Guatemala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Guatemalan', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kreye', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kreye', 'Region', 'Maranh', '\xc3', '\xa3', 'o', 'and', 'Par', '\xc3',  
'\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maranh',  
'\xc3', '\xa3', 'o', 'and', 'Par', '\xc3', '\xa1']], 10, True)

```

['Kreye', 'Language', 'codes', 'ISO', '639-3', 'xre']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Leco', 'Spoken', 'in', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Leco', 'Region', 'east', 'of', 'Lake', 'Titicaca']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'east', 'of',
'Lake', 'Titicaca']]], 6, True)

['Leco', 'Language', 'codes', 'ISO', '639-3', 'lec']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mal', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mal', 'Language', 'codes', 'ISO', '639-3', 'mkb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mapoyo', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mapoyo', 'Region', 'Suapure', 'River']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mapoyo', 'Language', 'codes', 'ISO', '639-3', 'mcg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nahari', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nahari', 'Region', 'Chhattisgarh', 'and', 'Orissa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Nahari', 'Language', 'codes', 'ISO', '639-3', 'nhh']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nk\xe1\xbb\x8dr\xe1\xbb\x8d\xe1\xbb\x8d', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nk\xe1\xbb\x8dr\xe1\xbb\x8d\xe1\xbb\x8d', 'Region', 'Rivers', 'State',  
,',', 'Opobo-Nkoro', 'LGA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rivers', 'State', ',,', 'Opobo-Nkoro', 'LGA']], 7, True)

['Nk\xe1\xbb\x8dr\xe1\xbb\x8d\xe1\xbb\x8d', 'Language', 'codes', 'ISO',  
'639-3', 'nkx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pacahuara', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pacahuara', 'Region', 'Magdalena']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Pacahuara', 'Language', 'codes', 'ISO', '639-3', 'pcp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ch\xc3\xa1cobo', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ch\xc3\xa1cobo', 'Region', 'Magdalena']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ch\xc3\xa1cobo', 'Language', 'codes', 'ISO', '639-3', 'cao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['P\xc3\xa9mono', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['P\xc3\xa9mono', 'Language', 'codes', 'ISO', '639-3', 'pev']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Reyesano', 'Spoken', 'in', 'Bolivia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Reyesano', 'Region', 'Beni']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Reyesano', 'Language', 'codes', 'ISO', '639-3', 'rey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sikiana', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sikiana', 'Spoken', 'in', 'Suriname']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sikiana', 'Spoken', 'in', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sikiana', 'Language', 'codes', 'ISO', '639-3', 'sik']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yabarana', 'Spoken', 'in', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yabarana', 'Region', 'Manapiare', 'River', 'basin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Yabarana', 'Language', 'codes', 'ISO', '639-3', 'yar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Labu', 'Region', 'Morobe', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Labu', 'Language', 'codes', 'ISO', '639-3', 'lbu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Macagu\xc3\xa1n', 'Spoken', 'in', 'Colombia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Macagu\xc3\xa1n', 'Region', 'between', 'the', 'Lipa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Macagu\xc3\xa1n', 'Region', 'Ele']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Macagu\xc3\xa1n', 'Region', 'and', 'Cuiloto', 'rivers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Macagu\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'mbn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guahibo', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Guahibo', 'Region', 'Casanare', ',', 'eastern', 'Meta', ',', 'Vichada',
',', 'Guaviare', ',', 'Guainia', 'states', '(', 'Colombia', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Casanare',
',', 'eastern', 'Meta', ',', 'Vichada', ',', 'Guaviare', ',', 'Guainia',
'states', '\\(', 'Colombia', '\\)']]], 16, True)

['Guahibo', 'Region', 'Orinoco', 'River', '(', 'Venezuela', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Orinoco',
'River', '\\(', 'Venezuela', '\\)']]], 7, True)

['Guahibo', 'Language', 'codes', 'ISO', '639-3', 'guh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cuiba', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

```

['Cuiba', 'Region', 'Meta', 'Casanare', 'and', 'Capanapara', 'rivers', '(',  
'Colombia', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Meta',  
'Casanare', 'and', 'Capanapara', 'rivers', '\\\\(', 'Colombia', '\\\\)']], 10,  
True)

['Cuiba', 'Region', 'Apure', 'Division', '(' , 'Venezuela', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Apure',  
'Division', '\\\\(', 'Venezuela', '\\\\)']], 7, True)

['Cuiba', 'Language', 'codes', 'ISO', '639-3', 'cui']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guayabero', 'Spoken', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Guayabero', 'Language', 'codes', 'ISO', '639-3', 'guo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Playero', 'Spoken', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Playero', 'Region', 'Arauca', 'River']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Playero', 'Language', 'codes', 'ISO', '639-3', 'gob']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Awad', 'Region', 'Astrolabe', 'Bay', 'area', ',', 'Madang', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Astrolabe',  
'Bay', 'area', ',', 'Madang', 'Province']], 8, True)

['Awad', 'Language', 'codes', 'ISO', '639-3', 'bcu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mindiri', 'Language', 'codes', 'ISO', '639-3', 'mpn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Wab', 'Region', 'Huon', 'Peninsula']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Wab', 'Region', 'Madang', 'Province']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Wab', 'Language', 'codes', 'ISO', '639-2', 'map']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Wab', 'Language', 'codes', 'ISO', '639-3', 'wab']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Marik', 'Region', 'around', 'Gogol', 'River', ',', 'Madang', 'Province']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'around',
'Gogol', 'River', ',', 'Madang', 'Province']], 8, True)
```

```
['Marik', 'Language', 'codes', 'ISO', '639-3', 'dad']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bil', 'Region', 'south', 'of', 'Madang', 'town', ',', 'Madang',
'Province']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'south', 'of',
'Madang', 'town', ',', 'Madang', 'Province']], 9, True)
```

```
['Bil', 'Language', 'codes', 'ISO', '639-3', 'brz']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gedaged', 'Language', 'codes', 'ISO', '639-3', 'gdd']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Matukar', 'Region', 'north', 'of', 'Madang', 'town', ',', 'Madang',
'Province']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'north', 'of',
'Madang', 'town', ',', 'Madang', 'Province']], 9, True)
```

```
['Matukar', 'Language', 'codes', 'ISO', '639-3', 'mjk']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Takia', 'Region', 'Madang', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Takia', 'Language', 'codes', 'ISO', '639-3', 'tbd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["U'wa", 'Spoken', 'in', 'Boyac', '\xc3', '\xa1', 'Department']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Boyac',
'\xc3', '\xa1', 'Department']], 7, True)
```

```
["U'wa", 'Spoken', 'in', 'Colombia', ';', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Colombia', ';', 'Venezuela']], 6, True)
```

```
["U'wa", 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["U'wa", 'Language', 'codes', 'ISO', '639-3', ':', 'tnd', '\xe2', '\x80',
'\x93', 'Angosturas', 'Tunebo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'tnd', '\xe2', '\x80', '\x93', 'Angosturas', 'Tunebo']],
['NUM', '639-3']], 12, True)
```

```
["U'wa", 'Language', 'codes', 'ISO', '639-3', 'tbn', '\xe2', '\x80',
'\x93', 'Barro', 'Negro', 'Tunebo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'tbn', '\xe2', '\x80', '\x93', 'Barro', 'Negro', 'Tunebo']],
['NUM', '639-3']], 12, True)
```

```
['Amara', 'Language', 'codes', 'ISO', '639-3', 'aie']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lamogai', 'Language', 'codes', 'ISO', '639-3', 'lmg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mouk-Aria', 'Language', 'codes', 'ISO', '639-3', 'mwh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shabaki', 'Spoken', 'in', 'Iraq']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shabaki', 'Region', 'Mosul']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Shabaki', 'Language', 'codes', 'ISO', '639-3', 'sdb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Polar', 'Spoken', 'in', 'Qaanaaq', 'area']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Polar', 'Spoken', 'in', 'Northern', 'Greenland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Polar', 'Spoken', 'in', 'Denmark']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polar', 'Region', 'North', 'America']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Polar', 'Official', 'language', 'in', 'Greenland', '(', 'Denmark', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Greenland', '\\(', 'Denmark', '\\)']]], 8, True)

['Polar', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Polar', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Polar', 'Language', 'codes', 'ISO', '639-3', 'kal-pol']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bukawa', 'Language', 'codes', 'ISO', '639-3', 'buk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yabem', 'Language', 'codes', 'ISO', '639-3', 'jae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kela', 'Language', 'codes', 'ISO', '639-3', 'kcl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mangareva', 'Spoken', 'in', 'French', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Mangareva', 'Language', 'codes', 'ISO', '639-3', 'mrv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Austral', 'Spoken', 'in', 'French', 'Polynesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Austral', 'Region', 'Austral', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Austral', 'Language', 'codes', 'ISO', '639-3', 'aut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Penrhyn', 'Spoken', 'in', 'Cook', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Penrhyn', 'Language', 'codes', 'ISO', '639-3', 'pnh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rakahanga-Manihiki', 'Spoken', 'in', 'Cook', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Rakahanga-Manihiki', 'Region', 'Rakahanga', 'and', 'Manihiki', 'islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rakahanga',

'and', 'Manihiki', 'islands']]], 6, True)

['Rakahanga-Manihiki', 'Official', 'language', 'in', 'Cook', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]]], 6, True)

['Rakahanga-Manihiki', 'Regulated', 'by', 'Kopapa', 'Reo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Rakahanga-Manihiki', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Rakahanga-Manihiki', 'Language', 'codes', 'ISO', '639-3', 'rkh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Maay', 'Spoken', 'in', 'Somalia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Maay', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Maay', 'Spoken', 'in', 'and', 'Yemen']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Maay', 'Total', 'speakers', '1', ',', '5-2', 'million', 'native', 'and',  
'at', 'least', '200', ',', '000', 'second', 'language', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'million', 'native', 'and', 'at', 'least', 'NUM', ',',  
'NUM', 'second', 'language', 'speakers']]], ['NUM', '1'], ['NUM', '5-2'],  
'NUM', '200'], ['NUM', '000']], 17, True)

['Maay', 'Official', 'language', 'in', 'Somalia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]]], 5, True)

['Maay', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]]], 6, True)



```
['Maay', 'Language', 'codes', 'ISO', '639-3', 'ymm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Maay', 'Language', 'codes', 'ISO', '639-1', 'so']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Gone', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Gone', 'Language', 'codes', 'ISO', '639-3', 'goo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lauan', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Lauan', 'Region', 'Lau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lauan', 'Region', 'Nayau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lauan', 'Region', 'Lakeba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lauan', 'Region', 'Oneata']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lauan', 'Region', 'Moce']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lauan', 'Region', 'Komo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Lauan', 'Region', 'Namuka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY'[]], 3, True)

['Lauan', 'Region', 'Kabara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lauan', 'Region', 'Vulaga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lauan', 'Region', 'Ogea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lauan', 'Region', 'Vatoa', 'islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Lauan', 'Language', 'codes', 'ISO', '639-3', 'llx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY'[]], ['NUM', '639-3']], 6, True)

['Lomaiviti', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 4, True)

['Lomaiviti', 'Region', 'Koro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lomaiviti', 'Region', 'Makogai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lomaiviti', 'Region', 'Levuka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lomaiviti', 'Region', 'Ovalau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

['Lomaiviti', 'Region', 'Batiki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY'[]]], 3, True)

```
['Lomaiviti', 'Region', 'Nairai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Lomaiviti', 'Region', 'Gau', 'islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lomaiviti', 'Language', 'codes', 'ISO', '639-3', 'lmv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kachari', 'Spoken', 'in', 'Assam', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Kachari', 'Language', 'codes', 'ISO', '639-3', 'xac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Western', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Western', 'Region', 'western', 'half', 'of', 'Viti', 'Levu', ',', 'Waya',  
'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western',  
'half', 'of', 'Viti', 'Levu', ',', 'Waya', 'Islands']]], 10, True)
```

```
['Western', 'Language', 'codes', 'ISO', '639-3', 'wyy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Namosi-Naitasiri-Serua', 'Spoken', 'in', 'Fiji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Namosi-Naitasiri-Serua', 'Region', 'Namosi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Namosi-Naitasiri-Serua', 'Region', 'Serua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Namosi-Naitasiri-Serua', 'Region', 'Naitasiri', 'provinces']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Namosi-Naitasiri-Serua', 'Language', 'codes', 'ISO', '639-3', 'bwb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Vangunu', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Vangunu', 'Region', 'Vangunu', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vangunu', 'Language', 'codes', 'ISO', '639-3', 'mpr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Duke', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Duke', 'Region', 'Kolombangara', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Duke', 'Language', 'codes', 'ISO', '639-3', 'nke']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Botolan', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Botolan', 'Region', 'Zambales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Botolan', 'Region', 'Olongapo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Botolan', 'Region', 'Metro', 'Manila']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Botolan', 'Official', 'language', 'in', 'None']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Botolan', 'Language', 'codes', 'ISO', '639-3', 'xsb']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Ghanongga', 'Spoken', 'in', 'Solomon', 'Islands']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Ghanongga', 'Region', 'north', 'Ranonga', 'Island']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Ghanongga', 'Language', 'codes', 'ISO', '639-3', 'ghn']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Hoava', 'Spoken', 'in', 'Solomon', 'Islands']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Hoava', 'Language', 'codes', 'ISO', '639-3', 'hoa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Ambala', 'Spoken', 'in', 'Philippines']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ambala', 'Region', 'Zambales']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Ambala', 'Region', 'Olongapo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Ambala', 'Region', 'Dinalupihan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Ambala', 'Language', 'codes', 'ISO', '639-3', 'abc']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kusaghe', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Kusaghe', 'Language', 'codes', 'ISO', '639-3', 'ksg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mag-indi/Baloga', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mag-indi/Baloga', 'Region', 'Floridablanca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Mag-indi/Baloga', 'Region', 'Porac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Mag-indi/Baloga', 'Region', 'San', 'Marcelino']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mag-indi/Baloga', 'Language', 'codes', 'ISO', '639-3', 'blx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Roviana', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Roviana', 'Language', 'codes', 'ISO', '639-3', 'rug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Simbo', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Simbo', 'Region', 'Simbo', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Simbo', 'Language', 'codes', 'ISO', '639-3', 'sbb']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ughele', 'Spoken', 'in', 'Solomon', 'Islands']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Ughele', 'Region', 'north', 'Rendova', 'Island']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Ughele', 'Language', 'codes', 'ISO', '639-3', 'uge']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sambal', 'Spoken', 'in', 'Philippines']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sambal', 'Region', 'Zambales']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Sambal', 'Region', 'Pangasinan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Sambal', 'Region', 'Metro', 'Manila']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sambal', 'Region', 'Palawan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Sambal', 'Official', 'language', 'in', 'None']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Sambal', 'Language', 'codes', 'ISO', '639-3', 'xsb']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uneapa', 'Language', 'codes', 'ISO', '639-3', 'bbn']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minaean', 'Spoken', 'in', 'Yemen']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Minaean', 'Total', 'speakers', 'Extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Minaean', 'Language', 'codes', 'ISO', '639-3', 'inm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Qatabanian', 'Spoken', 'in', 'Yemen']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Qatabanian', 'Total', 'speakers', 'Extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Qatabanian', 'Language', 'codes', 'ISO', '639-3', 'xqt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hadhramautic', 'Spoken', 'in', 'Yemen']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hadhramautic', 'Total', 'speakers', 'Extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Hadhramautic', 'Language', 'codes', 'ISO', '639-3', 'xhd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mag-antsi', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mag-antsi', 'Region', 'Zambales']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)



['Mag-antsi', 'Region', 'Tarlac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mag-antsi', 'Region', 'Mabalacat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mag-antsi', 'Region', 'Angeles']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mag-antsi', 'Language', 'codes', 'ISO', '639-3', 'sgb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abellen', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abellen', 'Region', 'Tarlac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abellen', 'Language', 'codes', 'ISO', '639-3', 'abp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hausa', 'Signed', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hausa', 'Total', 'signers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vivaro-Alpine', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Vivaro-Alpine', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vivaro-Alpine', 'Region', 'Southern', 'France']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Vivaro-Alpine', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Marivele\xc3\xb1o', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Marivele\xc3\xb1o', 'Region', 'Mariveles']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Marivele\xc3\xb1o', 'Language', 'codes', 'ISO', '639-3', 'ayt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aribwatsa', 'Spoken', 'in', 'New', 'Guinea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Aribwatsa', 'Total', 'speakers', '0']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aribwatsa', 'Language', 'codes', 'ISO', '639-3', 'laz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Duwet', 'Spoken', 'in', 'New', 'Guinea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Duwet', 'Total', 'speakers', '400']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Duwet', 'Language', 'codes', 'ISO', '639-3', 'gve']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guatuso', 'Spoken', 'in', 'Costa', 'Rica']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Guatuso', 'Language', 'codes', 'ISO', '639-3', 'gut']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Adabe', 'Language', 'codes', 'ISO', '639-3', 'adb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jewish', 'Language', 'codes', 'ISO', '639-3', 'tmr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['sign', 'Signed', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['sign', 'Total', 'signers', '3000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bali', 'Spoken', 'in', 'Demsu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bali', 'Spoken', 'in', 'Adamawa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bali', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bali', 'Language', 'codes', 'ISO', '639-3', 'bcn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gojri/Gujari', 'Spoken', 'in', 'Pakistan', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Gojri/Gujari', 'Language', 'codes', 'ISO', '639-2', 'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Gojri/Gujari', 'Language', 'codes', 'ISO', '639-3', 'gju']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Spoken', 'in', 'Italy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Southern', 'Region', 'Abruzzo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Apulia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Basilicata']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Calabria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Campania']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Lazio']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Marche']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Region', 'Molise']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'Dialects']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chayahuita', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]]], 6, True)

['Chayahuita', 'Language', 'codes', 'ISO', '639-3', 'cbt']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cocama', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]]], 6, True)

['Cocama', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Cocama', 'Spoken', 'in', 'Colombia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Cocama', 'Language', 'codes', 'ISO', '639-3', 'cod']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mursi', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Mursi', 'Region', 'Central', 'Omo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Mursi', 'Writing', 'system', 'Ethiopic']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Mursi', 'Language', 'codes', 'ISO', '639-3', 'muz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rama', 'Spoken', 'in', 'Nicaragua']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

```
['Rama', 'Region', 'Rama', 'Cay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Rama', 'Language', 'codes', 'ISO', '639-3', 'rma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Klao', 'Spoken', 'in', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Klao', 'Language', 'codes', 'ISO', '639-3', 'klu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Iteri', 'Region', 'Sandaun', 'Province', ',', 'Rocky', 'Peak',  
'Mountains', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sandaun',  
'Province', ',', 'Rocky', 'Peak', 'Mountains', '.', 'Coordinates']]], 10,  
True)
```

```
['Iteri', 'Language', 'codes', 'ISO', '639-3', 'itr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mantion', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mantion', 'Language', 'codes', 'ISO', '639-3', 'mnx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arikara', 'Total', 'speakers', '3']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arikara', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Arikara', 'Region', 'North-central', 'Oklahoma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Samoan', 'Spoken', 'in', 'Samoa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Samoan', 'Language', 'codes', 'ISO', '639-2', 'cpe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Samoan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Alutiiq', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Alutiiq', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Alutiiq', 'Language', 'codes', 'ISO', '639-3', 'ems']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Central', 'Region', 'western', 'and', 'southwestern', 'Alaska']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'western',  
'and', 'southwestern', 'Alaska']]], 6, True)

['Central', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Central', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'esu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bribri', 'Spoken', 'in', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Bribri', 'Language', 'codes', 'ISO', '639-3', 'bzd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wom', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Wom', 'Language', 'codes', 'ISO', '639-3', 'wmo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Patamona', 'Spoken', 'in', 'Guyana', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Patamona', 'Spoken', 'in', 'Patamona', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Patamona', 'Total', 'speakers', '5000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Patamona', 'Language', 'codes', 'ISO', '639-3', 'pbc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ansus', 'Total', 'speakers', '4600']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ansus', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ansus', 'Region', 'Miosnum', 'Island', 'and', 'the', 'south', 'coast',  
'of', 'Serui', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Miosnum',  
'Island', 'and', 'the', 'south', 'coast', 'of', 'Serui', 'Island']]], 11,  
True)

['Ansus', 'Language', 'codes', 'ISO', '639-3', 'and']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'and']], ['NUM', '639-3']], 6, True)

['Sudanese', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',



'\xe2', '\x80', '\x93']]], 6, True)

['Sudanese', 'Writing', 'system', 'English', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sudanese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Apala\xc3\xad', 'Spoken', 'in', 'Brazil', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Apala\xc3\xad', 'Spoken', 'in', 'Aparai', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Apala\xc3\xad', 'Total', 'speakers', '450']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Apala\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'apy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Pronunciation', '\xcb', '\x88', 'j', '\xc9', '\x9b', '\xcc',  
'\x83', 'z', '\xc9', '\xa8', 'k', '\xcb', '\x88', 'p', '\xc9', '\x94',  
'lski']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'j', '\xc9', '\x9b', '\xcc', '\x83', 'z', '\xc9', '\xa8', 'k',  
'\xcb', '\x88', 'p', '\xc9', '\x94', 'lski']]], 19, True)

['Old', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Old', 'Language', 'codes', 'ISO', '639-3', '0gi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tetelcingo', 'Spoken', 'in', 'Tetelcingo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tetelcingo', 'Spoken', 'in', 'Morelos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetelcingo', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetelcingo', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Tetelcingo', 'Pronunciation', 'm', '\xc9', '\x94', '\xe1', '\xb5', '\x83', 's', '\xe2', '\x81', '\xb1', 'e', '\xcb', '\x88', 'wal', '\xc9', '\xaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm', '\xc9', '\x94', '\xe1', '\xb5', '\x83', 's', '\xe2', '\x81', '\xb1', 'e', '\xcb', '\x88', 'wal', '\xc9', '\xaa']], 18, True)

['Tetelcingo', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tetelcingo', 'Language', 'codes', 'ISO', '639-3', 'nhg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Makhuwa-Shirima', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Makhuwa-Shirima', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Makhuwa-Shirima', 'Language', 'codes', 'ISO', '639-3', 'vmk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laiuse', 'Spoken', 'in', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Laiuse', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Laiuse', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Misantla', 'Spoken', 'in', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Misantla', 'Region', 'Veracruz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Misantla', 'Language', 'codes', 'ISO', '639-3', 'tlc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rutul', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Rutul', 'Official', 'language', 'in', 'Dagestan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Rutul', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Rutul', 'Language', 'codes', 'ISO', '639-3', 'rut']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wala', 'Spoken', 'in', 'Solomon', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Wala', 'Total', 'speakers', '7000']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wala', 'Language', 'codes', 'ISO', '639-3', 'lgl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Dagaare', 'Spoken', 'in', 'Ghana', ',', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ghana',  
,',', 'Burkina', 'Faso']]], 7, True)
```

```
['Dagaare', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dagaare', 'Language', 'codes', 'ISO', '639-3', 'dgd', '\xe2', '\x80',  
'\x93', 'Dagaari', 'Dioula']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'dgd', '\xe2', '\x80', '\x93', 'Dagaari', 'Type: ANY']],  
['NUM', '639-3']], 11, True)
```

```
['Pidgin', 'Spoken', 'in', 'The', 'Gambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Pidgin', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pidgin', 'Language', 'codes', 'ISO', '639-2', 'crp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Pidgin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Kadazan,', 'Total', 'speakers', 'second', 'language', ':']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'second', 'language', ':']]], 6, True)
```

```
['Kadazan,', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kadazan,', 'Region', 'Sabah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Kadazan,', 'Language', 'codes', 'ISO', '639-3', 'kzj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Western', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Western', 'Language', 'codes', 'ISO', '639-3', 'ssl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
["Ts'\xc3\xbcn-Lao", 'Spoken', 'in', 'Lai', 'Chau', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lai',  
'Chau', 'Province']]], 6, True)
```

```
["Ts'\xc3\xbcn-Lao", 'Language', 'codes', 'ISO', '639-3', 'tsl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sawi', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Sawi', 'Language', 'codes', 'ISO', '639-3', 'saw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tati', 'Total', 'speakers', 'ca', '.', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'ca', '.', '?']]], 6, True)
```

```
['Tati', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tati', 'Region', 'Northwest', 'of', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Tati', 'Language', 'codes', 'ISO', '639-3', 'tks']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Petats', 'Region', 'Buka', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Petats', 'Language', 'codes', 'ISO', '639-3', 'pex']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mojave', 'Spoken', 'in', 'California', 'and', 'Arizona', ',', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'California', 'and', 'Arizona', ',', 'USA']]]], 8, True)

['Mojave', 'Language', 'codes', 'ISO', '639-3', 'mov']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cocopah', 'Spoken', 'in', 'Mexico', ',', 'USA']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)

['Cocopah', 'Region', 'Cocopah', 'people']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Cocopah', 'Region', 'Baja', 'California']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Cocopah', 'Region', 'Arizona']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['Cocopah', 'Region', 'Sonora', 'Ethnicity']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Cocopah', 'Language', 'codes', 'ISO', '639-3', 'coc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jiao-Liao', 'Spoken', 'in', 'Shandong', 'and', 'Liaodong']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]]], 6, True)

['Jiao-Liao', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]]], 6, True)

['Jiao-Liao', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',

'639-3']], 10, True)

['Jiao-Liao', 'Language', 'codes', 'ISO', '639-3', 'cmn-jia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yinjibarndi', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yinjibarndi', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Yinjibarndi', 'Language', 'codes', 'ISO', '639-3', 'yij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Colorado', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Colorado', 'Region', 'Nevada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Colorado', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Colorado', 'Region', 'Utah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Colorado', 'Region', 'Arizona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Colorado', 'Region', 'Colorado']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Colorado', 'Language', 'codes', 'ISO', '639-3', 'ute']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Yerukala', 'Spoken', 'in', 'South', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Yerukala', 'Language', 'codes', 'ISO', '639-3', 'yeu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Batangas', 'Tagalog', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', '\xe2', '\x80', '\x93']]], 7, True)
```

```
['Batangas', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Batangas', 'Region', 'Batangas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Batangas', 'Writing', 'system', 'Historically', 'written', 'in',  
'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Historically', 'written', 'in', 'Type: ANY']]], 7, True)
```

```
['Batangas', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)
```

```
['Batangas', 'Language', 'codes', 'ISO', '639-3', 'tgl-bag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Shirvani', 'Spoken', 'in', 'Azerbaijan', ',', 'Dagestan', '(', 'Russia',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Azerbaijan', ',', 'Dagestan', '\\(', 'Russia', '\\)']]], 9, True)
```

```
['Shirvani', 'Region', 'Caucasus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Shirvani', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Shirvani', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Central', 'Spoken', 'in', 'Uzbekistan', 'and', 'Tajikistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Central', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Central', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rama', 'Spoken', 'in', 'Nicaragua']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Rama', 'Region', 'Rama', 'Cay']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['North', 'Spoken', 'in', 'Iraq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['North', 'Spoken', 'in', 'Syria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['North', 'Spoken', 'in', 'Turkey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['North', 'Writing', 'system', 'Arabic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

```

['North', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'ayp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tajiki', 'Spoken', 'in', 'Afghanistan', ',', 'Tajikistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)

['Tajiki', 'Language', 'codes', 'ISO', '639-3', 'abh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uzbeki', 'Spoken', 'in', 'Uzbekistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Uzbeki', 'Region', 'Bukhara', 'province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Uzbeki', 'Language', 'codes', 'ISO', '639-3', 'auz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Judeo-Iraqi', 'Spoken', 'in', 'Israel', ',', 'Iraq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)

['Judeo-Iraqi', 'Writing', 'system', 'Hebrew', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

['Judeo-Iraqi', 'Language', 'codes', 'ISO', '639-3', 'yhd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Qoqmon\xc4\x8daq', 'Total', 'speakers', '200']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Qoqmon\xc4\x8daq', 'Spoken', 'in', 'Xinjiang-Uyghur', 'Autonomous']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

```

ANY', 'Type: ANY']]], 5, True)

['Qoqmon\xc4\x8daq', 'Region', 'People's', 'Republic', 'of', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Qoqmon\xc4\x8daq', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aringa', 'Spoken', 'in', 'Northwest', 'corner', 'of', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Northwest', 'corner', 'of', 'Uganda']]], 7, True)

['Aringa', 'Spoken', 'in', 'north', 'of', 'Lake', 'Albert']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'north',  
'of', 'Lake', 'Albert']]], 7, True)

['Aringa', 'Spoken', 'in', 'Aringa', 'county']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Aringa', 'Spoken', 'in', 'north', 'of', 'Lugbara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'north',  
'of', 'Type: ANY']]], 6, True)

['Aringa', 'Language', 'codes', 'ISO', '639-3', 'luc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Twana', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Twana', 'Language', 'codes', 'ISO', '639-3', 'twa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Labrador', 'Spoken', 'in', 'Straits', 'of', 'Belle', 'Isle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Straits', 'of', 'Belle', 'Isle']]], 7, True)

['Labrador', 'Language', 'codes', 'ISO', '639-2', 'cpf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Labrador', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Qabiao', 'Spoken', 'in', 'H', '\xc3', '\xa0', 'Giang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'H',
'\xc3', '\xa0', 'Giang']]], 7, True)
```

```
['Qabiao', 'Spoken', 'in', 'Vietnam', ';', 'Wenshan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Vietnam', ';', 'Wenshan']]], 6, True)
```

```
['Qabiao', 'Spoken', 'in', 'Yunnan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Qabiao', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Qabiao', 'Language', 'codes', 'ISO', '639-3', 'laq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Amoy', 'Total', 'speakers', 'over', '10', 'million', '(', 'no', 'recent',
'data', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'over', 'NUM', 'million', '\\(', 'no', 'recent', 'data', '\\)']], ['NUM',
'10']], 11, True)
```

```
['Amoy', 'Spoken', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Amoy', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Amoy', 'Spoken', 'in', 'Singapore']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Amoy', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Amoy', 'Region', 'Southern', 'Fujian', 'province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Amoy', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Amoy', 'Language', 'codes', 'ISO', '639-3', 'nan-amo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wogamusin', 'Spoken', 'in', 'Ambunti', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)
```

```
['Wogamusin', 'Language', 'codes', 'ISO', '639-3', 'wog']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Selayar', 'Spoken', 'in', 'Selayar', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)
```

```
['Selayar', 'Spoken', 'in', 'South', 'Sulawesi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)
```

```
['Selayar', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Selayar', 'Language', 'codes', 'ISO', '639-3', 'sly']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Xakriab\x3\xa1', 'Language', 'extinction', 'Extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Xakriab\x3\xa1', 'Spoken', 'in', 'Minas', 'Gerais', ',', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Minas',  
'Gerais', ',', 'Brazil']]]], 7, True)
```

```

['Xakriab\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'xkr']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sangirese', 'Spoken', 'in', 'Indonesia', ',', 'Philippines']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Sangirese', 'Writing', 'system', 'Latin']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sangirese', 'Language', 'codes', 'ISO', '639-3', 'sxn']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Queensland', 'Spoken', 'in', 'Queensland', ',', 'Australia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Queensland', 'Total', 'speakers']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)

['Queensland', 'Writing', 'system', 'Latin', 'alphabet']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Queensland', 'Language', 'codes', 'ISO', '639-2', 'cpe']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Queensland', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Kangean', 'Spoken', 'in', 'Indonesia']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kangean', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
(C[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Kangean', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

```

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Kangean', 'Language', 'codes', 'ISO', '639-3', 'kkv']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bengkulu', 'Spoken', 'in', 'Bengkulu', 'Province']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Bengkulu', 'Language', 'codes', 'ISO', '639-2', 'msa']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bengkulu', 'Language', 'codes', 'ISO', '639-3', 'pse']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huitoto', 'Spoken', 'in', 'Colombia', 'and', 'Peru']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']], 6, True)

['Huitoto', 'Total', 'speakers', '6', ',', '245', 'in', 'Columbia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '6'], ['NUM', '245']], 8, True)

['Huitoto', 'Total', 'speakers', '1', ',', '133', '\xe2', '\x80', '\x93', '1', ',', '917', 'in', 'Peru']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\xe2', '\x80', '\x93', 'NUM', ',', 'NUM', 'in', 'Peru']], ['NUM', '1'], ['NUM', '133'], ['NUM', '1'], ['NUM', '917']], 14, True)

['Huitoto', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Huitoto', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huitoto', 'Language', 'codes', 'ISO', '639-3', ':', 'hto', '\xe2', '\x80', '\x93', 'Minica']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', ':', 'hto', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Huitoto', 'Language', 'codes', 'ISO', '639-3', 'huu', '\xe2', '\x80', '\x93', 'Murui']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'huu', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Maroon', 'Spoken', 'in', 'Jamaica']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maroon', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Maroon', 'Language', 'codes', 'ISO', '639-2', 'cpe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Maroon', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Abom', 'Language', 'codes', 'ISO', '639-3', 'aob']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cisalpine', 'Spoken', 'in', 'Cisalpine', 'Gaul']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Cisalpine', 'Language', 'codes', 'ISO', '639-3', 'xcg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Sa'a", 'Language', 'codes', 'ISO', '639-3', 'apb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Dori'o", 'Language', 'codes', 'ISO', '639-3', 'dor']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nigerian', 'Spoken', 'in', 'northern', 'Nigeria']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nigerian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Nigerian', 'Writing', 'system', 'Arabic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nigerian', 'Official', 'language', 'in', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Nigerian', 'Regulated', 'by', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nigerian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Eastern', 'Spoken', 'in', 'Hidalgo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Spoken', 'in', 'northern', 'Puebla', 'and', 'northern', 'Veracruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'northern', 'Puebla', 'and', 'northern', 'Veracruz']]], 8, True)

['Eastern', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Eastern', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'nhe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balochi', 'Spoken', 'in', 'Pakistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balochi', 'Spoken', 'in', 'Iran']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balochi', 'Spoken', 'in', 'Afghanistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balochi', 'Spoken', 'in', 'Turkmenistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balochi', 'Spoken', 'in', 'UAE']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Balochi', 'Spoken', 'in', 'and', 'Oman']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Balochi', 'Dialects']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']]], 2, True)

['Balochi', 'Official', 'language', 'in', 'Balochistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Balochi', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Balochi', 'Language', 'codes', 'ISO', '639-2', 'bal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Balochi', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Balochi', 'Language', 'codes', 'ISO', '639-3', ':', 'bal', '\xe2',

'\x80', '\x93', 'Baluchi', '(', 'generic', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'bal', '\xe2', '\x80', '\x93', 'Baluchi', '\\(',  
'generic', '\\)']], ['NUM', '639-3']], 14, True)

['Lola', 'Spoken', 'in', 'Aru', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Lola', 'Dialects', 'Warabal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Lola', 'Language', 'codes', 'ISO', '639-3', 'lcd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Temuan', 'Region', 'central', 'and', 'southern', 'Peninsular',  
'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central',  
'and', 'southern', 'Peninsular', 'Malaysia']]], 7, True)

['Temuan', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Temuan', 'Language', 'codes', 'ISO', '639-2', 'msa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Temuan', 'Language', 'codes', 'ISO', '639-3', 'tmw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sri', 'Spoken', 'in', 'Sri', 'Lanka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Sri', 'Spoken', 'in', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Sri', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sri', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sri', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Sri', 'Language', 'codes', 'ISO', '639-3', 'sci']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kunjen', 'Dialects', 'Uw', 'Oykangand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kunjen', 'Dialects', 'Uw', 'Olkola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kunjen', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kunjen', 'Language', 'codes', 'ISO', '639-3', 'kjn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piro', 'Pronunciation', '\xcb', '\x88', 'jine']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'jine']]], 5, True)

['Piro', 'Spoken', 'in', 'Brazil', ',', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Piro', 'Language', 'codes', 'ISO', '639-3', 'pib']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wasco-Wishram', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Wasco-Wishram', 'Region', 'Columbia', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Wasco-Wishram', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Wasco-Wishram', 'Language', 'codes', 'ISO', '639-3', 'wac']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lamba', 'Spoken', 'in', 'Zambia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lamba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Lamba', 'Region', 'Copperbelt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Lamba', 'Language', 'codes', 'ISO', '639-3', 'lam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Isoko', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Isoko', 'Language', 'codes', 'ISO', '639-3', 'iso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zulgo-Gemzek', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zulgo-Gemzek', 'Region', 'Far', 'North', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Zulgo-Gemzek', 'Language', 'codes', 'ISO', '639-3', 'gnd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Ebughu', 'Spoken', 'in', 'Akwa', 'Ibom', 'State', '(', 'Nigeria', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Akwa',
'Ibom', 'State', '\\(', 'Nigeria', '\\)']], 9, True)
```

```
['Ebughu', 'Total', 'speakers', '5000', '(', 'Ethnologue', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '5000']], 7, True)
```

```
['Ebughu', 'Language', 'codes', 'ISO', '639-3', 'ebg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['E', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['E', 'Language', 'codes', 'ISO', '639-3', 'eee']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Canadian', 'Total', 'speakers', '(', 'mother', 'tongue', ')', '7',
'million', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\
(', 'mother', 'tongue', '\\)', 'NUM', 'million', 'in', 'Canada']], ['NUM',
'7']], 11, True)
```

```
['Canadian', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']], 6, True)
```

```
['Canadian', 'Language', 'codes', 'ISO', '639-2', 'fre', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'fre', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Canadian', 'Language', 'codes', 'ISO', '639-2', 'fra', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'fra', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Canadian', 'Language', 'codes', 'ISO', '639-3', 'fra']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Canadian', 'Language', 'codes', 'ISO', '639-1', 'fr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

['Iaai', 'Spoken', 'in', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2', '\x80', '\x93']], 6, True)

['Iaai', 'Region', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Iaai', 'Language', 'codes', 'ISO', '639-3', 'iai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Huachipaeri', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']], 6, True)

['Huachipaeri', 'Language', 'codes', 'ISO', '639-3', 'hug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maritime', 'Signed', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Amal', 'Spoken', 'in', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Amal', 'Language', 'codes', 'ISO', '639-3', 'aad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ari', 'Region', 'Aramia', 'River', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Ari', 'Language', 'codes', 'ISO', '639-3', 'aac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gogodala', 'Region', 'Western', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gogodala', 'Region', 'north', 'bank', 'of', 'Fly', 'River']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'north', 'bank', 'of', 'Fly', 'River']]], 7, True)

['Gogodala', 'Region', 'Aramia', 'River', '.', '301', 'villages', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Aramia', 'River', '.', 'NUM', 'villages', '.', 'Coordinates']]], ['NUM', '301']], 9, True)

['Gogodala', 'Language', 'codes', 'ISO', '639-3', 'ggw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Minica', 'Spoken', 'in', 'Colombia', ',', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Colombia', ',', 'Per', '\xc3', '\xba']]], 8, True)

['Minica', 'Language', 'codes', 'ISO', '639-3', 'hto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Murui', 'Spoken', 'in', 'Colombia', ',', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Colombia', ',', 'Per', '\xc3', '\xba']]], 8, True)

['Murui', 'Language', 'codes', 'ISO', '639-3', 'huu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['N\xc3\xbcpcode', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['N\xc3\xbcpcode', 'Total', 'speakers', '100']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['N\xc3\xbcpcode', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['N\xc3\xbcpcode', 'Language', 'codes', 'ISO', '639-3', 'hux']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['I\xc3\xblapari', 'Spoken', 'in', 'Peru']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['I\\xc3\\xb1apari', 'Official', 'language', 'in', 'Per', '\\xc3', '\\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Per', '\\xc3', '\\xba']]], 7, True)

['I\\xc3\\xb1apari', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['I\\xc3\\xb1apari', 'Language', 'codes', 'ISO', '639-3', 'inp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chantyal', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chantyal', 'Region', 'Baglung', 'and', 'Myagdi', 'Districts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Baglung', 'and', 'Myagdi', 'Districts']]], 6, True)

['Chantyal', 'Language', 'codes', 'ISO', '639-3', 'chx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Borana', 'Spoken', 'in', 'Kenya', ',', 'Ethipia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Borana', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Borana', 'Language', 'codes', 'ISO', '639-3', 'gax']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Orma', 'Spoken', 'in', 'Tana', 'River', 'District', ',', 'Garissa', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Tana', 'River', 'District', ',', 'Garissa', 'District']]], 9, True)

['Orma', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'\xe2', '\x80', '\x93']]], 6, True)

['Orma', 'Language', 'codes', 'ISO', '639-2', 'orm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Orma', 'Language', 'codes', 'ISO', '639-3', 'orc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Orma', 'Language', 'codes', 'ISO', '639-1', 'om']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Pisabo', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pisabo', 'Language', 'codes', 'ISO', '639-3', 'pig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Teribe', 'Total', 'speakers', '3000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Teribe', 'Spoken', 'in', 'Panama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Teribe', 'Spoken', 'in', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Teribe', 'Region', 'In', 'Panama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Teribe', 'Region', ':', 'Bocas', 'del', 'Toro', 'Province', ',', 'Chiriqu', '\xc3', '\xad', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', ':', 'Bocas', 'del', 'Toro', 'Province', ',', 'Chiriqu', '\xc3', '\xad', 'Province']], 12, True)

['Teribe', 'Region', 'In', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',

'Type: ANY', 'Type: ANY']]], 5, True)

['Teribe', 'Region', ':', 'Lim', '\xc3', '\xb3', 'n', 'Province', ',', 'Puntarenas', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', ':', 'Lim', '\xc3', '\xb3', 'n', 'Province', ',', 'Puntarenas', 'Province']]], 11, True)

['Teribe', 'Language', 'codes', 'ISO', '639-3', 'tfr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chadian', 'Spoken', 'in', 'Chad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadian', 'Spoken', 'in', 'Cameroon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadian', 'Spoken', 'in', 'Niger']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadian', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadian', 'Spoken', 'in', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadian', 'Writing', 'system', 'Arabic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Chadian', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chadian', 'Language', 'codes', 'ISO', '639-3', 'shu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Talossan', 'Date', 'founded', '1980']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Talossan', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Talossan', 'Category', '(', 'purpose', ')', 'Talossan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']]], 6, True)

['Talossan', 'Created', 'by', 'R', '.', 'Ben', 'Madison']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'R',  
'.', 'Ben', 'Madison']]], 7, True)

['Talossan', 'Category', '(', 'sources', ')', 'a', 'posteriori',  
'language', '(', 'Romance', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'a', 'posteriori', 'language', '\\(', 'Romance', '\\  
)']]], 11, True)

['Talossan', 'Setting', 'and', 'usage', 'Talossa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Type: ANY']]], 5, True)

['Talossan', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Talossan', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Ika', 'Spoken', 'in', 'Delta', 'State', ',', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Ika', 'Language', 'codes', 'ISO', '639-3', 'ikk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Syria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['North', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

'\xe2', '\x80', '\x93']]], 6, True)

['North', 'Writing', 'system', 'Arabic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['North', 'Language', 'codes', 'ISO', '639-2', 'ara']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'acp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gunya', 'Spoken', 'in', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gunya', 'Total', 'speakers', '0']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gunya', 'Language', 'codes', 'ISO', '639-3', 'gyy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kullili', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Kullili', 'Language', 'codes', 'ISO', '639-3', 'zmu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lardil', 'Dialects', 'Damin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Lardil', 'Spoken', 'in', 'Bentinck', 'Island']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Lardil', 'Spoken', 'in', 'north', 'west', 'Mornington', 'Island']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'north',  
'west', 'Mornington', 'Island']]], 7, True)

['Lardil', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lardil', 'Pronunciation', 'le', '\xcb', '\x90', '\xc9', '\x96', '\xc9',  
'\xaa', 'l']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'le',  
'\xcb', '\x90', '\xc9', '\x96', '\xc9', '\xaa', 'l']]], 10, True)

['Lardil', 'Language', 'codes', 'ISO', '639-3', 'lbz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lardil', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of',  
'Wellesley', 'Islands', ',', 'the', 'area', 'traditionally', 'associated',  
'with', 'Lardil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Location', 'of', 'Wellesley', 'Islands', ',', 'the', 'area',  
'traditionally', 'associated', 'with', 'Type: ANY']], ['NUM', '639-3']],  
16, True)

['Ankave', 'Spoken', 'in', 'Kerema', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ankave', 'Spoken', 'in', 'Gulf', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ankave', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ankave', 'Language', 'codes', 'ISO', '639-3', 'aak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kayardild', 'Spoken', 'in', 'Bentinck', 'and', 'Sweers', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Bentinck', 'and', 'Sweers', 'Islands']]], 7, True)

['Kayardild', 'Spoken', 'in', 'north', 'west', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'north',  
'west', 'Type: ANY']]], 6, True)

['Kayardild', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kayardild', 'Language', 'codes', 'ISO', '639-3', 'gyd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Dalabon', 'Spoken', 'in', 'Arnhem', 'Land', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', ',', 'Australia']]], 7, True)

['Dalabon', 'Total', 'speakers', '20']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dalabon', 'Language', 'codes', 'ISO', '639-3', 'ngk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['lingua', 'Spoken', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['lingua', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['lingua', 'Spoken', 'in', 'Cyprus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['lingua', 'Region', 'Mediterranean', 'Basin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['lingua', 'Language', 'codes', 'ISO', '639-3', 'pml']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['lingua', 'Language', 'codes', 'Linguasphere', '51-AAB-c']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Chintang', 'Pronunciation', '\xcb', '\x88', 'ts', '\xca', '\xb0', 'inta', '\xcb', '\x90', '\xc5', '\x8b']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',

'\x88', 'ts', '\xca', '\xb0', 'inta', '\xcb', '\x90', '\xc5', '\x8b']]],  
12, True)

['Chintang', 'Spoken', 'in', 'Dhankuta', 'District', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Dhankuta', 'District', ',', 'Type: ANY']]], 7, True)

['Chintang', 'Language', 'codes', 'ISO', '639-3', 'ctn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anglo-Norman', 'Spoken', 'in', 'England', 'and', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', 'Wales']]], 6, True)

['Anglo-Norman', 'Spoken', 'in', 'Scotand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Anglo-Norman', 'Spoken', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Anglo-Norman', 'Language', 'codes', 'ISO', '639-3', 'xno']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Central', 'Spoken', 'in', 'Yellow', 'River', 'Plain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Yellow',  
'River', 'Plain']]], 6, True)

['Central', 'Total', 'speakers', '170', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '170']], 5, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'cmn-zho', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cmn-zho', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,  
True)

['Athpare', 'Language', 'codes', 'ISO', '639-3', 'aph']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dumi', 'Spoken', 'in', 'Khotang', 'district', ',', 'Nepal']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Khotang', 'district', ',', 'Type: ANY']]], 7, True)

['Dumi', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Dumi', 'Language', 'codes', 'ISO', '639-3', 'dus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khaling', 'Spoken', 'in', 'Khotang', 'district', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Khotang', 'district', ',', 'Type: ANY']]], 7, True)

['Khaling', 'Language', 'codes', 'ISO', '639-3', 'klr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ch'olti", 'Spoken', 'in', 'Guatemala', ',', 'Manche', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Guatemala', ',', 'Manche', 'region']]], 7, True)

["Ch'olti", 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Ch'olti", 'Language', 'codes', 'ISO', '639-3', 'qjt', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qjt', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Tsoly\xc3\xa1ni', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Tsoly\xc3\xa1ni', 'Created', 'by', 'MAR', 'Barker']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tsoly\xc3\xa1ni', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Tsoly\xc3\xa1ni', 'Category', '(', 'purpose', ')', 'Kh', '\xc3', '\xad',

```
'shan', '(', 'fictional', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Kh', '\\xc3', '\\xad', 'shan', '\\(', 'fictional', '\\
\\)']], 12, True)
```

```
['Tsoly\\xc3\\xa1ni', 'Category', '(', 'purpose', ')', 'Tsoly', '\\xc3',
'\\xa1', 'ni']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Tsoly', '\\xc3', '\\xa1', 'ni']], 9, True)
```

```
['Tsoly\\xc3\\xa1ni', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',
'\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Gorontalo', 'Region', 'Gorontalo', ',', 'Sulawesi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Gorontalo', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Gorontalo', 'Writing', 'system', 'Latin', 'Alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Gorontalo', 'Language', 'codes', 'ISO', '639-2', 'gor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Gorontalo', 'Language', 'codes', 'ISO', '639-3', 'gor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mainfr\\xc3\\xa4nkisch', 'Total', 'speakers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Mainfr\\xc3\\xa4nkisch', 'Spoken', 'in', 'Germany']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Mainfr\\xc3\\xa4nkisch', 'Region', 'Upper', 'Franconia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

['Mainfr\xc3\xa4nkisch', 'Region', 'Lower', 'Franconia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mainfr\xc3\xa4nkisch', 'Region', 'Middle', 'Franconia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mainfr\xc3\xa4nkisch', 'Region', 'northeast', 'Baden-W', '\xc3', '\xbc',  
'rttemberg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northeast',  
'Baden-W', '\xc3', '\xbc', 'rttemberg']]], 7, True)

['Mainfr\xc3\xa4nkisch', 'Region', 'southwest', 'Thuringia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mainfr\xc3\xa4nkisch', 'Language', 'codes', 'ISO', '639-3', 'vmf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guaym\xc3\xad', 'Spoken', 'in', 'Panam', '\xc3', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Panam',  
'\xc3', '\xa1']]], 6, True)

['Guaym\xc3\xad', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Guaym\xc3\xad', 'Pronunciation', 'Waim', '\xc3', '\xad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'Waim',  
'\xc3', '\xad']]], 5, True)

['Guaym\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'gym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jingpho', 'Region', 'Kachin', 'State', ',', 'Yingjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kachin',  
'State', ',', 'Yingjiang']]], 6, True)

['Jingpho', 'Language', 'codes', 'ISO', '639-2', 'kac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

```
['Jingpho', 'Language', 'codes', 'ISO', '639-3', 'kac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Luba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'Congo']]], 7, True)
```

```
['Luba', 'Region', 'Katanga', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Luba', 'Language', 'codes', 'ISO', '639-2', 'lub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Luba', 'Language', 'codes', 'ISO', '639-3', 'lub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Luba', 'Language', 'codes', 'ISO', '639-1', 'lu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Central', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Central', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Central', 'Language', 'codes', 'ISO', '639-3', 'nch']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mandar', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mandar', 'Region', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Mandar', 'Language', 'codes', 'ISO', '639-2', 'mdr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Mandar', 'Language', 'codes', 'ISO', '639-3', 'mdr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Highland', 'Spoken', 'in', 'M', '\xc3', '\xa9', 'xico', '(', 'Puebla',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'M',  
'\xc3', '\xa9', 'xico', '\\(', 'Puebla', '\\)']], 10, True)

['Highland', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Highland', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Highland', 'Language', 'codes', 'ISO', '639-3', 'azz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Orizaba', 'Spoken', 'in', 'Veracruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Orizaba', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Orizaba', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Orizaba', 'Language', 'codes', 'ISO', '639-3', 'nlv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bangi-me', 'Spoken', 'in', 'Dogon', 'cliffs', ',', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Dogon',  
'cliffs', ',', 'Mali']], 7, True)

['Bangi-me', 'Language', 'codes', 'ISO', '639-3', 'dba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'nhw', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nhw', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 9, True)

['Mezquital', 'Spoken', 'in', 'M', '\xc3', '\xa9', 'xico', '(', 'Hidalgo',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'M',  
'\xc3', '\xa9', 'xico', '\\(', 'Hidalgo', '\\)']]], 10, True)

['Mezquital', 'Spoken', 'in', 'United', 'States', 'North', 'Carolina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Mezquital', 'Spoken', 'in', 'Florida']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mezquital', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Mezquital', 'Language', 'codes', 'ISO', '639-2', 'oto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Mezquital', 'Language', 'codes', 'ISO', '639-3', 'ote']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Highland', 'Spoken', 'in', 'Mexico', '(', 'Puebla', ',', 'Veracruz',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',  
'\\(', 'Puebla', ',', 'Veracruz', '\\)']]], 9, True)

['Highland', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Type: ANY', 'Type: ANY']]], 5, True)

['Highland', 'Language', 'codes', 'ISO', '639-3', 'tos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tzotzil', 'Spoken', 'in', 'M', '\xc3', '\xa9', 'xico', '(', 'Chiapas',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'M',  
'\xc3', '\xa9', 'xico', '\\(', 'Chiapas', '\\)']]], 10, True)

['Tzotzil', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Tzotzil', 'Language', 'codes', 'ISO', '639-3', 'tzc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bachaj\xc3\xb3n', 'Spoken', 'in', 'Mexico', '(', 'Chiapas', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Bachaj\xc3\xb3n', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Bachaj\xc3\xb3n', 'Language', 'codes', 'ISO', '639-3', 'tzb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Loun', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Loun', 'Region', 'Maluku']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Loun', 'Writing', 'system', 'Latin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Loun', 'Language', 'codes', 'ISO', '639-3', 'lox']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Andoa', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Andoa', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Andoa', 'Language', 'codes', 'ISO', '639-3', 'anb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bustocco', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bustocco', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Comasco-Lecchese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Comasco-Lecchese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2',  
'\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Vallassinese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Vallassinese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2',  
'\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Omurano', 'Language', 'extinction', '1958']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```



```
['Omurano', 'Spoken', 'in', 'Peru', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Omurano', 'Spoken', 'in', 'Maina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Omurano', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Omurano', 'Language', 'codes', 'ISO', '639-3', 'omu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nuares', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nuares', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Varesino', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Varesino', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)
```

```
['Amuesha', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Amuesha', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Amuesha', 'Pronunciation', 'jane', '\xca', '\x83', 'a', '\xcb', '\x80',
```

't', '\xcd', '\xa1', '\xca', '\x83', '\xca', '\xb0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'jane',  
'\xca', '\x83', 'a', '\xcb', '\x80', 't', '\xcd', '\xa1', '\xca', '\x83',  
'\xca', '\xb0']]], 15, True)

['Amuesha', 'Language', 'codes', 'ISO', '639-3', 'ame']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cremonese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cremonese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Southwestern', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Southwestern', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2',  
'\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Tarantino', 'Spoken', 'in', 'Italy', ',', 'United', 'States', '(',  
'California', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Italy',  
',', 'United', 'States', '\\(', 'California', '\\)']]], 10, True)

['Tarantino', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Lecchese', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lecchese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',

'639-3']], 10, True)

['Comasco', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Comasco', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80', '\x93', 'Lombard']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Nyunga', 'Spoken', 'in', 'Western', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Nyunga', 'Language', 'extinction', 'disputed']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nyunga', 'Dialects', 'Tjapanmay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyunga', 'Dialects', 'Karlamay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyunga', 'Dialects', 'Pipelman', '(', 'Pipalman', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Nyunga', 'Dialects', 'Ngatjumay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyunga', 'Dialects', 'Kwetjman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyunga', 'Dialects', 'Mirnong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nyunga', 'Dialects', 'Kaniyang', 'Pindjarup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Nyunga', 'Dialects', 'Whadjuk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Nyunga', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nyunga', 'Language', 'codes', 'ISO', '639-3', 'nys']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lugbara', 'Spoken', 'in', 'Uganda', ',', 'DR', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Uganda', ',', 'DR', 'Congo']]], 7, True)

['Lugbara', 'Writing', 'system', '--']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lugbara', 'Language', 'codes', 'ISO', '639-3', 'lgg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Arosi', 'Spoken', 'in', 'Solomon', 'Islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Arosi', 'Total', 'speakers', '6750']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arosi', 'Dialects', 'Wango']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Arosi', 'Language', 'codes', 'ISO', '639-3', 'aia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lagh\xc3\xa9\xc3\xa9', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lagh\xc3\xa9\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'lmo',  
'\xe2', '\x80', '\x93', 'Lombard']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Pavese', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pavese', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Spasell', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Spasell', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']],  
3, True)

['Spasell', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
'\x93', 'Lombard']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lmo', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Zou', 'Spoken', 'in', 'Myanmar', ',', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Zou', 'Region', 'Tiddim']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Zou', 'Region', 'Chin', 'Hills', ';']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Zou', 'Region', 'In', 'India', ':', 'Manipur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'In', 'India',

['Zou', 'Region', 'Chandel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Zou', 'Region', 'Singngat', 'subdivision', 'and', 'Sungnu', 'area', ';', 'Churachandpur', 'districts', ';', 'Assam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Singngat', 'subdivision', 'and', 'Sungnu', 'area', ';', 'Churachandpur', 'districts', ';', 'Assam']], 12, True)

['Zou', 'Language', 'codes', 'ISO', '639-3', 'zom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Zo'\xc3\xa9", 'Spoken', 'in', 'Brazil', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

["Zo'\xc3\xa9", 'Region', 'Par', '\xc3', '\xa1']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

["Zo'\xc3\xa9", 'Region', 'Obidos', 'Municipality']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Zo'\xc3\xa9", 'Region', 'Cuminapanema', 'River', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

["Zo'\xc3\xa9", 'Language', 'codes', 'ISO', '639-3', 'pto']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nkutu', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Nkutu', 'Language', 'codes', 'ISO', '639-3', 'nkw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kara', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

```

['Kara', 'Language', 'codes', 'ISO', '639-3', 'leu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ashtiani', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Ashtiani', 'Language', 'codes', 'ISO', '639-3', 'atn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bolango', 'Dialects', 'Atinggola']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Bolango', 'Region', 'Gorontalo', ',', 'Sulawesi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Bolango', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bolango', 'Language', 'codes', 'ISO', '639-3', 'bld']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nangalami', 'Spoken', 'in', 'Afghanistan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nangalami', 'Language', 'codes', 'ISO', '639-3', 'nli']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Puxian', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Puxian', 'Spoken', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Puxian', 'Spoken', 'in', 'Singapore']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Puxian', 'Spoken', 'in', 'Taiwan', '(', 'Wuqiu', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Puxian', 'Spoken', 'in', 'United', 'States', '(', 'California', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United',  
'States', '\\(', 'Type: ANY', '\\)']]], 8, True)

['Puxian', 'Writing', 'system', 'Chinese', 'characters']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Puxian', 'Language', 'codes', 'ISO', '639-3', 'cpx']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mak\\xc3\\xa1', 'Total', 'speakers', '1500']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mak\\xc3\\xa1', 'Region', 'Presidente', 'Hayes', 'Department', ',',  
'Asunci', '\\xc3', '\\xb3', 'n']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Presidente',  
'Hayes', 'Department', ',', 'Asunci', '\\xc3', '\\xb3', 'n']]], 10, True)

['Mak\\xc3\\xa1', 'Spoken', 'in', 'Paraguay']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mak\\xc3\\xa1', 'Pronunciation', 'ma', '\\xcb', '\\x88', 'ka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ma',  
'\\xcb', '\\x88', 'ka']]], 6, True)

['Mak\\xc3\\xa1', 'Language', 'codes', 'ISO', '639-3', 'mca']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Central', 'Atlas', 'Tamazight', 'Total', 'speakers', 'Between', '3',  
'and', '5', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Atlas', 'Tamazight',  
'Total', 'speakers', 'Between', 'NUM', 'and', 'NUM', 'million']], ['NUM',  
'3'], ['NUM', '5']], 10, True)

['Central', 'Spoken', 'in', 'Morocco', ',', '']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ', ']]], 5, True)

['Central', 'Spoken', 'in', 'much', 'fewer', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'much', 'fewer', 'in']]]], 6, True)

['Central', 'Spoken', 'in', ':', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', ':', 'Type: ANY']]]], 5, True)

['Central', 'Writing', 'system', 'Tifinagh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Central', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Central', 'Writing', 'system', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Central', 'Pronunciation', 't', '\xc3', '\xa6', 'm', '\xc3', '\xa6', '\xcb', '\x88', 'z', '\xc9', '\xaa', '\xc9', '\xa3', 't', ', ', '\xce', '\xb8', '\xc3', '\xa6', 'm', '\xc3', '\xa6', '\xcb', '\x88', 'z', '\xc9', '\xaa', '\xc9', '\xa3', '\xce', '\xb8']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 't', '\xc3', '\xa6', 'm', '\xc3', '\xa6', '\xcb', '\x88', 'z', '\xc9', '\xaa', '\xc9', '\xa3', 't', ', ', '\xce', '\xb8', '\xc3', '\xa6', 'm', '\xc3', '\xa6', '\xcb', '\x88', 'z', '\xc9', '\xaa', '\xc9', '\xa3', '\xce', '\xb8']]]], 33, True)

['Central', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]]], 5, True)

['Central', 'Regulated', 'by', 'IRCAM']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]]], 4, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'tzm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mappila', 'Spoken', 'in', 'India', 'and', 'a', 'few', 'other',

```

'countries']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',
'and', 'a', 'few', 'other', 'countries']]], 9, True)

['Mappila', 'Total', 'speakers', 'Unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mappila', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Western', 'Spoken', 'in', 'Wales']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Western', 'Spoken', 'in', 'Welsh', 'Marches']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Western', 'Spoken', 'in', 'Pennines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Western', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Maasina', 'Fulfulde', 'Total', 'speakers', '1', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',
'speakers', 'NUM', 'million']]], ['NUM', '1']], 6, True)

['Maasina', 'Region', 'Inland', 'Niger', 'Delta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Maasina', 'Language', 'codes', 'ISO', '639-2', 'ful']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Maasina', 'Language', 'codes', 'ISO', '639-3', 'ffm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Maasina', 'Language', 'codes', 'ISO', '639-1', 'ff']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Butuanon', 'Spoken', 'in', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Butuanon', 'Language', 'codes', 'ISO', '639-3', 'btw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Swiss', 'Spoken', 'in', 'Switzerland', ',', 'northern', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Switzerland', ',', 'northern', 'Italy']], 7, True)

['Swiss', 'Regulated', 'by']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Swiss', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Solombala-English', 'Spoken', 'in', 'Solombala']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Solombala-English', 'Spoken', 'in', 'Arkhangelsk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Solombala-English', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Solombala-English', 'Writing', 'system', 'Russian', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Solombala-English', 'Language', 'codes', 'ISO', '639-2', 'crp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Solombala-English', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Toki', 'Pona', 'Total', 'speakers', 'three', 'said', 'to', 'be',  
'fluent', ';', 'several', 'dozen', 'with', 'internet', 'chat', 'ability']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pona', 'Total',  
'speakers', 'three', 'said', 'to', 'be', 'fluent', ';', 'several', 'dozen',  
'with', 'internet', 'chat', 'ability']]], 16, True)

['Toki', 'Date', 'founded', '2001']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Toki', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]], 7, True)

['Toki', 'Category', '(', 'purpose', ')', 'combining', 'elements', 'of',  
'the', 'subgenres', 'personal', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'combining', 'elements', 'of', 'the', 'subgenres',  
'personal', 'language']]], 12, True)

['Toki', 'Category', '(', 'purpose', ')', 'international', 'auxiliary',  
'language', 'and', 'philosophical', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'international', 'auxiliary', 'language', 'and',  
'philosophical', 'language']]], 11, True)

['Toki', 'Pronunciation', 'toki', 'pona']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Toki', 'Created', 'by', 'Sonja', 'Elen', 'Kisa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Sonja',  
'Elen', 'Kisa']]], 6, True)

['Toki', 'Category', '(', 'sources', ')', 'a', 'posteriori', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'a', 'posteriori', 'language']]], 8, True)

['Toki', 'Category', '(', 'sources', ')', 'with', 'elements', 'of',  
'English']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'with', 'elements', 'of', 'Type: ANY']]], 9, True)

['Toki', 'Category', '(', 'sources', ')', 'Tok', 'Pisin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',

'sources', '\\)', 'Type: ANY', 'Type: ANY']]], 7, True)

['Toki', 'Category', '(', 'sources', ')', 'Finnish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)

['Toki', 'Category', '(', 'sources', ')', 'Georgian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)

['Toki', 'Category', '(', 'sources', ')', 'Dutch']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)

['Toki', 'Category', '(', 'sources', ')', 'Acadian', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY', 'Type: ANY']]], 7, True)

['Toki', 'Category', '(', 'sources', ')', 'Esperanto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)

['Toki', 'Category', '(', 'sources', ')', 'Croatian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)

['Toki', 'Category', '(', 'sources', ')', 'Chinese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)

['Toki', 'Setting', 'and', 'usage', 'testing', 'principles', 'of',  
'minimalism', ',', 'the', 'Sapir', '\\xe2', '\\x80', '\\x93', 'Whorf',  
'hypothesis', 'and', 'pidgins']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'testing', 'principles', 'of', 'minimalism', ',', 'the', 'Sapir',  
'\\xe2', '\\x80', '\\x93', 'Whorf', 'hypothesis', 'and', 'pidgins']]], 18,  
True)

['Toki', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tigak', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

```
['Tigak', 'Language', 'codes', 'ISO', '639-2', 'map', '(', 'group', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'map', '\\(', 'group', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Tigak', 'Language', 'codes', 'ISO', '639-3', 'tgc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Iquito', 'Spoken', 'in', 'Peru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Iquito', 'Total', 'speakers', '25', 'native', '/', 'fluent', ',', '25',
'partial', '/', 'passive']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', 'native', '/', 'fluent', ',', 'NUM', 'partial', '/', 'passive']],
['NUM', '25'], ['NUM', '25']], 12, True)
```

```
['Iquito', 'Writing', 'system', 'Roman', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Iquito', 'Official', 'language', 'in', 'Peru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)
```

```
['Iquito', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)
```

```
['Iquito', 'Language', 'codes', 'ISO', '639-3', 'iqu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wuvulu-Aua', 'Region', 'Wuvulu', 'and', 'Aua', 'Islands', ',', 'Manus',
'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Wuvulu',
'and', 'Aua', 'Islands', ',', 'Manus', 'Province']]], 9, True)
```

```
['Wuvulu-Aua', 'Language', 'codes', 'ISO', '639-3', 'wuv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dhuwal', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)
```

```
['Dhuwal', 'Total', 'speakers', '3', ',', '650', '(', 'Ethnologue', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '3'], ['NUM',
'650']], 9, True)
```

```
['Dhuwal', 'Language', 'codes', 'ISO', '639-3', 'duj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Fagani', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Fagani', 'Total', 'speakers', '900']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Fagani', 'Dialects', 'Rihu'a']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Fagani', 'Dialects', 'Agufi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Fagani', 'Language', 'codes', 'ISO', '639-3', 'faf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kahua', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Kahua', 'Total', 'speakers', '5200']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kahua', 'Language', 'codes', 'ISO', '639-3', 'agw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bauro', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Bauro', 'Total', 'speakers', '5000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bauro', 'Dialects', 'Haununu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Bauro', 'Dialects', 'Ravo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Bauro', 'Language', 'codes', 'ISO', '639-3', 'bxa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Owa', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Owa', 'Total', 'speakers', '8400']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Owa', 'Dialects', 'Tawarafa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Owa', 'Dialects', 'Owa', 'Raha']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Owa', 'Dialects', 'Owa', 'Riki']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Owa', 'Language', 'codes', 'ISO', '639-3', 'stn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Indonesian', 'Signed', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Indonesian', 'Total', 'signers', '(', 'unknown', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\\
```



(', 'unknown', '\\\')]]], 6, True)

['Ayoreo', 'Spoken', 'in', 'Paraguay', ',', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Ayoreo', 'Region', 'Chaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ayoreo', 'Region', 'Alto', 'Paraguay', 'departments', '(', 'Paraguay', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alto', 'Paraguay', 'departments', '\\(', 'Paraguay', '\\\')']]], 8, True)

['Ayoreo', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-2']], 8, True)

['Ayoreo', 'Language', 'codes', 'ISO', '639-3', 'ayo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chamacoco', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chamacoco', 'Language', 'codes', 'ISO', '639-3', 'ceg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nyole', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nyole', 'Region', 'Tororo', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nyole', 'Language', 'codes', 'ISO', '639-3', 'nuj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kui', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Kui', 'Language', 'codes', 'ISO', '639-3', 'kxu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Muher', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Muher', 'Language', 'codes', 'ISO', '639-3', 'sgw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koya', 'Spoken', 'in', 'India', 'Chinturu', ',', '']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',
'Chinturu', ',', '']], 6, True)

['Koya', 'Spoken', 'in', 'Andhra', 'Pradesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Sauria', 'Spoken', 'in', 'Bihar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sauria', 'Spoken', 'in', 'West', 'Bengal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Sauria', 'Spoken', 'in', 'Bangladesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sauria', 'Writing', 'system', 'Devnagri']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sauria', 'Language', 'codes', 'ISO', '639-3', 'mjt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kullu', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Kullu', 'Spoken', 'in', 'India']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kullu', 'Region', 'Himachal', 'Pradesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kullu', 'Writing', 'system', 'Devanagari', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kullu', 'Language', 'codes', 'ISO', '639-3', 'kfx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['U\u005beice', 'dialect', 'Spoken', 'in', 'primarily', 'Serbia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Spoken', 'in', 'primarily', 'Serbia']]], 6, True)

['U\u005beice', 'Pronunciation', '\xcb', '\x88', 'u', '\xca', '\x83', '\xc9', '\x99', '\xcc', '\x86', 't', '\xca', '\x83', 'ki', '\xcb', '\x90', '\xcb', '\x88', '\xc9', '\xa1', '\xc9', '\x94', '\xca', '\x8b', '\xc9', '\x94', '\xcb', '\x90', 'r']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'u', '\xca', '\x83', '\xc9', '\x99', '\xcc', '\x86', 't', '\xca', '\x83', 'ki', '\xcb', '\x90', '\xcb', '\x88', '\xc9', '\xa1', '\xc9', '\x94', '\xca', '\x8b', '\xc9', '\x94', '\xcb', '\x90', 'r']]], 30, True)

['U\u005beice', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Aushiri', 'Language', 'extinction', 'extinct']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aushiri', 'Spoken', 'in', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aushiri', 'Language', 'codes', 'ISO', '639-3', 'avs']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hinukh', 'Region', 'Southern', 'Dagestan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hinukh', 'Spoken', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hinukh', 'Pronunciation', 'hi', '\xcb', '\x88', 'nuzas', 'm', '\xc9', '\x9b', 't', '\xcd', '\xa1', 's']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'hi', '\xcb', '\x88', 'nuzas', 'm', '\xc9', '\x9b', 't', '\xcd', '\xa1', 's']]], 13, True)

['Hinukh', 'Language', 'codes', 'ISO', '639-3', 'gin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Banat', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Banat', 'Language', 'codes', 'ISO', '639-3', 'bul-ban']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sanum\xc3\xa1', 'Spoken', 'in', 'Venezuela', ',', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Sanum\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'xsu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ghadam\xc3\xa8s', 'Total', 'speakers', '4000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ghadam\xc3\xa8s', 'Spoken', 'in', 'Libya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ghadam\xc3\xa8s', 'Region', 'Ghadam', '\xc3', '\xa8', 's']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ghadam', '\xc3', '\xa8', 's']]], 6, True)

['Ghadam\xc3\xa8s', 'Language', 'codes', 'ISO', '639-3', 'gha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karamanli', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Karamanli', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Karamanli', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Karamanli', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Macedonia']], 6, True)

['Karamanli', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Karamanli', 'Writing', 'system', 'Cyrillic', 'alphabet', ',', 'Greek',  
'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Karamanli', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Karamanli', 'Language', 'codes', 'ISO', '639-3', 'tur-kar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sadri', 'Spoken', 'in', 'India', ',', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Sadri', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Sadri', 'Language', 'codes', 'ISO', '639-2', 'bih']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sadri', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sadri', 'Language', 'codes', 'ISO', '639-3', ':', 'sck', '\xe2', '\x80',  
\x93', 'Sadri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'sck', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Sadri', 'Language', 'codes', 'ISO', '639-3', 'sdr', '\xe2', '\x80',  
\x93', 'Oraon', 'Sadri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sdr', '\xe2', '\x80', '\x93', 'Oraon', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Nafusi', 'Spoken', 'in', 'Libya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nafusi', 'Total', 'speakers', '184', ',', '000', 'inc', '.', 'Zuara',  
';']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'inc', '.', 'Type: ANY', 'Type: ANY']], ['NUM', '184'],  
['NUM', '000']], 10, True)

['Nafusi', 'Language', 'codes', 'ISO', '639-3', 'jbn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ningbo', 'dialect', 'Total', 'speakers', 'approx', '.', '5-6',  
'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '5-6']], 8, True)

['Ningbo', 'Region', 'Ningbo', ',', 'Zhejiang', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
,',', 'Zhejiang', 'province']]], 6, True)

['Ningbo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Singaporean', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Singaporean', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Singaporean', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Singaporean', 'Language', 'codes', 'ISO', '639-3', 'cmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Singaporean', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Kuku', 'Spoken', 'in', 'Southern', 'Sudan', 'and', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

['Kuku', 'Language', 'codes', 'ISO', '639-3', 'bfa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Noxilo', 'Total', 'speakers', '?', '?', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '?', '?', '?']], 6, True)

['Noxilo', 'Date', 'founded', '1997']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Noxilo', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY', 'language']], 7, True)

['Noxilo', 'Category', '(', 'purpose', ')', 'Noxilo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY']], 6, True)

```
['Noxilo', 'Created', 'by', 'Mizuta', 'Sentaro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'vocabulary', 'from',  
'English']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'vocabulary', 'from', 'Type: ANY']]], 8, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'Chinese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'Japanese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'Korean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)
```

```
['Noxilo', 'Category', '(', 'sources', ')', 'etc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'Type: ANY']]], 6, True)
```

```
['Noxilo', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Mongo', 'Spoken', 'in', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Mongo', 'Language', 'codes', 'ISO', '639-2', 'lol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```



```

['Mongo', 'Language', 'codes', 'ISO', '639-3', 'lol']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tayo', 'Language', 'codes', 'ISO', '639-3', 'cks']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bravanese', 'Spoken', 'in', 'Somalia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Bravanese', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']],
3, True)

['Bravanese', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Bravanese', 'Language', 'codes', 'ISO', '639-3', 'swl', '\xe2', '\x80',
'\x93', 'Chimwiini']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'swl', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Bravanese', 'Language', 'codes', 'ISO', '639-3', '(', 'generic', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\(', 'generic', '\\)']], ['NUM', '639-3']], 8, True)

['Gaddang', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Gaddang', 'Region', 'Luzon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Gaddang', 'Language', 'codes', 'ISO', '639-3', 'gad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Negeri', 'Spoken', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

```

```
['Negeri', 'Language', 'codes', 'ISO', '639-2', 'msa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Negeri', 'Language', 'codes', 'ISO', '639-3', 'zmi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Urhobo', 'Spoken', 'in', 'Delta', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Urhobo', 'Language', 'codes', 'ISO', '639-3', 'urh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Southern', 'Region', 'Yakutia', 'and', 'Magadan', 'Oblast']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Yakutia',
'and', 'Magadan', 'Oblast']], 6, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'yux']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mundari', 'Spoken', 'in', 'Southern', 'Sudan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Mundari', 'Language', 'codes', 'ISO', '639-3', 'mqu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Osing', 'Spoken', 'in', 'Java', ',', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Osing', 'Writing', 'system', 'Javanese']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Osing', 'Language', 'codes', 'ISO', '639-3', 'osi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pacoh', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pacoh', 'Total', 'speakers', '30', ',', '000', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'speakers']], ['NUM', '30'], ['NUM', '000']], 7, True)

['Pacoh', 'Language', 'codes', 'ISO', '639-3', 'pac']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kusunda', 'Spoken', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kusunda', 'Region', 'Gandaki', 'Zone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kusunda', 'Language', 'codes', 'ISO', '639-3', 'kkg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bantawa', 'Spoken', 'in', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bantawa', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Bantawa', 'Language', 'codes', 'ISO', '639-3', 'bap']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Colombian', 'Signed', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Colombian', 'Total', 'signers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\xe2', '\x80', '\x93']], 6, True)

['Ormuri', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ormuri', 'Language', 'codes', 'ISO', '639-3', 'oru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Parachi', 'Total', 'speakers', '600']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Parachi', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Parachi', 'Region', 'Kabul', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Parachi', 'Region', '6000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Parachi', 'Language', 'codes', 'ISO', '639-3', 'prc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Beurla', 'Reagaird', 'Total', 'speakers', 'Nearly', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Reagaird', 'Total', 'speakers', 'Nearly', 'extinct']], 6, True)

['Beurla', 'Pronunciation', '\xcb', '\x88', 'pj', '\xc9', '\xa4', '\xcb', '\x90', 'r', '\xcb', '\xa0', 'l', '\xcb', '\xa0', '\xcc', '\xaa', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xbe', '\xc9', '\x9b', 'k', '\xc9', '\x99', 'r', '\xcb', '\xa0', 't', '\xca', '\xb2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'pj', '\xc9', '\xa4', '\xcb', '\x90', 'r', '\xcb', '\xa0', 'l', '\xcb', '\xa0', '\xcc', '\xaa', '\xc9', '\x99', '\xcb', '\x88', '\xc9', '\xbe', '\xc9', '\x9b', 'k', '\xc9', '\x99', 'r', '\xcb', '\xa0', 't', '\xca', '\xb2']], 34, True)

['Beurla', 'Spoken', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Beurla', 'Official', 'language', 'in', 'Scotland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Beurla', 'Regulated', 'by']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Beurla', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Amurdag', 'Spoken', 'in', 'Oenpelli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Amurdag', 'Spoken', 'in', 'Goulburn', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Amurdag', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Amurdag', 'Total', 'speakers', '1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Amurdag', 'Language', 'codes', 'ISO', '639-3', 'amg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Canichana', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Canichana', 'Total', 'speakers', '(', 'extinct', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\  
(', 'extinct', '\\)']]], 6, True)
```

```
['Canichana', 'Language', 'codes', 'ISO', '639-3', 'caz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kallawaya', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Kallawaya', 'Language', 'codes', 'ISO', '639-3', 'caw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zaghawa', 'Dialects', 'Bidayat', '(', 'Tuba', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\(', 'Type: ANY', '\\)']]], 6, True)

['Zaghawa', 'Dialects', 'Wagi', '(', 'Twer', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\(', 'Type: ANY', '\\)']]], 6, True)

['Zaghawa', 'Dialects', 'Kube']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Zaghawa', 'Region', 'Chad', ',', 'northwestern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Chad', ',',  
'northwestern', 'Sudan']]], 6, True)

['Zaghawa', 'Spoken', 'in', 'Chad', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Zaghawa', 'Language', 'codes', 'ISO', '639-3', 'zag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tupuri', 'Spoken', 'in', 'Chad', ',', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Tupuri', 'Language', 'codes', 'ISO', '639-3', 'tui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Semelai', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Semelai', 'Region', 'Malaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

```

['Semelai', 'Language', 'codes', 'ISO', '639-3', 'sza']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bor\xc3\xb4ro', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bor\xc3\xb4ro', 'Region', 'central', 'Mato', 'Grosso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Bor\xc3\xb4ro', 'Language', 'codes', 'ISO', '639-3', 'bor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rusenu', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Itawis', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Itawis', 'Region', 'Luzon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Itawis', 'Language', 'codes', 'ISO', '639-3', 'itv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Salvadoran', 'Total', 'signers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Salvadoran', 'Language', 'codes', 'ISO', '639-2', 'sgn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Gokana', 'Spoken', 'in', 'Rivers', 'State', ',', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Gokana', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',

```

'\xe2', '\x80', '\x93']]], 6, True)

['Gokana', 'Language', 'codes', 'ISO', '639-3', 'gkn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Onge', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Onge', 'Total', 'speakers', '94']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Boga', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Boga', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Boga', 'Language', 'codes', 'ISO', '639-3', 'bvw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ga'anda", 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Ga'anda", 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Ga'anda", 'Language', 'codes', 'ISO', '639-3', 'gqa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jara', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jara', 'Region', 'Borno', 'and', 'Gombe', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Borno', 'and',  
'Gombe', 'States']]], 6, True)



```
['Jara', 'Region', 'Biu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Jara', 'Region', 'Kwaya-Kusar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Jara', 'Region', 'Akko', 'and', 'Yamaltu-Deba', 'LGAs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Akko', 'and',
'Yamaltu-Deba', 'LGAs']]], 6, True)
```

```
['Jara', 'Language', 'codes', 'ISO', '639-3', 'jaf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cibak', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Cibak', 'Region', 'Borno', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Cibak', 'Language', 'codes', 'ISO', '639-3', 'ckl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kofa', 'Total', 'speakers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kofa', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kofa', 'Region', 'Adamawa', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kofa', 'Language', 'codes', 'ISO', '639-3', 'kso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pandan', 'Spoken', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Pandan', 'Language', 'codes', 'ISO', '639-3', 'cts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iriga', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Iriga', 'Region', 'Bicol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Iriga', 'Writing', 'system', 'Historically', 'written', 'in',  
'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Iriga', 'Official', 'language', 'in', 'Regional', 'language', 'in',  
'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Iriga', 'Language', 'codes', 'ISO', '639-3', 'bto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tswa', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tswa', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Tswa', 'Language', 'codes', 'ISO', '639-3', 'tsc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ronga', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ronga', 'Spoken', 'in', 'South', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Ronga', 'Region', 'South', 'of', 'Maputo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Ronga', 'Language', 'codes', 'ISO', '639-3', 'rng']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yeyi', 'Spoken', 'in', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yeyi', 'Spoken', 'in', 'Botswana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yeyi', 'Region', 'along', 'the', 'Okavango', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'along', 'the',  
'Okavango', 'River']]], 6, True)

['Yeyi', 'Language', 'codes', 'ISO', '639-3', 'yey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ofo', 'Total', 'speakers', '\xe2', '\x80', '\x94']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'\xe2', '\x80', '\x94']]], 6, True)

['Ofo', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Ofo', 'Region', 'Mississippi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ofo', 'Language', 'codes', 'ISO', '639-3', 'of']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lau', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

```
['Lau', 'Region', 'Northeast', 'Malaita', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Lau', 'Language', 'codes', 'ISO', '639-3', 'llu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Lore diakarkar', 'Total', 'speakers', '340']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lore diakarkar', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lore diakarkar', 'Region', 'Santo', 'Island', ',', 'east', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Santo',  
'Island', ',', 'east', 'coast']]], 7, True)
```

```
['Lore diakarkar', 'Language', 'codes', 'ISO', '639-3', 'lnn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Shark', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Shark', 'Language', 'codes', 'ISO', '639-3', 'ssv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kamoro', 'Total', 'speakers', '8000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kamoro', 'Language', 'codes', 'ISO', '639-3', 'kgq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Yaur', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Yaur', 'Region', 'Irian', 'Jaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Yaur', 'Language', 'codes', 'ISO', '639-3', 'jau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baropasi', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Baropasi', 'Total', 'speakers', '2500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Baropasi', 'Language', 'codes', 'ISO', '639-3', 'brp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Banggai', 'Spoken', 'in', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Banggai', 'Language', 'codes', 'ISO', '639-3', 'bgz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iwam', 'Spoken', 'in', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Iwam', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Iwam', 'Language', 'codes', 'ISO', '639-3', 'iwm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rennellese', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Rennellese', 'Total', 'speakers', '4400']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Rennellese', 'Dialects', 'Munggava', '(', 'Mugaba', ',', 'Rennell', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Munggava',  
'\(', 'Mugaba', ',', 'Rennell', '\)']]]], 8, True)
```

```
['Rennellese', 'Dialects', 'Mungiki', '(', 'Mugiki', ',', 'Bellona', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Mungiki', '\(  
, 'Mugiki', ',', 'Bellona', '\)']]]], 8, True)
```

```
['Rennellese', 'Language', 'codes', 'ISO', '639-3', 'mnv']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Yao', 'Spoken', 'in', 'Malawi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Yao', 'Spoken', 'in', 'Mozambique']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Yao', 'Spoken', 'in', 'Tanzania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Yao', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]]], 5, True)
```

```
['Yao', 'Language', 'codes', 'ISO', '639-2', 'yao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Yao', 'Language', 'codes', 'ISO', '639-3', 'yao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Parkari', 'Spoken', 'in', 'Sindh', 'province', ',', 'Pakistan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Sindh',  
'province', ',', 'Pakistan']]]], 7, True)
```

```
['Parkari', 'Language', 'codes', 'ISO', '639-3', 'kvx']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gnau', 'Region', 'Sandaun', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Gnau', 'Region', 'Namblo', 'Census', 'Division']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Gnau', 'Region', 'northwest', 'of', 'Maimai', '.', 'Coordinates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northwest',  
'of', 'Maimai', '.', 'Coordinates']]], 7, True)

['Gnau', 'Language', 'codes', 'ISO', '639-3', 'gnu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Juang', 'Spoken', 'in', 'India', '(', 'Orissa', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)

['Juang', 'Language', 'codes', 'ISO', '639-3', 'jun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tai', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tai', 'Region', 'Lopburi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Tai', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Ngawun', 'Language', 'extinction', '1977']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ngawun', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ngawun', 'Language', 'codes', 'ISO', '639-3', 'nxn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kesen', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kesen', 'Region', 'Iwate', 'Prefecture', ':', 'Kamaishi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Iwate', 'Prefecture', ':', 'Kamaishi']], 6, True)

['Kesen', 'Region', '\xc5', '\x8c', 'funato']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kesen', 'Region', 'Rikuzentakata']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kesen', 'Region', 'and', 'Sumita']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kesen', 'Official', 'language', 'in', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Kesen', 'Regulated', 'by', '(', 'Harutsugu', 'Yamaura', 'described', 'the', 'verbal', 'speech', 'and', 'created', 'the', 'orthography', '.', ')]':  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', '\\(', 'Harutsugu', 'Yamaura', 'described', 'the', 'verbal', 'speech', 'and', 'created', 'the', 'orthography', '.', '\\)']], 16, True)

['Kesen', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Zaiwa', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zaiwa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Zaiwa', 'Language', 'codes', 'ISO', '639-3', 'atb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



```
['Ai-Cham', 'Spoken', 'in', 'China']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ai-Cham', 'Region', 'Qiannan', 'Prefecture', ',', 'Guizhou', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Qiannan',  
'Prefecture', ',', 'Guizhou', 'Province']]], 7, True)
```

```
['Ai-Cham', 'Language', 'codes', 'ISO', '639-3', 'aih']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Dhudhuroa', 'Spoken', 'in', 'North-eastern', 'Victoria', ',',  
'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)
```

```
['Dhudhuroa', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Paniya', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Paniya', 'Region', 'Kerala', 'State']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Paniya', 'Language', 'codes', 'ISO', '639-3', 'tcy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bushi', 'Spoken', 'in', 'Madagascar', ',', 'Mayotte']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Bushi', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Bushi', 'Language', 'codes', 'ISO', '639-3', 'buc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Ndom', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ndom', 'Total', 'speakers', '1200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ndom', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Ndom', 'Language', 'codes', 'ISO', '639-3', 'nqm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kedah', 'Spoken', 'in', 'Malaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kedah', 'Language', 'codes', 'ISO', '639-2', 'may', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'may', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)
```

```
['Kedah', 'Language', 'codes', 'ISO', '639-2', 'msa', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'msa', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)
```

```
['Kedah', 'Language', 'codes', 'ISO', '639-3', 'meo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kedah', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Palawa', 'Spoken', 'in', 'Tasmania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Palawa', 'Total', 'speakers', 'native', ':', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'native', ':', 'none']]], 6, True)
```

```
['Palawa', 'Language', 'codes', 'ISO', '639-3', 'xtz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rikbakts\xc3\xa1,', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Rikbakts\xc3\xa1,', 'Region', 'Mato', 'Grosso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Rikbakts\xc3\xa1,', 'Language', 'codes', 'ISO', '639-3', 'rkb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abron', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Abron', 'Official', 'language', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']], 5, True)

['Abron', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']], 6, True)

['Abron', 'Language', 'codes', 'ISO', '639-3', 'abr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Burgundian', 'Region', 'Burgundy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Burgundian', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Burgundian', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Burgundian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Glosa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Glosa', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Glosa', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Glosa', 'Category', '(', 'purpose', ')', 'Glosa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Glosa', 'Created', 'by', 'Ronald', 'Clark', 'and', 'Wendy', 'Ashby', ',',  
'based', 'on', 'the', 'Interglossa', 'of', 'Lancelot', 'Hogben']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
'Ronald', 'Clark', 'and', 'Wendy', 'Ashby', ',', 'based', 'on', 'the',  
'Interglossa', 'of', 'Lancelot', 'Hogben']]], 16, True)

['Glosa', 'Setting', 'and', 'usage', 'international', 'auxiliary',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Glosa', 'Language', 'codes', 'ISO', '639-3', 'igs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Loko', 'Spoken', 'in', 'Sierra', 'Leone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Loko', 'Language', 'codes', 'ISO', '639-3', 'lok']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abui', 'Spoken', 'in', 'Alor', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Abui', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',

'Type: ANY', 'Type: ANY']]], 5, True)

['Abui', 'Language', 'codes', 'ISO', '639-3', 'abz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Majang', 'Total', 'speakers', '15', ',', '341', '(', 'Ethnologue', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '15'], ['NUM',  
'341']], 9, True)

['Majang', 'Spoken', 'in', 'Ethiopia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Majang', 'Region', 'Godere', ',', 'Gambela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Majang', 'Language', 'codes', 'ISO', '639-3', 'mpe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Beray', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Beray', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Beray', 'Language', 'codes', 'ISO', '639-3', 'bve']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Beray', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Barmoodan', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']]], 6, True)

['Barmoodan', 'Setting', 'and', 'usage']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Barmoodan', 'Created', 'by', 'John', 'Taylor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Barmoodan', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Barmoodan', 'Category', '(', 'purpose', ')', 'Barmoodan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)

['Barmoodan', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Barmoodan', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Beti', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beti', 'Spoken', 'in', 'Equatorial', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Beti', 'Spoken', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beti', 'Total', 'speakers', '2', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '2']], 5, True)

['Beti', 'Dialects', 'Bebele']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Beti', 'Dialects', 'Bebil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Beti', 'Dialects', 'Bulu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Beti', 'Dialects', 'Eton']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Beti', 'Dialects', 'Ewondo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Beti', 'Dialects', 'Fang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Beti', 'Dialects', 'Mengisa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Beti', 'Language', 'codes', 'ISO', '639-3', 'btb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kako', 'Language', 'codes', 'ISO', '639-3', 'kkj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lop', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Lop', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lop', 'Language', 'codes', 'ISO', '639-2', 'uig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lop', 'Language', 'codes', 'ISO', '639-3', 'uig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lop', 'Language', 'codes', 'ISO', '639-1', 'ug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Daur', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Daur', 'Region', 'Inner', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Daur', 'Region', 'Hailar', 'District', ';', 'Heilongjiang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Hailar', 'District', ';', 'Heilongjiang', 'Province']], 7, True)

['Daur', 'Region', 'Qiqihar', 'Prefecture', ';', 'Xinjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Qiqihar', 'Prefecture', ';', 'Xinjiang']], 6, True)

['Daur', 'Region', 'Tacheng', 'Prefecture']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Daur', 'Language', 'codes', 'ISO', '639-3', 'dta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Achagua', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Achagua', 'Language', 'codes', 'ISO', '639-3', 'aca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Andaqui', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Andaqui', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Andaqui', 'Region', 'southern', 'highlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Andaqui', 'Language', 'codes', 'ISO', '639-3', 'ana']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Low', 'Lusatian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']], 7, True)

['Low', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Low', 'Region', 'Brandenburg', ',', 'Saxony']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Low', 'Language', 'codes', 'ISO', '639-2', 'ger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Low', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Bartangi', 'Spoken', 'in', 'Afghanistan', ',', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Bartangi', 'Writing', 'system', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bartangi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Bartangi', 'Language', 'codes', 'ISO', '639-3', 'sgl-bar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mesmes', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mesmes', 'Total', 'speakers', '0']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Mesmes', 'Language', 'codes', 'ISO', '639-3', 'mys']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Qimant', 'Dialects', 'Qwara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Qimant', 'Dialects', 'Kayla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Qimant', 'Region', 'Amhara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Qimant', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Qimant', 'Language', 'codes', 'ISO', '639-3', 'ahg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Greenlandic', 'Language', 'codes', 'ISO', '639-3', 'non-gre', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'non-gre', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)
```

```
['Yahi', 'Total', 'speakers', 'Extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Yahi', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Yahi', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Yahi', 'Language', 'codes', 'ISO', '639-3', 'ynn']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nukak', 'Dialects', 'Kakwa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Nukak', 'Region', 'Department', 'of', 'Guaviare', ',', 'Amazon', 'Basin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Department', 'of', 'Guaviare', ',', 'Amazon', 'Basin']], 8, True)

['Nukak', 'Spoken', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nukak', 'Regulated', 'by', 'Instituto', 'Colombiano', 'de', 'Antropolog', '\xc3', '\xad', 'a', 'e', 'Historia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Instituto', 'Colombiano', 'de', 'Antropolog', '\xc3', '\xad', 'a', 'e', 'Historia']], 12, True)

['Nukak', 'Language', 'codes', 'ISO', '639-3', 'mbr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kawaiisu', 'Region', 'California']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kawaiisu', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Kawaiisu', 'Pronunciation', 'n', '\xc9', '\xa8', 'w', '\xc9', '\xa8', '\xca', '\x94', 'abi', '\xc9', '\xa3', 'id', '\xc9', '\xa8', ',', 'n', '\xc9', '\xa8', 'w', '\xc9', '\xa8', '\xca', '\x94', 'abi', '\xc9', '\xa3', 'ip', '\xc9', '\xa8']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'n', '\xc9', '\xa8', 'w', '\xc9', '\xa8', '\xca', '\x94', 'abi', '\xc9', '\xa3', 'id', '\xc9', '\xa8', ',', 'n', '\xc9', '\xa8', 'w', '\xc9', '\xa8', '\xca', '\x94', 'abi', '\xc9', '\xa3', 'ip', '\xc9', '\xa8']], 31, True)

['Kawaiisu', 'Language', 'codes', 'ISO', '639-3', 'xaw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Bonda', 'Spoken', 'in', 'India']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Bonda', 'Language', 'codes', 'ISO', '639-3', 'bfb']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Iwam', 'Spoken', 'in', 'Sandaun', 'Province']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Iwam', 'Total', 'speakers', '430']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Iwam', 'Language', 'codes', 'ISO', '639-3', 'ayq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Gataq', 'Spoken', 'in', 'India']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Gataq', 'Language', 'codes', 'ISO', '639-3', 'gaq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Cemhu\xc3\xae', 'Total', 'speakers', '2000']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Cemhu\xc3\xae', 'Spoken', 'in', 'New', 'Caledonia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Cemhu\xc3\xae', 'Region', 'Touho', ':', 'east', 'coast', 'from',
'Congouma', 'to', 'Wagap', 'and', 'inland', 'valleys']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Touho', ':',
'east', 'coast', 'from', 'Congouma', 'to', 'Wagap', 'and', 'inland',
'valleys']]], 13, True)
```

```
['Cemhu\xc3\xae', 'Language', 'codes', 'ISO', '639-3', 'cam']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

['Kpala', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Kpala', 'Language', 'codes', 'ISO', '639-3', 'kpl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sarawak', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sarawak', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Sarawak', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Sarawak', 'Language', 'codes', 'ISO', '639-3', 'zlm-sar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cha\xe2\x80\x99', 'Spoken', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cha\xe2\x80\x99', 'Language', 'codes', 'ISO', '639-3', 'cbi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mara', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mara', 'Spoken', 'in', 'Mizoram', ',', 'India', ';', 'Chin', 'State',  
';', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Mizoram', ',', 'India', ';', 'Chin', 'State', ';', 'Type: ANY']]], 11,  
True)

['Mara', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Mara', 'Pronunciation', 'm', '\xcc', '\xa5', 'ara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm', '\xcc', '\xa5', 'a']]], 6, True)

['Mara', 'Official', 'language', 'in', 'none', ',', 'recognized', 'as', 'a', 'minority', 'language', 'in', 'Mizoram']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'none', ',', 'recognized', 'as', 'a', 'minority', 'language', 'in', 'Mizoram']]], 13, True)

['Mara', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Mara', 'Language', 'codes', 'ISO', '639-3', 'mrh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wellemedan', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wellemedan', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wellemedan', 'Region', 'Sahara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wellemedan', 'Language', 'codes', 'ISO', '639-3', 'ttq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Burji', 'Total', 'speakers', '46', ',', '419', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '46'], ['NUM', '419']], 8, True)

['Burji', 'Total', 'speakers', '10', ',', '400', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '10'], ['NUM', '400']], 8, True)

['Burji', 'Spoken', 'in', 'Ethiopia', ',', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Burji', 'Writing', 'system', 'Ethiopic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Burji', 'Language', 'codes', 'ISO', '639-3', 'bjj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gedeo', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gedeo', 'Region', 'Central', 'highland', 'area']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gedeo', 'Language', 'codes', 'ISO', '639-3', 'drs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dirasha', 'Total', 'speakers', '49', ',', '927', 'of', 'Gidole', 'dialect']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'of', 'Gidole', 'dialect']], ['NUM', '49'], ['NUM', '927']], 9, True)

['Dirasha', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dirasha', 'Region', 'Omo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dirasha', 'Region', 'in', 'the', 'hills', 'west', 'of', 'Lake', 'Chamo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'in', 'the', 'hills', 'west', 'of', 'Lake', 'Chamo']]], 9, True)

['Dirasha', 'Region', 'around', 'Gidole', 'town']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dirasha', 'Language', 'codes', 'ISO', '639-3', 'gdl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nambikwara', 'Spoken', 'in', 'Mato', 'Grosso', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mato',  
'Grosso', ',', 'Brazil']]], 7, True)

['Nambikwara', 'Language', 'codes', 'ISO', '639-3', 'nab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Siwu', 'Dialects', 'Akpafu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Siwu', 'Dialects', 'Lolobi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Siwu', 'Region', 'Volta', ',', 'north', 'of', 'Hohoe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
,',', 'Type: ANY', 'of', 'Type: ANY']]], 7, True)

['Siwu', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Siwu', 'Language', 'codes', 'ISO', '639-3', 'akp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wolio', 'Spoken', 'in', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Wolio', 'Total', 'speakers', '65', ',', '000', 'plus', 'L2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'plus', 'L2']], ['NUM', '65'], ['NUM', '000']], 8, True)

['Wolio', 'Language', 'codes', 'ISO', '639-3', 'wlo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Moronene', 'Spoken', 'in', 'Indonesia']:



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Moronene', 'Region', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Moronene', 'Dialects', 'Wita', 'Ea', '(', 'Rumbia', ',', 'Poleang', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Wita', 'Ea', '\\(', 'Rumbia', ',', 'Poleang', '\\)']]], 9, True)
```

```
['Moronene', 'Dialects', 'Tokotu', '\\xe2', '\\x80', '\\x99', 'a', '(', 'Kabaena', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Tokotu', '\\xe2', '\\x80', '\\x99', 'a', '\\(', 'Kabaena', '\\)']]], 10, True)
```

```
['Moronene', 'Language', 'codes', 'ISO', '639-3', 'mgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Umatilla', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Umatilla', 'Region', 'Oregon', ',', 'Washington']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Umatilla', 'Language', 'codes', 'ISO', '639-3', 'uma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Prekmurian', 'Spoken', 'in', 'Slovenia', ',', 'Hungary', 'and', 'emigrant', 'groups', 'in', 'various', 'countries']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Slovenia', ',', 'Hungary', 'and', 'emigrant', 'groups', 'in', 'various', 'countries']]], 12, True)
```

```
['Prekmurian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)
```

```
['Prekmurian', 'Language', 'codes', 'ISO', '639-3', 'slv-pre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ollari', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ollari', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ollari', 'Language', 'codes', 'ISO', '639-3', 'gau', '\xe2', '\x80', '\x93', 'Kondekor', '(', 'Gadaba', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gau', '\xe2', '\x80', '\x93', 'Kondekor', '\\(', 'Gadaba', '\\)']], ['NUM', '639-3']], 13, True)

['Thuringian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Thuringian', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Thuringian', 'Region', 'Thuringia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Thuringian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Late', 'Spoken', 'in', 'Japan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Late', 'Writing', 'system', 'Hiragana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Late', 'Writing', 'system', 'Katakana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Late', 'Writing', 'system', 'and', 'Han']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Late', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Dawera-Daweloor', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dawera-Daweloor', 'Region', 'Maluku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dawera-Daweloor', 'Language', 'codes', 'ISO', '639-3', 'ddw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngaju', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngaju', 'Region', 'Kalimantan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ngaju', 'Language', 'codes', 'ISO', '639-3', 'nij']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Melanau', 'Spoken', 'in', 'Malaysia', ',', 'Brunei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Melanau', 'Language', 'codes', 'ISO', '639-3', 'mel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buyang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Buyang', 'Region', 'Yunnan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Buyang', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buyang', 'Language', 'codes', 'ISO', '639-3', 'yln', '\xe2', '\x80',  
'\x93', 'Langnian', 'Buyang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'yln', '\xe2', '\x80', '\x93', 'Langnian', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Buyang', 'Language', 'codes', 'ISO', '639-3', 'yha', '\xe2', '\x80',  
'\x93', 'Baha', 'Buyang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'yha', '\xe2', '\x80', '\x93', 'Baha', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Phula', 'Spoken', 'in', 'China', ',', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Phula', 'Language', 'codes', 'ISO', '639-3', 'phh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saintongeais', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Saintongeais', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saintongeais', 'Region', 'Charente']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Saintongeais', 'Region', 'Charente-Maritime']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Saintongeais', 'Region', 'Northern', 'Gironde']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Saintongeais', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',

```

'\x93', 'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Saintongeais', 'Language', 'codes', 'ISO', '639-3', 'fra-san']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Poitevin', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)

['Poitevin', 'Spoken', 'in', 'France']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Poitevin', 'Region', 'Poitou', ',', 'Vend', '\xc3', '\xa9', 'e']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Poitou', ',',
'Vend', '\xc3', '\xa9', 'e']]], 8, True)

['Poitevin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Poitevin', 'Language', 'codes', 'ISO', '639-3', 'fra-poi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Poitevin', 'Language', 'codes', 'Linguasphere', '51-AAA-ha']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Filipino', 'Signed', 'in', 'Philippines']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Filipino', 'Total', 'signers', '(', 'unknown', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\
(', 'unknown', '\\)']]], 6, True)

['Nguon', 'Region', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

```
['Nguon', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nguon', 'Language', 'codes', 'ISO', '639-3', 'nuo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nguon', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of', 'Qu',  
'\xe1', '\xba', '\xa3', 'ng', 'B', '\xc3', '\xac', 'nh', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Location', 'of', 'Qu', '\xe1', '\xba', '\xa3', 'ng', 'B',  
'\xc3', '\xac', 'nh', 'Province']], ['NUM', '639-3']], 17, True)
```

```
['Laghuu', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Laghuu', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Laghuu', 'Language', 'codes', 'ISO', '639-3', 'lgh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Biak', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Biak', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Biak', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Biak', 'Language', 'codes', 'ISO', '639-3', 'bhw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taos', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```

['Taos', 'Language', 'codes', 'ISO', '639-3', 'twf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bugan', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bugan', 'Region', 'Yunnan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Bugan', 'Language', 'codes', 'ISO', '639-3', 'bbh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Arin', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Arin', 'Region', 'Yenisei', 'River']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Arin', 'Language', 'codes', 'ISO', '639-3', 'xrn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Finnish', 'Spoken', 'in', 'Finland', ',', 'Sweden']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Finnish', 'Language', 'codes', 'ISO', '639-2', 'rom']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Finnish', 'Language', 'codes', 'ISO', '639-3', 'rmf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Afade', 'Spoken', 'in', 'Nigeria', ',', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Afade', 'Region', 'Borno', 'State']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Afade', 'Region', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Afade', 'Language', 'codes', 'ISO', '639-3', 'aal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lagwan', 'Spoken', 'in', 'Cameroon', ',', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Lagwan', 'Language', 'codes', 'ISO', '639-3', 'kot']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Malgbe', 'Spoken', 'in', 'Cameroon', ',', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Malgbe', 'Language', 'codes', 'ISO', '639-3', 'mxf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Maslam', 'Total', 'speakers', '723']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maslam', 'Spoken', 'in', 'Cameroon', ',', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Maslam', 'Language', 'codes', 'ISO', '639-3', 'msv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mpade', 'Spoken', 'in', 'Cameroon', ',', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Mpade', 'Language', 'codes', 'ISO', '639-3', 'mpi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)



['Mser', 'Spoken', 'in', 'Cameroon', ',', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Mser', 'Language', 'codes', 'ISO', '639-3', 'kqx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cakfem-Mushere', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cakfem-Mushere', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cakfem-Mushere', 'Language', 'codes', 'ISO', '639-3', 'cky']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Jorto', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jorto', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jorto', 'Language', 'codes', 'ISO', '639-3', 'jrt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kofyar', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kofyar', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kofyar', 'Language', 'codes', 'ISO', '639-3', 'kwl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Miship', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Miship', 'Region', 'Plateau', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Miship', 'Language', 'codes', 'ISO', '639-3', 'mjs']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mwaghavul', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mwaghavul', 'Region', 'Plateau', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mwaghavul', 'Language', 'codes', 'ISO', '639-3', 'sur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngas', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngas', 'Region', 'Plateau', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ngas', 'Language', 'codes', 'ISO', '639-3', 'anc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngas', 'Language', 'codes', 'ISO', '639-3', 'Ethnic', 'territories', '(', 'tan', ')', 'of', 'the', 'Ngas-speaking', 'people', '(', 'Angas', ')', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Ethnic', 'territories', '\\\\(', 'tan', '\\\\)', 'of', 'the', 'Ngas-speaking', 'people', '\\\\(', 'Angas', '\\\\)', 'in', 'Nigeria']]], ['NUM', '639-3']], 19, True)

['Koenom', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Koenom', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Koenom', 'Language', 'codes', 'ISO', '639-3', 'kcs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Montol', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Montol', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Montol', 'Language', 'codes', 'ISO', '639-3', 'mtl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pyapun', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Pyapun', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Pyapun', 'Language', 'codes', 'ISO', '639-3', 'pcw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tal', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Tal', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Tal', 'Language', 'codes', 'ISO', '639-3', 'tal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ajawa', 'Total', 'speakers', 'Extinct']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ajawa', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ajawa', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ajawa', 'Language', 'codes', 'ISO', '639-3', 'ajw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ciwogai', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ciwogai', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ciwogai', 'Language', 'codes', 'ISO', '639-3', 'tgd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Diri', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Diri', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Diri', 'Language', 'codes', 'ISO', '639-3', 'dwa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kariya', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kariya', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kariya', 'Language', 'codes', 'ISO', '639-3', 'kil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mburku', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Mburku', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Mburku', 'Language', 'codes', 'ISO', '639-3', 'bbt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Miya', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Miya', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Miya', 'Language', 'codes', 'ISO', '639-3', 'mkf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pa'a', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pa'a', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pa'a', 'Language', 'codes', 'ISO', '639-3', 'pqa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Siri', 'Total', 'speakers', 'a', 'few', 'older', 'adults']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a',  
'few', 'older', 'adults']], 7, True)

['Siri', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Siri', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Siri', 'Language', 'codes', 'ISO', '639-3', 'sir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warji', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Warji', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Warji', 'Language', 'codes', 'ISO', '639-3', 'wji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piruris', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Piruris', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Piruris', 'Language', 'codes', 'ISO', '639-3', 'twf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auyokawa', 'Total', 'speakers', 'Extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Auyokawa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Auyokawa', 'Region', 'Jigawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Auyokawa', 'Language', 'codes', 'ISO', '639-3', 'auo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Duwai', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Duwai', 'Region', 'Yobe', 'State', ',', 'Kano', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Yobe',
'State', ',', 'Kano', 'State']]], 7, True)

['Duwai', 'Language', 'codes', 'ISO', '639-3', 'dbp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Alaba-K'abeena", 'Total', 'speakers', '226', ',', '604', 'of', 'Alaba']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'of', 'Type: ANY']], ['NUM', '226'], ['NUM', '604']], 8,
True)

["Alaba-K'abeena", 'Total', 'speakers', '51', ',', '333', 'of', 'Qebena']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'of', 'Type: ANY']], ['NUM', '51'], ['NUM', '333']], 8,
True)

["Alaba-K'abeena", 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Alaba-K'abeena", 'Region', 'Rift', 'Valley', 'southwest', 'of', 'Lake',
'Shala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rift',
'Valley', 'southwest', 'of', 'Lake', 'Shala']]], 8, True)

["Alaba-K'abeena", 'Language', 'codes', 'ISO', '639-3', 'alw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kambaata', 'Total', 'speakers', '30', ',', '098', 'of', 'Donga',
'dialect']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'of', 'Donga', 'dialect']], ['NUM', '30'], ['NUM',
'098']], 9, True)

```

```
['Kambaata', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kambaata', 'Region', 'southwest', 'Gurage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kambaata', 'Region', 'Kambaata']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Kambaata', 'Region', 'Hadiyya', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kambaata', 'Writing', 'system', 'Ethiopic', '/', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Kambaata', 'Language', 'codes', 'ISO', '639-3', 'ktb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Southern', 'Region', 'New', 'Mexico', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Southern', 'Region', 'Tiwa', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'tix']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arbore', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arbore', 'Region', 'extreme', 'southwest', ',', 'Omo', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'extreme',
```



'southwest', ',', 'Omo', 'region']]], 7, True)

['Arbore', 'Language', 'codes', 'ISO', '639-3', 'arv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baiso', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Baiso', 'Region', 'in', 'region', 'of', 'Lake', 'Abaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'in', 'region',  
'of', 'Lake', 'Abaya']], 7, True)

['Baiso', 'Language', 'codes', 'ISO', '639-3', 'bsw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Daasanach', 'Total', 'speakers', '47', ',', '986', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '47'], ['NUM', '986']], 8,  
True)

['Daasanach', 'Total', 'speakers', '8', ',', '000', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '8'], ['NUM', '000']], 8,  
True)

['Daasanach', 'Spoken', 'in', 'Ethiopia', ',', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Daasanach', 'Language', 'codes', 'ISO', '639-3', 'dsh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['El', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['El', 'Region', 'Lake', 'Turkana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['El', 'Language', 'codes', 'ISO', '639-3', 'elo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alagwa', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Alagwa', 'Region', 'Dodoma', 'region']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Alagwa', 'Language', 'codes', 'ISO', '639-3', 'wbj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Burunge', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Burunge', 'Region', 'Dodoma']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Burunge', 'Language', 'codes', 'ISO', '639-3', 'bds']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gorowa', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gorowa', 'Region', 'Dodoma', 'and', 'Arusha', 's']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Dodoma', 'and', 'Arusha', 's']], 6, True)

['Gorowa', 'Language', 'codes', 'ISO', '639-3', 'gow']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iraqw', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iraqw', 'Region', 'Arusha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

```

['Iraqw', 'Language', 'codes', 'ISO', '639-3', 'irk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Kw'adza", 'Total', 'speakers', 'Extinct']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Kw'adza", 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Kw'adza", 'Region', 'Mbulu', 'District']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Kw'adza", 'Language', 'codes', 'ISO', '639-3', 'wka']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dime', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Dime', 'Language', 'codes', 'ISO', '639-3', 'dim']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karo', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karo', 'Region', 'South', 'Omo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Karo', 'Language', 'codes', 'ISO', '639-3', 'kxh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bana', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bana', 'Region', 'Far', 'North', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',

```

'Type: ANY', 'Type: ANY']]], 5, True)

['Bana', 'Language', 'codes', 'ISO', '639-3', 'bcw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hya', 'Spoken', 'in', 'Cameroon', ',', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Hya', 'Language', 'codes', 'ISO', '639-3', 'hya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamwe', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kamwe', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kamwe', 'Language', 'codes', 'ISO', '639-3', 'hig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Psikye', 'Spoken', 'in', 'Cameroon', ',', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Psikye', 'Language', 'codes', 'ISO', '639-3', 'kvj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baldemu', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Baldemu', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Baldemu', 'Language', 'codes', 'ISO', '639-3', 'bdn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Cuvok', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cuvok', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Cuvok', 'Language', 'codes', 'ISO', '639-3', 'cuv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dugwor', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dugwor', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Dugwor', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Dugwor', 'Language', 'codes', 'ISO', '639-3', 'dme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mada', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mada', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mada', 'Language', 'codes', 'ISO', '639-3', 'mxu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mafa', 'Spoken', 'in', 'Cameroon', ',', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Mafa', 'Language', 'codes', 'ISO', '639-3', 'maf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Matal', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Matal', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Matal', 'Language', 'codes', 'ISO', '639-3', 'mfh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mbuko', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mbuko', 'Region', 'Far', 'North']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mbuko', 'Language', 'codes', 'ISO', '639-3', 'mqb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mefe', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mefe', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Mefe', 'Language', 'codes', 'ISO', '639-3', 'mfj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Meray', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Meray', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

```
['Merey', 'Language', 'codes', 'ISO', '639-3', 'meq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mofu-Gudur', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Mofu-Gudur', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Mofu-Gudur', 'Language', 'codes', 'ISO', '639-3', 'mif']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Moloko', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Moloko', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Moloko', 'Language', 'codes', 'ISO', '639-3', 'mlw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Muyang', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Muyang', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Muyang', 'Language', 'codes', 'ISO', '639-3', 'muy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['North', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['North', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'gis']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['North', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'mfk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'giz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vame', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Vame', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Vame', 'Language', 'codes', 'ISO', '639-3', 'mlr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wuzlam', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



```
['Wuzlam', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Wuzlam', 'Language', 'codes', 'ISO', '639-3', 'udl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bacama', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bacama', 'Region', 'Adamawa', 'State', ',', 'Kaduna', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Adamawa',  
'State', ',', 'Kaduna', 'State']]], 7, True)
```

```
['Bacama', 'Language', 'codes', 'ISO', '639-3', 'bcy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bata', 'Spoken', 'in', 'Nigeria', ',', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Bata', 'Language', 'codes', 'ISO', '639-3', 'bta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bata', 'Language', 'codes', 'ISO', '639-3', 'Ethnic', 'territories',  
'of', 'the', 'Bata-speaking', 'people', '(', 'Batta', ')', 'in', 'Nigeria',  
,',', 'in', 'blue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Ethnic', 'territories', 'of', 'the', 'Bata-speaking',  
'people', '\\\\(', 'Batta', '\\\\)', 'in', 'Nigeria', ',', 'in', 'blue']]],  
['NUM', '639-3']], 19, True)
```

```
['Fali', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Fali', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Fali', 'Language', 'codes', 'ISO', '639-3', 'fli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gude', 'Spoken', 'in', 'Nigeria', ',', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Gude', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gude', 'Region', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gude', 'Language', 'codes', 'ISO', '639-3', 'gde']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gudu', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gudu', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gudu', 'Language', 'codes', 'ISO', '639-3', 'gdu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Holma', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Holma', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Holma', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Holma', 'Language', 'codes', 'ISO', '639-3', 'hod']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Jimi', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jimi', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Jimi', 'Language', 'codes', 'ISO', '639-3', 'jim']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngwaba', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ngwaba', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ngwaba', 'Language', 'codes', 'ISO', '639-3', 'ngw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nzanyi', 'Spoken', 'in', 'Nigeria', ',', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Nzanyi', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nzanyi', 'Language', 'codes', 'ISO', '639-3', 'nja']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sharwa', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sharwa', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sharwa', 'Language', 'codes', 'ISO', '639-3', 'swq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsuva', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsuva', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tsuva', 'Language', 'codes', 'ISO', '639-3', 'tsh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buduma', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Buduma', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Buduma', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Buduma', 'Language', 'codes', 'ISO', '639-3', 'bdm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jilbe', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jilbe', 'Region', 'Borno', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jilbe', 'Language', 'codes', 'ISO', '639-3', 'jie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngizim', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Ngizim', 'Region', 'Yobe', 'State']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ngizim', 'Language', 'codes', 'ISO', '639-3', 'ngi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Teshenawa', 'Total', 'speakers', 'extinct']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Teshenawa', 'Spoken', 'in', 'Nigeria']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Teshenawa', 'Region', 'Jigawa', 'State']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Teshenawa', 'Language', 'codes', 'ISO', '639-3', 'twc']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dr\xc3\xa8ents,', 'Spoken', 'in', 'Netherlands']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Dr\xc3\xa8ents,', 'Region', 'Drenthe', 'and', 'a', 'strip', 'of',
'Overijssel', 'just', 'south', 'of', 'Drenthe']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Drenthe',
'and', 'a', 'strip', 'of', 'Overijssel', 'just', 'south', 'of',
'Drenthe']]], 12, True)
```

```
['Dr\xc3\xa8ents,', 'Regulated', 'by', 'Provincial', 'States', 'of',
'Drenthe', '(', 'spelling', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Provincial', 'States', 'of', 'Drenthe', '\\(', 'spelling', '\\)']]], 10,
True)
```

```
['Dr\xc3\xa8ents,', 'Language', 'codes', 'ISO', '639-3', 'drt']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chara', 'Spoken', 'in', 'Ethiopia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Chara', 'Writing', 'system', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chara', 'Language', 'codes', 'ISO', '639-3', 'cra']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Buwal', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Buwal', 'Region', 'Far', 'North', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Buwal', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Buwal', 'Language', 'codes', 'ISO', '639-3', 'bhs']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Daba', 'Spoken', 'in', 'Cameroon', ',', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Daba', 'Language', 'codes', 'ISO', '639-3', 'dbq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gavar', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gavar', 'Region', 'Far', 'North', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gavar', 'Language', 'codes', 'ISO', '639-3', 'gou']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Mazagway', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mazagway', 'Language', 'codes', 'ISO', '639-3', 'dkx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Mbedam', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mbedam', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mbedam', 'Language', 'codes', 'ISO', '639-3', 'xmd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Mina', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mina', 'Region', 'Far', 'North', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mina', 'Language', 'codes', 'ISO', '639-3', 'hna']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bidiyo', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bidiyo', 'Region', 'south', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bidiyo', 'Language', 'codes', 'ISO', '639-3', 'bid']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Dangal\xc3\xa9at', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Dangal\xc3\xa9at', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Dangal\xc3\xa9at', 'Language', 'codes', 'ISO', '639-3', 'daa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Jonkor', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jonkor', 'Region', 'Salamat', 'Prefecture']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jonkor', 'Language', 'codes', 'ISO', '639-3', 'jeu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mabire', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mabire', 'Region', 'Gu', '\xc3', '\xa9', 'ra', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Gu', '\xc3', '\xa9', 'ra', 'Province']]], 7, True)

['Mabire', 'Language', 'codes', 'ISO', '639-3', 'muj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mawa', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mawa', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mawa', 'Language', 'codes', 'ISO', '639-3', 'mcw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)



```
['Migaama', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Migaama', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Migaama', 'Language', 'codes', 'ISO', '639-3', 'mmy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mogum', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mogum', 'Region', 'south', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mogum', 'Language', 'codes', 'ISO', '639-3', 'mou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ubi', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ubi', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Ubi', 'Language', 'codes', 'ISO', '639-3', 'ubi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Official', 'Language', 'codes', 'ISO', '639-2', 'arc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Official', 'Language', 'codes', 'ISO', '639-3', 'arc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'oar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Putai', 'Total', 'speakers', '50']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Putai', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Putai', 'Region', 'Borno', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Putai', 'Language', 'codes', 'ISO', '639-3', 'mfl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buso', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Buso', 'Region', 'west']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Buso', 'Language', 'codes', 'ISO', '639-3', 'bso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Boor', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Boor', 'Region', 'south']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Boor', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Boor', 'Language', 'codes', 'ISO', '639-3', 'bvf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Miltu', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Miltu', 'Region', 'southwest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Miltu', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Miltu', 'Language', 'codes', 'ISO', '639-3', 'mlj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sarua', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sarua', 'Region', 'southwest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Sarua', 'Language', 'codes', 'ISO', '639-3', 'swy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Birgit', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Birgit', 'Region', 'southeast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Birgit', 'Language', 'codes', 'ISO', '639-3', 'btf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kajakse', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kajakse', 'Region', 'east']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY'']], 3, True)

['Kajakse', 'Language', 'codes', 'ISO', '639-3', 'ckq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Masmaje', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Masmaje', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Masmaje', 'Language', 'codes', 'ISO', '639-3', 'mes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mubi', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Mubi', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Mubi', 'Language', 'codes', 'ISO', '639-3', 'mub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Toram', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Toram', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Toram', 'Language', 'codes', 'ISO', '639-3', 'trj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zirenkel', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Zirenkel', 'Language', 'codes', 'ISO', '639-3', 'zrn']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Barein', 'Spoken', 'in', 'Chad']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Barein', 'Region', 'south', 'central']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Barein', 'Language', 'codes', 'ISO', '639-3', 'bva']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Saba', 'Spoken', 'in', 'Chad']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Saba', 'Region', 'south', 'central']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Saba', 'Language', 'codes', 'ISO', '639-3', 'saa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sokoro', 'Spoken', 'in', 'Chad']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Sokoro', 'Region', 'central']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Sokoro', 'Language', 'codes', 'ISO', '639-3', 'sok']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tamki', 'Spoken', 'in', 'Chad']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tamki', 'Region', 'central']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY'[]], 3, True)

['Tamki', 'Language', 'codes', 'ISO', '639-3', 'tax']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bure', 'Total', 'speakers', '500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Bure', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Bure', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Bure', 'Language', 'codes', 'ISO', '639-3', 'bvh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karakare', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Karakare', 'Region', 'Bauchi', 'State', ',', 'Yobe', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Bauchi',  
'State', ',', 'Yobe', 'State']], 7, True)

['Karakare', 'Language', 'codes', 'ISO', '639-3', 'kai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karakare', 'Language', 'codes', 'ISO', '639-3', 'Ethnic', 'territories',  
'(', 'pink', ')', 'of', 'the', 'Karekare-speaking', 'people', '(', 'Kare',  
)', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Ethnic', 'territories', '\\(', 'pink', '\\)', 'of', 'the',  
'Karekare-speaking', 'people', '\\(', 'Kare', '\\)', 'in', 'Nigeria']],  
['NUM', '639-3']], 19, True)

['Beele', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Beele', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Beele', 'Language', 'codes', 'ISO', '639-3', 'bxq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bole', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bole', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bole', 'Region', 'Gombe', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bole', 'Region', 'Yobe', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bole', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bole', 'Language', 'codes', 'ISO', '639-3', 'bol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Deno', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Deno', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Deno', 'Language', 'codes', 'ISO', '639-3', 'dbb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Galambu', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Galambu', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Galambu', 'Language', 'codes', 'ISO', '639-3', 'glo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gera', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gera', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gera', 'Language', 'codes', 'ISO', '639-3', 'gew']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Geruma', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Geruma', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Geruma', 'Language', 'codes', 'ISO', '639-3', 'gea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Giiwo', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Giiwo', 'Region', 'Bauchi', 'State']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Giiwo', 'Language', 'codes', 'ISO', '639-3', 'kks']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Kholok', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kholok', 'Region', 'Taraba', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kholok', 'Language', 'codes', 'ISO', '639-3', 'ktc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kubi', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kubi', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kubi', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kubi', 'Language', 'codes', 'ISO', '639-3', 'kof']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Maaka', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maaka', 'Region', 'Borno', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Maaka', 'Language', 'codes', 'ISO', '639-3', 'mew']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngamo', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ngamo', 'Region', 'Gombe', 'State', ',', 'Yobe', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Gombe',

'State', ',', 'Yobe', 'State']]], 7, True)

['Ngamo', 'Language', 'codes', 'ISO', '639-3', 'nbh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyam', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nyam', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nyam', 'Region', 'Taraba', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Nyam', 'Language', 'codes', 'ISO', '639-3', 'nmi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Boghom', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Boghom', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Boghom', 'Language', 'codes', 'ISO', '639-3', 'bux']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kir-Balar', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kir-Balar', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kir-Balar', 'Language', 'codes', 'ISO', '639-3', 'kkkr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mangas', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mangas', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mangas', 'Language', 'codes', 'ISO', '639-3', 'zns']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aweer', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aweer', 'Language', 'codes', 'ISO', '639-3', 'bob']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aweer', 'Language', 'codes', 'Linguasphere', '14-GAF-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Rendille', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Rendille', 'Region', 'Eastern', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Rendille', 'Language', 'codes', 'ISO', '639-3', 'rel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Basketo', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Basketo', 'Language', 'codes', 'ISO', '639-3', 'bst']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Male', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Male', 'Region', 'Omo', 'region', ',', 'southeast', 'of', 'Jinka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Omo',  
'region', ',', 'southeast', 'of', 'Jinka']]], 8, True)

['Male', 'Writing', 'system', 'Ethiopic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Male', 'Language', 'codes', 'ISO', '639-3', 'mdy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Boro', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Boro', 'Language', 'codes', 'ISO', '639-3', 'bwo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Algerian', 'Spoken', 'in', 'Algeria', ',', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Algerian', 'Region', 'southern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'southern']]],  
3, True)

['Algerian', 'Language', 'codes', 'ISO', '639-3', 'aao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sapo', 'Spoken', 'in', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sapo', 'Language', 'codes', 'ISO', '639-3', 'krn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Moose', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```

['Moose', 'Region', 'Ontario']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Moose', 'Language', 'codes', 'ISO', '639-3', 'crm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Moose', 'Language', 'codes', 'Linguasphere', '62-ADA-ae']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Dhofari', 'Spoken', 'in', 'Oman']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Dhofari', 'Writing', 'system', 'Arabic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Dhofari', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dhofari', 'Language', 'codes', 'ISO', '639-3', 'adf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yalunka', 'Spoken', 'in', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yalunka', 'Spoken', 'in', 'Mali']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yalunka', 'Spoken', 'in', 'Senegal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yalunka', 'Spoken', 'in', 'Sierra', 'Leone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Yalunka', 'Language', 'codes', 'ISO', '639-3', 'yal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bedawi', 'Spoken', 'in', 'Egypt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bedawi', 'Spoken', 'in', 'Israel']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bedawi', 'Spoken', 'in', 'Jordan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bedawi', 'Spoken', 'in', 'Palestine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bedawi', 'Spoken', 'in', 'Syria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bedawi', 'Writing', 'system', 'Arabic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Bedawi', 'Language', 'codes', 'ISO', '639-2', 'ara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bedawi', 'Language', 'codes', 'ISO', '639-3', 'avl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dewoin', 'Spoken', 'in', 'Liberia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dewoin', 'Language', 'codes', 'ISO', '639-3', 'dee']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gbii', 'Spoken', 'in', 'Liberia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Gbii', 'Language', 'codes', 'ISO', '639-3', 'ggb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tajuasohn', 'Spoken', 'in', 'Liberia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tajuasohn', 'Language', 'codes', 'ISO', '639-3', 'tja']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gliio-Oubi', 'Language', 'codes', 'ISO', '639-3', 'oub']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuwa', 'Spoken', 'in', 'Liberia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Kuwa', 'Language', 'codes', 'ISO', '639-3', 'blh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Omani', 'Spoken', 'in', 'Oman']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Omani', 'Region', 'Hajar', 'mountains', 'and', 'a', 'few', 'coastal',
'towns']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Hajar',
'mountains', 'and', 'a', 'few', 'coastal', 'towns']]], 9, True)

['Omani', 'Writing', 'system', 'Arabic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Omani', 'Language', 'codes', 'ISO', '639-3', 'acx']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sa`idi', 'Spoken', 'in', 'Egypt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sa`idi', 'Writing', 'system', 'Arabic', 'alphabet']:

```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sa`idi', 'Language', 'codes', 'ISO', '639-3', 'aec']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sanaani', 'Spoken', 'in', 'Yemen']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Sanaani', 'Region', 'north']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Sanaani', 'Writing', 'system', 'Arabic', 'alphabet']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Sanaani', 'Language', 'codes', 'ISO', '639-3', 'ayn']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
["Ta'izzi-Adeni", 'Spoken', 'in', 'Yemen', ',', 'Djibouti']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)
```

```
["Ta'izzi-Adeni", 'Region', 'south']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
["Ta'izzi-Adeni", 'Writing', 'system', 'Arabic', 'alphabet']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
["Ta'izzi-Adeni", 'Language', 'codes', 'ISO', '639-3', 'acq']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shihhi', 'Spoken', 'in', 'United', 'Arab', 'Emirates', ',', 'Oman']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'Arab', 'Emirates', ',', 'Oman']]], 8, True)
```

```
['Shihhi', 'Language', 'codes', 'ISO', '639-3', 'ssh']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```



['Judeo-Tripolitanian', 'Spoken', 'in', 'Israel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Judeo-Tripolitanian', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Judeo-Tripolitanian', 'Language', 'codes', 'ISO', '639-3', 'yud']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Judeo-Tunisian', 'Spoken', 'in', 'Israel', ',', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Judeo-Tunisian', 'Writing', 'system', 'Hebrew', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Judeo-Tunisian', 'Language', 'codes', 'ISO', '639-3', 'ajt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hoby\\xc3\\xb3t', 'Spoken', 'in', 'Oman', ',', 'Yemen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Hoby\\xc3\\xb3t', 'Language', 'codes', 'ISO', '639-3', 'hoh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shehri', 'Spoken', 'in', 'Oman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Shehri', 'Language', 'codes', 'ISO', '639-3', 'shv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cineni', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cineni', 'Region', 'Borno', 'State']:

[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Cineni', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Cineni', 'Language', 'codes', 'ISO', '639-3', 'cie']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dghwede', 'Spoken', 'in', 'Nigeria']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dghwede', 'Region', 'Borno', 'State']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dghwede', 'Language', 'codes', 'ISO', '639-3', 'dgh']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Glavda', 'Spoken', 'in', 'Nigeria', ',', 'Cameroon']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Glavda', 'Language', 'codes', 'ISO', '639-3', 'glw']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guduf-Gava', 'Spoken', 'in', 'Nigeria']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Guduf-Gava', 'Region', 'Borno', 'State']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Guduf-Gava', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Guduf-Gava', 'Language', 'codes', 'ISO', '639-3', 'gdf']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gvoko', 'Spoken', 'in', 'Nigeria', ',', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Gvoko', 'Language', 'codes', 'ISO', '639-3', 'ngs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Duhwa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Duhwa', 'Region', 'Nasarawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Duhwa', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Duhwa', 'Language', 'codes', 'ISO', '639-3', 'kbz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kulere', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kulere', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kulere', 'Language', 'codes', 'ISO', '639-3', 'kul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mundat', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mundat', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Mundat', 'Language', 'codes', 'ISO', '639-3', 'mmf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ron', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Ron', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Ron', 'Language', 'codes', 'ISO', '639-3', 'cla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sha', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Sha', 'Region', 'Plateau', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Sha', 'Language', 'codes', 'ISO', '639-3', 'scw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Carolina', 'Algonquian', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Algonquian', 'Total',  
'speakers', 'extinct']], 5, True)
```

```
['Carolina', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Carolina', 'Region', 'North', 'Carolina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Carolina', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Carolina', 'Language', 'codes', 'ISO', '639-3', 'pmk', '\xe2', '\x80',  
'\x93', 'Pamlico']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pmk', '\xe2', '\x80', '\x93', 'Pamlico']], ['NUM', '639-3']], 10, True)
```

```
['Mahican', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Mahican', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mahican', 'Region', 'New', 'York']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Mahican', 'Language', 'codes', 'ISO', '639-3', 'mjj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Piscataway', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Piscataway', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Piscataway', 'Region', 'Maryland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maryland']]], 3, True)
```

```
['Piscataway', 'Language', 'codes', 'ISO', '639-3', 'psy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Naskapi', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Naskapi', 'Region', 'Quebec', '', 'Labrador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Naskapi', 'Language', 'codes', 'ISO', '639-3', 'nsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Naskapi', 'Language', 'codes', 'Linguasphere', '62-ADA-ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Northern', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Region', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', 'cr1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northern', 'Language', 'codes', 'Linguasphere', '62-ADA-af']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Southern', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Southern', 'Region', 'Quebec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'crj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Language', 'codes', 'Linguasphere', '62-ADA-ag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Woods', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Woods', 'Region', 'Manitoba', ',', 'Saskatchewan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

```

['Woods', 'Language', 'codes', 'ISO', '639-3', 'cwd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Woods', 'Language', 'codes', 'Linguasphere', '62-ADA-ab']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)

['Central', 'Ojibwa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',
'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Central', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Central', 'Region', 'Ontario']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'ojc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chippewa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Chippewa', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Chippewa', 'Language', 'codes', 'ISO', '639-3', 'ciw']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Ojibwa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',
'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Eastern', 'Spoken', 'in', 'Canada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Eastern', 'Region', 'Ontario']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

```

ANY']]], 3, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'ojg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northwestern', 'Ojibwa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', '\xe2', '\x80', '\x93']], 7, True)

['Northwestern', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Northwestern', 'Region', 'Ontario', ',', 'Manitoba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Northwestern', 'Language', 'codes', 'ISO', '639-3', 'obj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Ojibwa', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', '\xe2', '\x80', '\x93']], 7, True)

['Western', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Western', 'Region', 'Manitoba', ',', 'Saskatchewan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'ojw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Malecite-Passamaquoddy', 'Spoken', 'in', 'Canada', ';', 'United',  
'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',  
';', 'United', 'States']], 7, True)

['Malecite-Passamaquoddy', 'Language', 'codes', 'ISO', '639-3', 'pqm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



```
['Mohegan-Montauk-Narragansett', 'Total', 'speakers', 'extinct']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mohegan-Montauk-Narragansett', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Mohegan-Montauk-Narragansett', 'Region', 'southern', 'New', 'England']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mohegan-Montauk-Narragansett', 'Language', 'codes', 'ISO', '639-3',
'mof']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Nanticoke', 'Total', 'speakers', 'extinct']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Nanticoke', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Nanticoke', 'Region', 'Delaware', ',', 'Maryland']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Nanticoke', 'Language', 'codes', 'ISO', '639-3', 'nnt']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Gros', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Gros', 'Region', 'Montana']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Gros', 'Language', 'codes', 'ISO', '639-3', 'ats']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Nawathinehena', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Nawathinehena', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nawathinehena', 'Language', 'codes', 'ISO', '639-3', 'nwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Daza', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Daza', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Daza', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Daza', 'Language', 'codes', 'ISO', '639-3', 'dzd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dass', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dass', 'Language', 'codes', 'ISO', '639-3', 'dot']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jimi', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jimi', 'Region', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jimi', 'Language', 'codes', 'ISO', '639-3', 'jmi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gawwada', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gawwada', 'Region', 'Dirashe', 'special', 'woreda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Gawwada', 'Region', 'Southern', 'Nations']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gawwada', 'Region', 'Nationalities']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Gawwada', 'Region', 'and', "People's"]:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gawwada', 'Writing', 'system', 'Ethiopic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gawwada', 'Language', 'codes', 'ISO', '639-3', 'gwd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tsamai', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tsamai', 'Region', 'Omo', 'River', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Tsamai', 'Language', 'codes', 'ISO', '639-3', 'tsb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dabarre', 'Spoken', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Dabarre', 'Language', 'codes', 'ISO', '639-3', 'dbr']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Garre', 'Spoken', 'in', 'Somalia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Garre', 'Region', 'southern']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'southern']]],
3, True)
```

```
['Garre', 'Language', 'codes', 'ISO', '639-3', 'gex']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jiiddu', 'Spoken', 'in', 'Somalia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Jiiddu', 'Language', 'codes', 'ISO', '639-3', 'jii']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tunni', 'Spoken', 'in', 'Somalia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tunni', 'Language', 'codes', 'ISO', '639-3', 'tqq']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Birgid', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Birgid', 'Spoken', 'in', 'Sudan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Birgid', 'Region', 'west']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Birgid', 'Language', 'codes', 'ISO', '639-3', 'brk']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mukulu', 'Spoken', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mukulu', 'Region', 'central']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Mukulu', 'Language', 'codes', 'ISO', '639-3', 'moz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dorze', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dorze', 'Language', 'codes', 'ISO', '639-3', 'doz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gamo-Gofa-Dawro', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gamo-Gofa-Dawro', 'Region', 'Omo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Gamo-Gofa-Dawro', 'Writing', 'system', 'Ethiopic', '/', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 6, True)

['Gamo-Gofa-Dawro', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gamo-Gofa-Dawro', 'Language', 'codes', 'ISO', '639-3', ':', 'gmv', '\xe2', '\x80', '\x93', 'Gamo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'gmv', '\xe2', '\x80', '\x93', 'Gamo']], ['NUM', '639-3']], 11, True)

['Gamo-Gofa-Dawro', 'Language', 'codes', 'ISO', '639-3', 'gof', '\xe2', '\x80', '\x93', 'Gofa']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gof', '\xe2', '\x80', '\x93', 'Gofa']], ['NUM', '639-3']], 10, True)

['Gamo-Gofa-Dawro', 'Language', 'codes', 'ISO', '639-3', 'dwr', '\xe2', '\x80', '\x93', 'Dawro']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'dwr', '\xe2', '\x80', '\x93', 'Dawro']], ['NUM', '639-3']], 10, True)

['Melo', 'Spoken', 'in', 'Ethiopia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Melo', 'Region', 'Gamo', 'Gofa', 'Zone']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Melo', 'Language', 'codes', 'ISO', '639-3', 'mfx']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oyda', 'Spoken', 'in', 'Ethiopia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oyda', 'Region', 'Northwest', 'Omo']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oyda', 'Language', 'codes', 'ISO', '639-3', 'oyd']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Middle', 'Spoken', 'in', 'Mongolia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Middle', 'Spoken', 'in', 'China']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Middle', 'Spoken', 'in', 'Russia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Middle', 'Language', 'codes', 'ISO', '639-3', 'xng']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Maipure', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Maipure', 'Region', 'Orinoco']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Maipure', 'Language', 'codes', 'ISO', '639-3', 'qij']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-3', 'uzn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Spoken', 'in', 'Afghanistan', ',', 'Turkey']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'uzs']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Northern', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Northern', 'Region', 'Altai', 'Mountains']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Northern', 'Language', 'codes', 'ISO', '639-3', 'atv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Southern', 'Region', 'Gorno-Altai', 'Ao', 'mountains']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Southern', 'Language', 'codes', 'ISO', '639-2', 'alt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Southern', 'Language', 'codes', 'ISO', '639-3', 'alt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Azerbaijan', 'and', 'neighboring', 'Caucasian',  
'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Azerbaijan', 'and', 'neighboring', 'Caucasian', 'region']]], 8, True)

['North', 'Total', 'speakers', '7', ',', '059', ',', '529', 'worldwide',  
'including', '4', ',', '000', ',', '000', 'monolinguals']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', ',', 'NUM', 'worldwide', 'including', 'NUM', ',', 'NUM',  
',', 'NUM', 'monolinguals']]], ['NUM', '7'], ['NUM', '059'], ['NUM', '529'],  
['NUM', '4'], ['NUM', '000'], ['NUM', '000']], 16, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'azj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['South', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Spoken', 'in', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Region', 'northwestern', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



['South', 'Region', 'northern', 'Iraq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Region', 'eastern', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'azb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Salchuq', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Salchuq', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Salchuq', 'Language', 'codes', 'ISO', '639-3', 'slq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Achterhooks', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Achterhooks', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Achterhooks', 'Region', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Achterhooks', 'Language', 'codes', 'ISO', '639-3', 'act']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amto', 'Region', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Amto', 'Language', 'codes', 'ISO', '639-3', 'amt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

['ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Siawi', 'Region', 'Sandaun', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Siawi', 'Language', 'codes', 'ISO', '639-3', 'mmp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fyer', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fyer', 'Region', 'Plateau', 'State']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fyer', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Fyer', 'Language', 'codes', 'ISO', '639-3', 'fie']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tambas', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tambas', 'Region', 'Plateau', 'State']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tambas', 'Language', 'codes', 'ISO', '639-3', 'tdk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mundari', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mundari', 'Spoken', 'in', 'Bangladesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mundari', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mundari', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mundari', 'Language', 'codes', 'ISO', '639-3', ':', 'unr', '\xe2', '\x80', '\x93', 'Mundari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'unr', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Mundari', 'Language', 'codes', 'ISO', '639-3', 'unx', '\xe2', '\x80', '\x93', 'Munda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'unx', '\xe2', '\x80', '\x93', 'Munda']], ['NUM', '639-3']], 10, True)

['Erromangan', 'Spoken', 'in', 'Southern', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Erromangan', 'Total', 'speakers', '1900']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Erromangan', 'Language', 'codes', 'ISO', '639-3', 'erg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Auk\xc5\xa1taitian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Auk\xc5\xa1taitian', 'Region', 'Auk', '\xc5', '\xa1', 'taitija', ',', 'Dz', '\xc5', '\xab', 'kija', 'and', 'Suvalkija']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Auk', '\xc5', '\xa1', 'taitija', ',', 'Dz', '\xc5', '\xab', 'kija', 'and', 'Suvalkija']], 13, True)

['Auk\xc5\xa1taitian', 'Spoken', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Auk\\xc5\\xa1\\taitian', 'Writing', 'system', 'Roman', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Auk\\xc5\\xa1\\taitian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',  
'\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Rodriguan', 'Spoken', 'in', 'Rodrigues']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Rodriguan', 'Language', 'codes', 'ISO', '639-3', 'mfe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Rodriguan', 'Language', 'codes', 'Linguasphere', '51-AAC-ced']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Chagossian', 'Spoken', 'in', 'Mauritius', ',', 'Seychelles']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Chagossian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',  
'\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Agalega', 'Spoken', 'in', 'Agalega', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Agalega', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
["Ma'di", 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
["Ma'di", 'Spoken', 'in', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```

["Ma'di", 'Language', 'codes', 'ISO', '639-3', 'mhi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Batak', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Batak', 'Region', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Batak', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Batak',
'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Batak', 'Official', 'language', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Batak', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Batak', 'Language', 'codes', 'ISO', '639-3', 'btx']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sunam', 'Spoken', 'in', 'Himachal', 'Pradesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Sunam', 'Language', 'codes', 'ISO', '639-3', 'ssk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gahri', 'Spoken', 'in', 'Himachal', 'Pradesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Gahri', 'Language', 'codes', 'ISO', '639-3', 'bfu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pattani', 'Spoken', 'in', 'Himachal', 'Pradesh']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Pattani', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Pattani', 'Language', 'codes', 'ISO', '639-3', 'lae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Paama', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Paama', 'Language', 'codes', 'ISO', '639-3', 'pma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Zia', 'Pronunciation', 'd', '\xcc', '\xaa', 'ia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'd', '\xcc', '\xaa', 'ia']]], 6, True)

['Zia', 'Spoken', 'in', 'lower', 'Waria', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'lower', 'Waria', 'Valley']]], 6, True)

['Zia', 'Spoken', 'in', 'Morobe', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Zia', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zia', 'Language', 'codes', 'ISO', '639-3', 'zia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sheko', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sheko', 'Language', 'codes', 'ISO', '639-3', 'she']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Camling', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Camling', 'Language', 'codes', 'ISO', '639-3', 'rab']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Puinave', 'Pronunciation', 'w', '\xc3', '\xa3', 'ns', '\xc9', '\xa4',  
'h', '\xc9', '\xa4', 't']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'w',  
'\xc3', '\xa3', 'ns', '\xc9', '\xa4', 'h', '\xc9', '\xa4', 't']]], 12,  
True)
```

```
['Puinave', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Puinave', 'Language', 'codes', 'ISO', '639-3', 'pui']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aka-Kede', 'Total', 'speakers', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aka-Kede', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aka-Kede', 'Region', 'Andaman', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aka-Kede', 'Language', 'codes', 'ISO', '639-3', 'akx']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aka-Kol', 'Total', 'speakers', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aka-Kol', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Aka-Kol', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aka-Kol', 'Language', 'codes', 'ISO', '639-3', 'aky']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oko-Juwoi', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oko-Juwoi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oko-Juwoi', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oko-Juwoi', 'Language', 'codes', 'ISO', '639-3', 'okj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bo', 'Language', 'extinction', '2010']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bo', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bo', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bo', 'Language', 'codes', 'ISO', '639-3', 'akm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka-Cari', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aka-Cari', 'Spoken', 'in', 'India']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Cari', 'Region', 'Andaman', 'Islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Cari', 'Language', 'codes', 'ISO', '639-3', 'aci']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka-Jeru', 'Total', 'speakers', 'disputed']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Jeru', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Jeru', 'Region', 'Andaman', 'Islands']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Jeru', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka-Jeru', 'Language', 'codes', 'ISO', '639-3', ':', 'akj', '\xe2', '\x80', '\x93', 'Aka-Jeru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'akj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Aka-Jeru', 'Language', 'codes', 'ISO', '639-3', 'gac', '\xe2', '\x80', '\x93', 'Great', 'Andamanese', 'creole']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'gac', '\xe2', '\x80', '\x93', 'Great', 'Andamanese', 'creole']], ['NUM', '639-3']], 12, True)

['Aka-Kora', 'Language', 'extinction', '2009']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Kora', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Aka-Kora', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aka-Kora', 'Language', 'codes', 'ISO', '639-3', 'ack']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jarawa', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jarawa', 'Region', 'Andaman', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jarawa', 'Language', 'codes', 'ISO', '639-3', 'anq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jemez', 'Total', 'speakers', '1300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jemez', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Jemez', 'Region', 'New', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jemez', 'Language', 'codes', 'ISO', '639-3', 'tow']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Awa\xe2\x80\x93Cuaiquer', 'Total', 'speakers', '21', ',', '000', '(', '1986', ',', '1991', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'Type: ANY', 'NUM', ',', 'NUM', 'Type: ANY']], ['NUM', '21'], ['NUM', '000'], ['NUM', '1986'], ['NUM', '1991']], 11, True)

['Awa\xe2\x80\x93Cuaiquer', 'Language', 'codes', 'ISO', '639-3', 'kwi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cat\xc3\xado', 'Spoken', 'in', 'Colombia', ',', 'Panama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Cat\xc3\xado', 'Language', 'codes', 'ISO', '639-3', 'cto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cham\xc3\xad', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cham\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'cmi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kulung', 'Spoken', 'in', 'Nepal', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Kulung', 'Language', 'codes', 'ISO', '639-3', 'kle']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ndali', 'Spoken', 'in', 'Tanzania', ',', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Ndali', 'Language', 'codes', 'ISO', '639-3', 'ndh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Siuslaw', 'Language', 'extinction', '1970s']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Siuslaw', 'Spoken', 'in', 'Oregon', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Siuslaw', 'Spoken', 'in', 'Siuslaw', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Siuslaw', 'Language', 'codes', 'ISO', '639-3', 'sis']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spanish', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Spoken', 'in', 'Spain']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'osp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dizin', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dizin', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Dizin', 'Language', 'codes', 'ISO', '639-3', 'mdx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Andoque', 'Spoken', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Andoque', 'Language', 'codes', 'ISO', '639-3', 'ano']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Barasana\xe2\x80\x93Eduria', 'Spoken', 'in', 'Colombia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Barasana\xe2\x80\x93Eduria', 'Language', 'codes', 'ISO', '639-3', 'bsn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abui', 'Spoken', 'in', 'Alor', 'Island']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Abui', 'Language', 'codes', 'ISO', '639-3', 'adn']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cabiyari', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Cabiyari', 'Language', 'codes', 'ISO', '639-3', 'cbb']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kakwa', 'Spoken', 'in', 'Colombia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kakwa', 'Spoken', 'in', 'Brazil']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kakwa', 'Region', 'South', 'American', 'Northwest']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Kakwa', 'Writing', 'system', 'Latin', 'alphabet']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Kakwa', 'Language', 'codes', 'ISO', '639-3', 'cbv']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kakwa', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of', 'Vaupes', 'Department', 'in', 'Colombia']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Location', 'of', 'Vaupes', 'Department', 'in', 'Colombia']], ['NUM', '639-3']], 11, True)

['Khinalug', 'Region', 'Quba']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Khinalug', 'Spoken', 'in', 'Azerbaijan']:  
 ([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Khinalug', 'Pronunciation', 'k', '\xc3', '\xa6', 't', '\xca', '\x83',  
'mit', '\xcd', '\xa1', 's', '\xca', '\xbc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k',  
'\xc3', '\xa6', 't', '\xca', '\x83', 'mit', '\xcd', '\xa1', 's', '\xca',  
'\xbc']]], 14, True)
```

```
['Khinalug', 'Language', 'codes', 'ISO', '639-3', 'kjj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gumuz', 'Total', 'speakers', '179', ',', '348', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '179'], ['NUM', '348']], 8,  
True)
```

```
['Gumuz', 'Total', 'speakers', '40', ',', '000', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '40'], ['NUM', '000']], 8,  
True)
```

```
['Gumuz', 'Spoken', 'in', 'Ethiopia', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Gumuz', 'Region', 'Benishangul-Gumuz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Gumuz', 'Language', 'codes', 'ISO', '639-3', 'guk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngiemboon', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ngiemboon', 'Language', 'codes', 'ISO', '639-3', 'nnh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Burgundian', 'Spoken', 'in', 'Gaul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Burgundian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
```

```

'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Burgundian', 'Language', 'codes', 'ISO', '639-3', '1em']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sio', 'Region', 'Morobe', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Sio', 'Language', 'codes', 'ISO', '639-3', 'xsi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['K\x3\xa2te', 'Pronunciation', 'k', '\xc9', '\x94', 't', '\xc9', '\x9b']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k',
'\xc9', '\x94', 't', '\xc9', '\x9b']], 8, True)

['K\x3\xa2te', 'Spoken', 'in', 'Huon', 'Peninsula']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['K\x3\xa2te', 'Spoken', 'in', 'Morobe', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['K\x3\xa2te', 'Writing', 'system', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['K\x3\xa2te', 'Language', 'codes', 'ISO', '639-3', 'kmg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pare', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Pare', 'Language', 'codes', 'ISO', '639-3', 'asa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Davani', 'Spoken', 'in', 'Iran']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Davani', 'Region', 'Middle', 'East']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Davani', 'Language', 'codes', 'ISO', '639-3', 'fay']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tetserret', 'Total', 'speakers', 'few', 'thousand', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'few', 'thousand', '?']]], 6, True)

['Tetserret', 'Spoken', 'in', 'Niger']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tetserret', 'Region', 'Sahara']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tetserret', 'Language', 'codes', 'ISO', '639-3', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Busa', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Busa', 'Language', 'codes', 'ISO', '639-3', 'bqp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Singapore', 'Spoken', 'in', 'Singapore']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Singapore', 'Region', 'Singapore']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Singapore', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)



```
['Singapore', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Singapore', 'Language', 'codes', 'ISO', '639-3', 'nan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Singapore', 'Language', 'codes', 'ISO', '639-1', 'zh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)
```

```
['Kriva', 'Region', 'Kriva', 'Palanka']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Kriva', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Republic', 'of', 'Macedonia']], 6, True)
```

```
['Kriva', 'Writing', 'system', 'Cyrillic', '(', 'Macedonian', 'variant',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)
```

```
['Kriva', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Mirpur', 'Punjabi', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Mirpur', 'Spoken', 'in', 'Pakistan', ',', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']], 6, True)
```

```
['Mirpur', 'Region', 'Azad', 'Kashmir', ',', 'Punjab', 'region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Azad',
'Kashmir', ',', 'Punjab', 'region']], 7, True)
```

```
['Mirpur', 'Language', 'codes', 'ISO', '639-2', 'lah']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Mirpur', 'Language', 'codes', 'ISO', '639-3', 'pmu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Omi', 'Spoken', 'in', 'Orientale', 'Province', ',', 'Democratic',
'Republic', 'of', 'the', 'Congo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Orientale', 'Province', ',', 'Democratic', 'Republic', 'of', 'the',
'Congo']]], 11, True)
```

```
['Omi', 'Language', 'codes', 'ISO', '639-3', 'omi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Baima', 'Spoken', 'in', 'Sichuan', ',', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Baima', 'Language', 'codes', 'ISO', '639-3', 'bqh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kamu', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']]], 7, True)
```

```
['Kamu', 'Total', 'speakers', '0']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kamu', 'Language', 'codes', 'ISO', '639-3', 'xmu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bujhyal', 'Spoken', 'in', 'Tanahu', 'District', ',', 'Nepal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Tanahu',
'District', ',', 'Type: ANY']]], 7, True)
```

```
['Bujhyal', 'Language', 'codes', 'ISO', '639-3', 'byh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chepang', 'Spoken', 'in', 'Nepal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Chepang', 'Writing', 'system', 'Devanagari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Chepang', 'Language', 'codes', 'ISO', '639-3', 'cdm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Wayu', 'Spoken', 'in', 'Janakpur', 'Zone', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Janakpur', 'Zone', ',', 'Type: ANY']]], 7, True)
```

```
['Wayu', 'Language', 'codes', 'ISO', '639-3', 'vay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Deilami', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Deilami', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Mangbutu', 'Spoken', 'in', 'Congo', '(', 'DRC', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)
```

```
['Mangbutu', 'Language', 'codes', 'ISO', '639-3', 'mdk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Mangbetu', 'Spoken', 'in', 'Congo', '(', 'DRC', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 7, True)
```

```
['Mangbetu', 'Language', 'codes', 'ISO', '639-3', 'mdj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Efe', 'Spoken', 'in', 'Ituri', 'forest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Efe', 'Language', 'codes', 'ISO', '639-3', 'efe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lese', 'Spoken', 'in', 'Congo', '(', 'DRC', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Lese', 'Language', 'codes', 'ISO', '639-3', 'les']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Malawi', 'Spoken', 'in', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Malawi', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Malawi', 'Language', 'codes', 'ISO', '639-3', 'lon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kumarbhadg', 'Writing', 'system', 'Devnagri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kumarbhadg', 'Language', 'codes', 'ISO', '639-3', 'kmj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jakati', 'Language', 'codes', 'ISO', '639-3', 'jat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pardhan', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pardhan', 'Total', 'speakers', '117', ',', '000', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'people']], ['NUM', '117'], ['NUM', '000']], 7, True)

['Pardhan', 'Writing', 'system', 'Indic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```

['Pardhan', 'Language', 'codes', 'ISO', '639-3', 'pch']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jordanian', 'Spoken', 'in', 'Jordan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Jordanian', 'Total', 'speakers', 'Second', 'language', ':', '15',
'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'Second', 'language', ':', 'NUM', 'million']], ['NUM', '15']], 8, True)

['Jordanian', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Jordanian', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Jordanian', 'Language', 'codes', 'ISO', '639-3', 'ajp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guanano', 'Spoken', 'in', 'Brazil', ',', 'Colombia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Guanano', 'Total', 'speakers', '1000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Guanano', 'Language', 'codes', 'ISO', '639-3', 'gvc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Barbare\xc3\xb1o', 'Language', 'codes', 'ISO', '639-3', 'boi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ocaina', 'Spoken', 'in', 'Peru', ',', 'Colombia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Ocaina', 'Total', 'speakers', '66']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ocaina', 'Language', 'codes', 'ISO', '639-3', 'oca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Transpiranto', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Transpiranto', 'Created', 'by', 'Gr', '\xc3', '\xb6', 'nk', '\xc3', '\xb6', 'pings', 'Veckoblad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Gr', '\xc3', '\xb6', 'nk', '\xc3', '\xb6', 'pings', 'Veckoblad']]], 11, True)

['Transpiranto', 'Category', '(', 'purpose', ')', 'jest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Transpiranto', 'Category', '(', 'purpose', ')', 'int', '.', 'auxiliary', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'int', '.', 'auxiliary', 'languages']]], 9, True)

['Transpiranto', 'Category', '(', 'purpose', ')', 'Transpiranto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Transpiranto', 'Date', 'founded', '1929']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Transpiranto', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Arji', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arji', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

```
['Arji', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Shompen', 'Spoken', 'in', 'Great', 'Nicobar', 'Island']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Great',  
'Nicobar', 'Island']]], 6, True)
```

```
['Shompen', 'Language', 'codes', 'ISO', '639-3', 'sii']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Muinane', 'Spoken', 'in', 'Colombia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Muinane', 'Total', 'speakers', '150']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Muinane', 'Language', 'codes', 'ISO', '639-3', 'bmr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Fasu', 'Spoken', 'in', 'New', 'Guinea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Fasu', 'Dialects', 'Some']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Fasu', 'Dialects', 'Kaibu', '(', 'Kaipu', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\(', 'Type: ANY', '\\)']]], 6, True)
```

```
['Fasu', 'Dialects', 'Namome', '(', 'Namumi', ',', 'Namuni', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Namome', '\\  
(', 'Namumi', ',', 'Namuni', '\\)']]], 8, True)
```

```
['Fasu', 'Language', 'codes', 'ISO', '639-3', 'faa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tokelauan', 'Spoken', 'in', 'Tonga']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Tokelauan', 'Language', 'codes', 'ISO', '639-3', 'nkp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alo-Teqel', 'Spoken', 'in', 'Vanua', 'Lava']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Alo-Teqel', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Rerep', 'Dialects', 'Tisman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Rerep', 'Language', 'codes', 'ISO', '639-3', 'pgk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balomar', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Balomar', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Balomar', 'Region', 'Province', 'of', 'Entre', 'R', '\xc3', '\xad',  
'os']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Province',  
'of', 'Entre', 'R', '\xc3', '\xad', 'os']]], 9, True)

['Balomar', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chan\xc3\xa1', 'Spoken', 'in', 'Uruguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chan\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'qs1', '\xe2',  
'\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'qs1', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Baenan', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Baenan', 'Region', 'Bahia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Baenan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Baenan', 'Language', 'codes', 'ISO', '639-3', '0iu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mire', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mire', 'Region', 'Southwest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mire', 'Language', 'codes', 'ISO', '639-3', 'mvh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Culle', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Culle', 'Region', 'La', 'Libertad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Culle', 'Region', 'Cajabamba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Culle', 'Region', 'Pallasca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

ANY']]], 3, True)

['Culle', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Culle', 'Language', 'codes', 'ISO', '639-3', '0iq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ndam', 'Spoken', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ndam', 'Region', 'Southwest']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ndam', 'Language', 'codes', 'ISO', '639-3', 'ndm']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cuitlatec', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cuitlatec', 'Region', 'Guerrero']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Cuitlatec', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Cuitlatec', 'Language', 'codes', 'ISO', '639-3', 'qpb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Somrai', 'Spoken', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Somrai', 'Region', 'Southwest']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Somrai', 'Language', 'codes', 'ISO', '639-3', 'sor']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tumak', 'Spoken', 'in', 'Chad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tumak', 'Region', 'Southwest']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Tumak', 'Language', 'codes', 'ISO', '639-3', 'tmc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abduyi', 'Spoken', 'in', 'Iran']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Abduyi', 'Region', 'Middle', 'East']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Abduyi', 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Abduyi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Singdarin', 'Spoken', 'in', 'Singapore']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Singdarin', 'Total', 'speakers', 'no', 'statistics', 'yet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'no', 'statistics', 'yet']]], 6, True)

['Singdarin', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Oriya', 'Region', 'Orissa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Oriya', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oriya', 'Writing', 'system', 'Oriya', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Oriya', 'Official', 'language', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Oriya', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-2', 'ori']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-3', 'ori']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oriya', 'Language', 'codes', 'ISO', '639-1', 'or']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Andi', 'Spoken', 'in', 'Southern', 'Dagestan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Andi', 'Language', 'codes', 'ISO', '639-3', 'ani']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Akhvakh', 'Language', 'codes', 'ISO', '639-3', 'akv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Nagarchal', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Nagarchal', 'Total', 'speakers', 'few', 'if', 'any']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'few', 'if', 'any']]], 6, True)
```

```
['Nagarchal', 'Language', 'codes', 'ISO', '639-3', 'nbg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Koibal', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Koibal', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Koibal', 'Language', 'codes', 'ISO', '639-3', 'zkb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Solresol', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Solresol', 'Date', 'founded', '1827']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Solresol', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Solresol', 'Category', '(', 'purpose', ')', 'musical', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)
```

```
['Solresol', 'Category', '(', 'purpose', ')', 'Solresol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']]], 6, True)
```

```
['Solresol', 'Created', 'by', 'Fran', '\xc3', '\xa7', 'ois', 'Sudre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Fran',
```

'\xc3', '\xa7', 'ois', 'Sudre']]], 8, True)

['Solresol', 'Category', '(', 'sources', ')', 'a', 'priori']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)'], 'Type: ANY', 'Type: ANY']]], 7, True)

['Solresol', 'Setting', 'and', 'usage', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', '\xe2', '\x80', '\x93']]], 7, True)

['Solresol', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Jaminjung', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Jaminjung', 'Region', 'Victoria', 'River', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'River', ',', 'Type: ANY']]], 6, True)

['Jaminjung', 'Language', 'codes', 'ISO', '639-3', 'djd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nungali', 'Total', 'speakers', 'Probably', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'Probably', 'extinct']]], 5, True)

['Nungali', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nungali', 'Language', 'codes', 'ISO', '639-3', 'nug']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Irantxe', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Irantxe', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Irantxe', 'Region', 'Mato', 'Grosso', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Irantxe', 'Region', 'Irantxe', ',', 'Mynky']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Irantxe', 'Language', 'codes', 'ISO', '639-3', 'irn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Movima', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Movima', 'Region', 'Beni', 'Department']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Movima', 'Official', 'language', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Movima', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Movima', 'Language', 'codes', 'ISO', '639-3', 'mzp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aikan\xc3\xa3', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aikan\xc3\xa3', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aikan\xc3\xa3', 'Region', 'Rond', '\xc3', '\xb4', 'nia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rond', '\xc3',  
'\xb4', 'nia']]], 6, True)

['Aikan\xc3\xa3', 'Language', 'codes', 'ISO', '639-2', 'min', '(', 'B',  
)']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'min', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Aikan\\xc3\\xa3', 'Language', 'codes', 'ISO', '639-2', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 8, True)

['Aikan\\xc3\\xa3', 'Language', 'codes', 'ISO', '639-3', 'tba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Torgut', 'Spoken', 'in', 'Mongolia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Torgut', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Torgut', 'Region', 'Xinjiang']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Torgut', 'Region', 'Khovd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Torgut', 'Region', 'eastern', 'Kalmykia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Torgut', 'Language', 'codes', 'ISO', '639-2', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Torgut', 'Language', 'codes', 'ISO', '639-3', 'xal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kol', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kol', 'Language', 'codes', 'ISO', '639-3', 'biw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Cua', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cua', 'Total', 'speakers', '27766', '(', 'SIL', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\(', 'SIL', '\\)']], ['NUM', '27766']], 7, True)

['Cua', 'Language', 'codes', 'ISO', '639-3', 'cua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aja/Adja', 'Spoken', 'in', 'Ghana', ',', 'Togo', 'and', 'Benin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ghana', ',', 'Togo', 'and', 'Benin']], 8, True)

['Aja/Adja', 'Language', 'codes', 'ISO', '639-3', 'ajg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wuxi', 'dialect', 'Total', 'speakers', 'approx', '.', '2-4', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total', 'speakers', 'approx', '.', 'NUM', 'million']], ['NUM', '2-4']], 8, True)

['Wuxi', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Wuxi', 'Language', 'codes', 'ISO', '639-3', 'wu-wux']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Temein', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Temein', 'Region', 'Nuba', 'Hills']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Temein', 'Language', 'codes', 'ISO', '639-3', 'teq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Illyrian', 'Spoken', 'in', 'Once', 'Illyria', 'and', 'some', 'lands', 'adjacent']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Once', 'Illyria', 'and', 'some', 'lands', 'adjacent']]], 9, True)

['Illyrian', 'Region', 'Western', 'Balkans']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Illyrian', 'Language', 'codes', 'ISO', '639-3', 'xil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bahnar', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bahnar', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bahnar', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Bahnar', 'Language', 'codes', 'ISO', '639-3', 'bdq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Old', 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Urarina', 'Spoken', 'in', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Urarina', 'Region', 'Loreto', ',', 'Urarinas', 'district']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Loreto', ',', 'Urarinas', 'district']]], 6, True)

['Urarina', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Urarina', 'Language', 'codes', 'ISO', '639-3', 'ura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Guat\xc3\xb3', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Guat\xc3\xb3', 'Region', 'Mato', 'Grosso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Guat\xc3\xb3', 'Language', 'codes', 'ISO', '639-3', 'gta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sechura', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sechura', 'Region', 'Piura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Sechura', 'Language', 'codes', 'ISO', '639-3', 'qfi', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qfi', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 9, True)

['Laghu', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Laghu', 'Language', 'codes', 'ISO', '639-3', 'lgb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gothic', 'Spoken', 'in', 'Oium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gothic', 'Spoken', 'in', 'Dacia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gothic', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gothic', 'Spoken', 'in', 'Gallia', 'Narbonensis']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gothic', 'Spoken', 'in', 'Hispania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gothic', 'Spoken', 'in', 'Crimea']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gothic', 'Writing', 'system', 'Gothic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Gothic', 'Language', 'codes', 'ISO', '639-2', 'got']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Gothic', 'Language', 'codes', 'ISO', '639-3', 'got']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Oran', 'Spoken', 'in', 'Malaysia', ',', 'Singapore']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Oran', 'Language', 'codes', 'ISO', '639-2', 'ors', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ors', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Oran', 'Language', 'codes', 'ISO', '639-2', 'ors', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ors', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)

['Oran', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

```
['Hakha', 'Spoken', 'in', 'Burma', '/', 'Myanmar']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Burma',  
('/', 'Type: ANY')]]], 6, True)
```

```
['Hakha', 'Spoken', 'in', 'India']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Hakha', 'Spoken', 'in', 'Bangladesh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Hakha', 'Language', 'codes', 'ISO', '639-2', 'cnh', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cnh', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)
```

```
['Hakha', 'Language', 'codes', 'ISO', '639-2', 'cnh', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cnh', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)
```

```
['Hakha', 'Language', 'codes', 'ISO', '639-3', 'cnh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kassite', 'Spoken', 'in', 'Babylon', '(', 'extinct', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']]]], 7, True)
```

```
['Kassite', 'Region', 'Middle', 'East']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Kassite', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']]], ['NUM',  
'639-3']], 10, True)
```

```
['Kassite', 'Language', 'codes', 'ISO', '639-3', '0ss']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Pasto', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```

['Pasto', 'Language', 'codes', 'ISO', '639-3', 'bpb']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tama', 'Total', 'speakers', '0']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tama', 'Spoken', 'in', 'Colombia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tama', 'Region', 'Orteguaza', 'River']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Tama', 'Language', 'codes', 'ISO', '639-3', 'ten']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wagdi', 'Spoken', 'in', 'India', '(', 'Vagad', 'region', 'of',
'Rajasthan', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',
'\\(', 'Vagad', 'region', 'of', 'Rajasthan', '\\)']]], 10, True)

['Wagdi', 'Language', 'codes', 'ISO', '639-3', 'wbr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Caura', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Caura', 'Region', 'Chaura', 'Island']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Caura', 'Language', 'codes', 'ISO', '639-3', 'crv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Universal', 'Total', 'speakers', 'none']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

```
['Universal', 'Date', 'founded', '1923']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Universal', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Universal', 'Category', '(', 'purpose', ')', 'Esperantido']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']]], 6, True)
```

```
['Universal', 'Category', '(', 'purpose', ')', 'Universal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']]], 6, True)
```

```
['Universal', 'Pronunciation', 'univer', '\\xcb', '\\x88', 'sal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation',  
'univer', '\\xcb', '\\x88', 'sal']]], 6, True)
```

```
['Universal', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)
```

```
['Zhaba', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zhaba', 'Language', 'codes', 'ISO', '639-3', 'zhab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Rangas', 'Spoken', 'in', 'Uttaranchal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Rangas', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']]], 6, True)
```

```
['Rangas', 'Language', 'codes', 'ISO', '639-3', 'rgk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Turung', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Turung', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Turung', 'Region', 'Assam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Turung', 'Language', 'codes', 'ISO', '639-3', 'try']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Degema', 'Spoken', 'in', 'Rivers', 'State', ',', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.'], 'Type: ANY']]], 7, True)

['Degema', 'Total', 'speakers', '10', ',', '000', 'in', '1999']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '10'], ['NUM', '000']], 8, True)

['Degema', 'Language', 'codes', 'ISO', '639-3', 'deg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Judaео-Spanish,', 'Total', 'speakers', '8', ',', '000', 'in', 'Turkey']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '8'], ['NUM', '000']], 8, True)

['Judaео-Spanish,', 'Total', 'speakers', '1', ',', '000', 'in', 'Greece']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '1'], ['NUM', '000']], 8, True)

['Judaео-Spanish,', 'Total', 'speakers', '300', 'in', 'the', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'in', 'the', 'United', 'States']]], ['NUM', '300']], 8, True)

['Judaео-Spanish,', 'Total', 'speakers', 'unknown', 'numbers', 'elsewhere', ',', 'steady', 'decline', 'in', 'all', 'those', 'places']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'unknown', 'numbers', 'elsewhere', ',', 'steady', 'decline', 'in', 'all',



'those', 'places']]], 13, True)

['Judaean-Spanish,', 'Pronunciation', 'd', '\xca', '\x92', 'u', '\xc3',  
\xb0', 'eo', 'espa', '\xc9', '\xb2', 'ol']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'd',  
\xca', '\x92', 'u', '\xc3', '\xb0', 'eo', 'espa', '\xc9', '\xb2', 'ol']]],  
13, True)

['Judaean-Spanish,', 'Spoken', 'in', 'Israel']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Judaean-Spanish,', 'Spoken', 'in', 'Turkey']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Judaean-Spanish,', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Judaean-Spanish,', 'Spoken', 'in', 'Greece']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Judaean-Spanish,', 'Spoken', 'in', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
Type: ANY']]], 4, True)

['Judaean-Spanish,', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Judaean-Spanish,', 'Official', 'language', 'in', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
in', 'Type: ANY']]], 5, True)

['Judaean-Spanish,', 'Language', 'codes', 'ISO', '639-2', 'lad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Judaean-Spanish,', 'Language', 'codes', 'ISO', '639-3', 'lad']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cayuse', 'Spoken', 'in', 'United', 'States']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Cayuse', 'Region', 'Oregon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Cayuse', 'Language', 'codes', 'ISO', '639-3', 'xcy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lojban', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Lojban', 'Date', 'founded', '1987']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lojban', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Lojban', 'Category', '(', 'purpose', ')', 'engineered', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Lojban', 'Category', '(', 'purpose', ')', 'logical', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Lojban', 'Category', '(', 'purpose', ')', 'Lojban']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']]], 6, True)

['Lojban', 'Writing', 'system', 'mixed', 'a', 'priori', '/', 'a', 'posteriori']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'mixed', 'a', 'priori', '/', 'a', 'posteriori']]], 9, True)

['Lojban', 'Pronunciation', 'la', '\xcb', '\x88', 'lo', '\xca', '\x92', 'ban']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'la', '\xcb', '\x88', 'lo', '\xca', '\x92', 'ban']]], 9, True)

['Lojban', 'Setting', 'and', 'usage', 'A', 'logically', 'engineered', 'language', 'for', 'various', 'usages']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'A', 'logically', 'engineered', 'language', 'for', 'various',  
'usages']]], 11, True)
```

```
['Lojban', 'Language', 'codes', 'ISO', '639-2', 'jbo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Lojban', 'Language', 'codes', 'ISO', '639-3', 'jbo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Albanian', 'Pronunciation', '\xca', '\x83', 'cip']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xca',  
'\x83', 'cip']]], 5, True)
```

```
['Albanian', 'Official', 'language', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Albanian', 'Official', 'language', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Albanian', 'Official', 'language', 'in', 'and', 'recognised', 'as', 'a',  
'minority', 'language', 'in']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'and', 'recognised', 'as', 'a', 'minority', 'language', 'in']]], 11,  
True)
```

```
['Albanian', 'Official', 'language', 'in', ':', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Albanian', 'Official', 'language', 'in', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Albanian', 'Official', 'language', 'in', 'Montenegro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Albanian', 'Official', 'language', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Albanian', 'Official', 'language', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Albanian', 'Regulated', 'by', 'officially', 'by', 'the', 'Social',  
'Sciences', 'and', 'Albanological', 'Section', 'of', 'the', 'Academy',  
'of', 'Sciences', 'of', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'officially', 'by', 'the', 'Social', 'Sciences', 'and', 'Albanological',  
'Section', 'of', 'the', 'Academy', 'of', 'Sciences', 'of', 'Albania']]],  
18, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-2', 'alb', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'alb', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-2', 'sqi', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sqi', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-3', 'sqi', '\\xe2', '\\x80',  
'\\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'sqi', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM',  
'639-3']], 10, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-3', ':', 'aln', '\\xe2',  
'\\x80', '\\x93', 'Gheg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'aln', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-3', 'aae', '\\xe2', '\\x80',  
'\\x93', 'Arb', '\\xc3', '\\xab', 'resh', '\\xc3', '\\xab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'aae', '\\xe2', '\\x80', '\\x93', 'Arb', '\\xc3', '\\xab', 'resh',  
'\\xc3', '\\xab']], ['NUM', '639-3']], 15, True)
```

```
['Albanian', 'Language', 'codes', 'ISO', '639-3', 'aat', '\\xe2', '\\x80',  
'\\x93', 'Arvanitika']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'aat', '\\xe2', '\\x80', '\\x93', 'Type: ANY']], ['NUM',
```

'639-3']], 10, True)

['Albanian', 'Language', 'codes', 'ISO', '639-3', 'als', '\xe2', '\x80', '\x93', 'Tosk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'als', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Albanian', 'Language', 'codes', 'ISO', '639-1', 'sq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Mbukushu', 'Spoken', 'in', 'Okavango', 'River']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Mbukushu', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Mbukushu', 'Language', 'codes', 'ISO', '639-3', 'mhw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kwangali', 'Spoken', 'in', 'Namibia', '(', 'Okavango', 'River', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Namibia', '\\(', 'Okavango', 'River', '\\)']], 8, True)

['Kwangali', 'Spoken', 'in', 'Angola']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kwangali', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Kwangali', 'Language', 'codes', 'ISO', '639-3', 'kwn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gciriku', 'Spoken', 'in', 'Okavango', 'River', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Okavango', 'River', 'Ethnicity']], 6, True)

['Gciriku', 'Spoken', 'in', 'Vagciriku']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Gciriku', 'Spoken', 'in', 'Vamanyo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gciriku', 'Spoken', 'in', 'Vashambyu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Gciriku', 'Dialects', 'Manyo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gciriku', 'Dialects', 'Gciriku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gciriku', 'Dialects', 'Shambyu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Gciriku', 'Language', 'codes', 'ISO', '639-3', 'diu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hokkien', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hokkien', 'Spoken', 'in', 'and', 'other', 'areas', 'of', 'Hoklo', 'settlement']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'other', 'areas', 'of', 'Hoklo', 'settlement']]], 9, True)

['Hokkien', 'Region', 'Southern', 'Fujian', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hokkien', 'Region', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hokkien', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hokkien', 'Regulated', 'by', 'None', '(', 'Republic', 'of', 'China', 'Ministry', 'of', 'Education', 'and', 'some', 'NGOs', 'are', 'influential', 'in', 'Taiwan', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'None', '\\(', 'Republic', 'of', 'China', 'Ministry', 'of', 'Education', 'and', 'some', 'NGOs', 'are', 'influential', 'in', 'Taiwan', '\\)']]], 19, True)

['Hokkien', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Hokkien', 'Language', 'codes', 'ISO', '639-3', 'nan-hok']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hokkien', 'Language', 'codes', 'ISO', '639-3', 'nan-fuj', '(',  
'redundant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nan-fuj', '\\(', 'redundant', '\\)']], ['NUM', '639-3']], 9,  
True)

['Hokkien', 'Language', 'codes', 'ISO', '639-3', 'nan-fuk', '(',  
'redundant', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nan-fuk', '\\(', 'redundant', '\\)']], ['NUM', '639-3']], 9,  
True)

['Puguli', 'Total', 'speakers', '15000']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Puguli', 'Region', 'West', 'Africa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Puguli', 'Spoken', 'in', 'Burkina', 'Faso']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Puguli', 'Language', 'codes', 'ISO', '639-3', 'pug']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lote', 'Region', 'Pomio', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lote', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Lote', 'Language', 'codes', 'ISO', '639-3', 'uvl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Orok', 'Total', 'speakers', '67']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Orok', 'Spoken', 'in', 'Russia']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Orok', 'Region', 'Sakhalin', 'Oblast', ',', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sakhalin', 'Oblast', ',', 'Russia']]], 6, True)

['Orok', 'Language', 'codes', 'ISO', '639-3', 'oaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ugaritic', 'Spoken', 'in', 'ancient', 'Ugarit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ugaritic', 'Language', 'codes', 'ISO', '639-2', 'uga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Ugaritic', 'Language', 'codes', 'ISO', '639-3', 'uga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Aheu', 'Spoken', 'in', 'Laos', ',', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Aheu', 'Language', 'codes', 'ISO', '639-3', 'thm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Akha', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Akha', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Akha', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Akha', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Akha', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Akha', 'Language', 'codes', 'ISO', '639-3', 'ahk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nayi', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nayi', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Nayi', 'Language', 'codes', 'ISO', '639-3', 'noz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alak', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Alak', 'Language', 'codes', 'ISO', '639-3', 'alk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Arem', 'Spoken', 'in', 'Laos', ',', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Arem', 'Language', 'codes', 'ISO', '639-3', 'aem']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bit', 'Spoken', 'in', 'Laos', ',', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Bit', 'Language', 'codes', 'ISO', '639-3', 'bgk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pinyin', 'Spoken', 'in', 'Northwest']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Pinyin', 'Region', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Pinyin', 'Language', 'codes', 'ISO', '639-3', 'pny']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Koroshi', 'Spoken', 'in', 'Iran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koroshi', 'Region', 'Middle', 'East']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koroshi', 'Official', 'language', 'in', 'None']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Koroshi', 'Language', 'codes', 'ISO', '639-3', 'ktl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bo', 'Spoken', 'in', 'Laos']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bo', 'Total', 'speakers', '0', '(', 'extinct', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\\\(', 'extinct', '\\\\)']]], ['NUM', '0']], 7, True)

['Bo', 'Language', 'codes', 'ISO', '639-3', 'bgl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Eastern', 'Spoken', 'in', 'Laos']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Eastern', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Eastern', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Eastern', 'Language', 'codes', 'ISO', '639-3', 'bru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chut', 'Spoken', 'in', 'Vietnam', ',', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Chut', 'Language', 'codes', 'ISO', '639-3', 'scb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Con', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Con', 'Language', 'codes', 'ISO', '639-3', 'cno']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Halang', 'Spoken', 'in', 'Laos', ',', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Halang', 'Language', 'codes', 'ISO', '639-3', 'hld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Korouni', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Korouni', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Korouni', 'Region', 'Middle', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Korouni', 'Official', 'language', 'in', 'None']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Korouni', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Tikopia', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tikopia', 'Region', 'Tikopia', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tikopia', 'Language', 'codes', 'ISO', '639-3', 'tkp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anuta', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Anuta', 'Region', 'Anuta', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Anuta', 'Language', 'codes', 'ISO', '639-3', 'aud']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alamblak', 'Language', 'codes', 'ISO', '639-3', 'amp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hung', 'Spoken', 'in', 'Laos', ',', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Hung', 'Language', 'codes', 'ISO', '639-3', 'hnu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ir', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ir', 'Language', 'codes', 'ISO', '639-3', 'irr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jeh', 'Spoken', 'in', 'Vietnam', ',', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Jeh', 'Language', 'codes', 'ISO', '639-3', 'jeh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jeng', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Jeng', 'Language', 'codes', 'ISO', '639-3', 'jeg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Savosavo', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Savosavo', 'Region', 'Savo', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Savosavo', 'Region', 'north', 'of', 'Guadalcanal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Savosavo', 'Region', 'Central', 'Solomons']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Savosavo', 'Language', 'codes', 'ISO', '639-3', 'svs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Touo', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Touo', 'Language', 'codes', 'ISO', '639-3', 'tqu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tetepare', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tetepare', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tetepare', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bilua', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Bilua', 'Language', 'codes', 'ISO', '639-3', 'blb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Santa', 'Spoken', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Santa', 'Language', 'codes', 'ISO', '639-3', 'stc', '\xe2', '\x80', '\x93', 'Macrolanguage']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'stc', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Santa', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Santa', 'Language', 'codes', 'ISO', '639-3', ':', 'ntu', '\xe2', '\x80', '\x93', 'Nat', '\xc3', '\xbc', 'gu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ntu', '\xe2', '\x80', '\x93', 'Nat', '\xc3', '\xbc', 'gu']], ['NUM', '639-3']], 14, True)

['Santa', 'Language', 'codes', 'ISO', '639-3', 'nlz', '\xe2', '\x80', '\x93', 'Nal', '\xc3', '\xb6', 'go']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

['ISO', 'NUM', 'nlz', '\xe2', '\x80', '\x93', 'Nal', '\xc3', '\xb6', 'go']],  
['NUM', '639-3']], 13, True)

['Nanggu', 'Spoken', 'in', 'Solomon', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Nanggu', 'Region', 'Nendo', 'Island', 'Coordinates']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]]], 5, True)

['Nanggu', 'Language', 'codes', 'ISO', '639-3', 'ngr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Moxos', 'Spoken', 'in', 'Bolivia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Moxos', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Moxos', 'Language', 'codes', 'ISO', '639-3', ':', 'igj', '\xe2', '\x80',  
'\x93', 'Ignaciano', 'Moxos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'igj', '\xe2', '\x80', '\x93', 'Ignaciano', 'Type:  
ANY']]], ['NUM', '639-3']], 12, True)

['Moxos', 'Language', 'codes', 'ISO', '639-3', 'trn', '\xe2', '\x80',  
'\x93', 'Trinitario', 'Moxos']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'trn', '\xe2', '\x80', '\x93', 'Trinitario', 'Type: ANY']]],  
['NUM', '639-3']], 11, True)

['Yur\xc3\xad', 'Total', 'speakers', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Yur\xc3\xad', 'Spoken', 'in', 'Brazil', ',', 'extending', 'slightly',  
'into', 'Colombia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Brazil',  
, ',', 'extending', 'slightly', 'into', 'Colombia']]]], 9, True)

['Yur\xc3\xad', 'Region', 'Caquet', '\xc3', '\xa1', 'river',  
'Coordinates']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Caquet',  
'\xc3', '\xa1', 'river', 'Coordinates']]], 7, True)

['Yur\xc3\xad', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Yur\xc3\xad', 'Language', 'codes', 'ISO', '639-3', '0io']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tuwali', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Tuwali', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tuwali', 'Region', 'Kiangán', ',', 'Ifugao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Tuwali', 'Official', 'language', 'in', 'Regional', 'language', 'in',  
'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Tuwali', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tuwali', 'Language', 'codes', 'ISO', '639-3', 'ifk', 'Ph', 'locator',  
'ifugao', 'kiangan', '.', 'png']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ifk', 'Ph', 'locator', 'ifugao', 'kiangan', '.', 'png']],  
['NUM', '639-3']], 12, True)

['Parya', 'Spoken', 'in', 'Tajikistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Parya', 'Spoken', 'in', 'Uzbekistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Parya', 'Spoken', 'in', 'Afghanistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Parya', 'Region', 'Gissar', 'Valley', ',', 'Surkhandarya', 'basin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Gissar', 'Valley', ',', 'Surkhandarya', 'basin']]], 7, True)

['Parya', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Parya', 'Language', 'codes', 'ISO', '639-3', 'paq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Larteh', 'Spoken', 'in', 'Ghana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Larteh', 'Region', 'Anum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Larteh', 'Region', 'Lete']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Larteh', 'Region', 'Chiripon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Larteh', 'Language', 'codes', 'ISO', '639-3', 'lar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Busuu', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Busuu', 'Region', 'North', 'West', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Busuu', 'Region', 'Menchum', 'Division']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Busuu', 'Region', 'Furu-Awa', 'Subdivision']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Busuu', 'Region', 'Furu-Awa', 'and', 'Furu-Nangwa', 'villages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Furu-Awa',
'and', 'Furu-Nangwa', 'villages']]], 6, True)
```

```
['Busuu', 'Language', 'codes', 'ISO', '639-3', 'bju']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Siraya', 'Total', 'speakers', '(', 'prev', '.', 'extinct', ';',
'revitalization', 'movement', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\
(', 'prev', '.', 'extinct', ';', 'revitalization', 'movement', '\\)']]],
11, True)
```

```
['Siraya', 'Spoken', 'in', 'Taiwan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Siraya', 'Region', 'Southwestern']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Siraya', 'Region', 'around', 'present-day', 'Tainan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Siraya', 'Region', 'from', 'Peimen', 'to', 'Hengchun', 'to', 'Tapu', '.',
'Coordinates']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'from',
'Peimen', 'to', 'Hengchun', 'to', 'Tapu', '.', 'Coordinates']]], 10, True)
```

```
['Siraya', 'Language', 'codes', 'ISO', '639-3', 'fos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Orang', 'Pronunciation', 'ora', '\\xc5', '\\x8b', 'kanak']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ora',
'\\xc5', '\\x8b', 'kanak']]], 6, True)
```

['Orang', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Orang', 'Language', 'codes', 'ISO', '639-2', 'msa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Orang', 'Language', 'codes', 'ISO', '639-3', 'orn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lun', 'Total', 'speakers', 'second', 'language', ':']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'second', 'language', ':']], 6, True)

['Lun', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Lun', 'Region', 'Sarawak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Lun', 'Language', 'codes', 'ISO', '639-3', 'lnd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Swedish', 'Writing', 'system', 'SignWriting']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Swedish', 'Total', 'signers', '8', ',', '000', 'deaf', 'users', ';', 'also', ',', 'their', 'children']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', 'NUM', ',', 'NUM', 'deaf', 'users', ';', 'also', ',', 'their', 'children']], ['NUM', '8'], ['NUM', '000']], 13, True)

['Pajapan', 'Spoken', 'in', 'Pajapan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pajapan', 'Spoken', 'in', 'Veracruz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pajapan', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pajapan', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Pajapan', 'Language', 'codes', 'ISO', '639-3', 'nhp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kemezung', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kemezung', 'Region', 'Northwest', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kemezung', 'Region', 'Donga-Mantung', 'Division']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kemezung', 'Region', 'Southwest', 'corner', 'of', 'Ako', 'Subdivision']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Southwest', 'corner', 'of', 'Ako', 'Subdivision']], 7, True)

['Kemezung', 'Region', 'Northwest', 'of', 'Nkamb', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Northwest', 'of', 'Nkamb', '\xc3', '\xa9']], 7, True)

['Kemezung', 'Region', 'town', 'of', 'Dumbu', 'and', 'village', 'of', 'Kwei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'town', 'of', 'Dumbu', 'and', 'village', 'of', 'Kwei']], 9, True)

['Kemezung', 'Language', 'codes', 'ISO', '639-3', 'dmo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sena', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Sena', 'Language', 'codes', 'ISO', '639-3', 'seh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jola-Fonyi', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)

['Jola-Fonyi', 'Spoken', 'in', 'Senegal']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Jola-Fonyi', 'Region', 'West', 'Africa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Jola-Fonyi', 'Language', 'codes', 'ISO', '639-3', 'dyo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kokota', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Kokota', 'Language', 'codes', 'ISO', '639-3', 'kkk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Marchigiano', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Marchigiano', 'Region', 'central', 'Marche', '(', 'provinces', 'of',
'Ancona', ',', 'Macerata', 'and', 'Fermo', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central',
'Marche', '\\(', 'provinces', 'of', 'Ancona', ',', 'Macerata', 'and',
'Fermo', '\\)']]], 13, True)

['Marchigiano', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93', 'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Marchigiano', 'Language', 'codes', 'ISO', '639-3', 'ita-cen']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cicipu', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cicipu', 'Region', 'Kebbi', 'State', ',', 'Niger', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kebbi', 'State', ',', 'Niger', 'State']], 7, True)

['Cicipu', 'Language', 'codes', 'ISO', '639-3', 'awc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Irish', 'Region', 'Gaeltachta', '\xc3', '\xad', ',', 'but', 'also', 'spoken', 'throughout', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Gaeltachta', '\xc3', '\xad', ',', 'but', 'also', 'spoken', 'throughout', 'Ireland']], 11, True)

['Irish', 'Writing', 'system', 'Latin', '(', 'Irish', 'variant', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Irish', 'Pronunciation', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\xcb', '\x90', 'l', '\xca', '\xb2', '\xc9', '\x9f', '\xc9', '\x99']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', '\xc9', '\xa1', 'e', '\xcb', '\x90', 'l', '\xca', '\xb2', '\xc9', '\x9f', '\xc9', '\x99']], 16, True)

['Irish', 'Official', 'language', 'in', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Irish', 'Official', 'language', 'in', 'EU']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Irish', 'Regulated', 'by', 'Foras', 'na', 'Gaeilge']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Foras', 'na', 'Gaeilge']], 6, True)

['Irish', 'Recognised', 'minority', 'language', 'in', 'UK-Northern', 'Ireland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority', 'language', 'in', 'UK-Northern', 'Ireland']], 7, True)

```

['Irish', 'Language', 'codes', 'ISO', '639-2', 'gle']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Irish', 'Language', 'codes', 'ISO', '639-3', 'gle']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Irish', 'Language', 'codes', 'Linguasphere', '50-AAA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']], 5, True)

['Irish', 'Language', 'codes', 'ISO', '639-1', 'ga']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Vedda', 'Total', 'speakers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Vedda', 'Language', 'codes', 'ISO', '639-3', 'ved']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kelantanese', 'Spoken', 'in', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Kelantanese', 'Region', 'Kelantan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Kelantanese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93', 'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Kelantanese', 'Language', 'codes', 'ISO', '639-3', 'zlm-kel']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baekje', 'Spoken', 'in', 'Paekche']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

```



```

['Baekje', 'Region', 'Korea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Baekje', 'Language', 'codes', 'ISO', '639-3', 'pkc', '\xe2', '\x80', '\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pkc', '\xe2', '\x80', '\x93', 'Macrolanguage']], ['NUM', '639-3']], 10, True)

['Baekje', 'Language', 'codes', 'ISO', '639-3', 'individual', 'code']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Baekje', 'Language', 'codes', 'ISO', '639-3', ':', 'xpp', '\xe2', '\x80', '\x93', 'Buyeo-Baekje']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'xpp', '\xe2', '\x80', '\x93', 'Buyeo-Baekje']], ['NUM', '639-3']], 11, True)

['Quileute', 'Spoken', 'in', 'Olympic', 'Peninsula', ',', 'Washington']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Olympic', 'Peninsula', ',', 'Washington']]], 7, True)

['Quileute', 'Total', 'speakers', 'moribund']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quileute', 'Language', 'codes', 'ISO', '639-3', 'qui']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Finnish', 'Total', 'speakers', 'about', '6', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'about', 'NUM', 'million']]], ['NUM', '6']], 6, True)

['Finnish', 'Region', 'Northern', 'Europe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Spoken', 'in', 'Finland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Spoken', 'in', 'Estonia']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Spoken', 'in', 'Ingria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Spoken', 'in', 'Karelia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Spoken', 'in', 'Norway']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Spoken', 'in', 'Sweden']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Finnish', 'Pronunciation', '\xcb', '\x88', 'suo', '\xcc', '\xaf', 'mi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'suo', '\xcc', '\xaf', 'mi']]], 8, True)

['Finnish', 'Official', 'language', 'in', 'Finland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Finnish', 'Official', 'language', 'in', 'European', 'Union']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Finnish', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'recognised', 'as', 'minority', 'language', 'in']]], 9, True)

['Finnish', 'Official', 'language', 'in', ':', 'Sweden']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Finnish', 'Official', 'language', 'in', 'Republic', 'of', 'Karelia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Republic', 'of', 'Karelia']]], 7, True)

['Finnish', 'Language', 'codes', 'ISO', '639-2', 'fin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Finnish', 'Language', 'codes', 'ISO', '639-3', 'fin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Finnish', 'Language', 'codes', 'ISO', '639-1', 'fi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Dutch', 'Total', 'speakers', '(', 'EU', 'alone', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\  
(', 'EU', 'alone', '\\)']], 7, True)

['Dutch', 'Total', 'speakers', '23', 'million', '(', 'native', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', '\\((', 'native', '\\)']], ['NUM', '23']], 8, True)

['Dutch', 'Total', 'speakers', '28', 'million', '(', 'total', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'million', '\\((', 'total', '\\)']], ['NUM', '28']], 8, True)

['Dutch', 'Spoken', 'in', 'mainly', 'the', 'Netherlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'mainly',  
'the', 'Netherlands']], 6, True)

['Dutch', 'Spoken', 'in', 'Belgium']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Dutch', 'Spoken', 'in', 'and', 'Suriname']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Dutch', 'Spoken', 'in', 'also', 'in', 'Aruba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'also',  
'in', 'Aruba']], 6, True)

['Dutch', 'Spoken', 'in', 'Cura', '\\xc3', '\\xa7', 'ao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Cura',  
'\\xc3', '\\xa7', 'ao']], 7, True)

['Dutch', 'Spoken', 'in', 'and', 'Sint-Maarten']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

```
['Dutch', 'Spoken', 'in', 'as', 'well', 'as', 'the', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'as',  
'well', 'as', 'the', 'United', 'States']]], 9, True)
```

```
['Dutch', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dutch', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dutch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dutch', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dutch', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dutch', 'Pronunciation', '\xcb', '\x88', 'ne', '\xcb', '\x90', 'd',  
'\xc9', '\x99', 'rl', '\xc9', '\x91', 'nts', '(', 'listen', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb',  
'\x88', 'ne', '\xcb', '\x90', 'd', '\xc9', '\x99', 'rl', '\xc9', '\x91',  
'nts', '\\(', 'listen', '\\)']]], 17, True)
```

```
['Dutch', 'Official', 'language', 'in', 'Aruba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Dutch', 'Official', 'language', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

```
['Dutch', 'Official', 'language', 'in', 'Cura', '\xc3', '\xa7', 'ao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Cura', '\xc3', '\xa7', 'ao']]], 8, True)
```

```
['Dutch', 'Official', 'language', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)
```

['Dutch', 'Official', 'language', 'in', 'Sint', 'Maarten']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Dutch', 'Official', 'language', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Dutch', 'Official', 'language', 'in', 'Benelux']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Dutch', 'Official', 'language', 'in', 'European', 'Union']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Dutch', 'Regulated', 'by', 'Nederlandse', 'Taalunie']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Dutch', 'Language', 'codes', 'ISO', '639-2', 'dut', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'dut', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Dutch', 'Language', 'codes', 'ISO', '639-2', 'nld', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nld', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Dutch', 'Language', 'codes', 'ISO', '639-3', 'nld']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dutch', 'Language', 'codes', 'ISO', '639-1', 'nl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Salish', 'Spoken', 'in', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Salish', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Salish', 'Language', 'codes', 'ISO', '639-3', ':', 'fla', '\\xe2', '\\x80',  
'\\x93', 'Flathead']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'fla', '\xe2', '\x80', '\x93', 'Flathead']], ['NUM', '639-3']], 11, True)

['Salish', 'Language', 'codes', 'ISO', '639-3', 'spo', '\xe2', '\x80', '\x93', 'Spokane']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'spo', '\xe2', '\x80', '\x93', 'Spokane']], ['NUM', '639-3']], 10, True)

['Acehnese', 'Total', 'speakers', '3', 'million']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '3']], 5, True)

['Acehnese', 'Spoken', 'in', 'Indonesia', ',', 'Malaysia']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Acehnese', 'Region', 'Aceh', ',', 'Sumatra']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Acehnese', 'Language', 'codes', 'ISO', '639-2', 'ace']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Acehnese', 'Language', 'codes', 'ISO', '639-3', 'ace']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Acehnese', 'Language', 'codes', 'ISO', '639-3', 'Aceh', 'province', ',', 'Sumatra']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Aceh', 'province', ',', 'Sumatra']], ['NUM', '639-3']], 9, True)

['Icelandic', 'Spoken', 'in', 'Iceland']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Icelandic', 'Spoken', 'in', 'Denmark']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Icelandic', 'Spoken', 'in', 'Norway']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Icelandic', 'Writing', 'system', 'Latin', '(', 'Icelandic', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Icelandic', 'Pronunciation', 'is', '(', 't', ')', 'l', '\xc9', '\x9b', 'nska']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'is', '\(', 't', '\)'], 'l', '\xc9', '\x9b', 'nska']]], 10, True)

['Icelandic', 'Official', 'language', 'in', 'Iceland', 'de', 'facto']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Iceland', 'de', 'facto']]], 7, True)

['Icelandic', 'Language', 'codes', 'ISO', '639-2', 'ice', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'ice', '\(\(', 'B', '\)\)']], ['NUM', '639-2']], 9, True)

['Icelandic', 'Language', 'codes', 'ISO', '639-2', 'isl', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'isl', '\(\(', 'T', '\)\)']], ['NUM', '639-2']], 9, True)

['Icelandic', 'Language', 'codes', 'ISO', '639-3', 'isl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Icelandic', 'Language', 'codes', 'Linguasphere', '52-AAA-aa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Icelandic', 'Language', 'codes', 'ISO', '639-1', 'is']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Haifeng', 'dialect', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haifeng', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Haifeng', 'Region', 'Mainly', 'in', 'Shanwei', ',', 'eastern', 'Guangdong', 'province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mainly', 'in', 'Shanwei', ',', 'eastern', 'Guangdong', 'province']]], 9, True)

['Haifeng', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Sobei', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sobei', 'Region', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Sobei', 'Language', 'codes', 'ISO', '639-3', 'sob']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bikol', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bikol', 'Region', 'Bicol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bikol', 'Writing', 'system', 'Historically', 'written', 'in', 'Baybayin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Historically', 'written', 'in', 'Type: ANY']]], 7, True)

['Bikol', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']]], 9, True)

['Bikol', 'Language', 'codes', 'ISO', '639-2', 'bik']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Jalapa', 'Spoken', 'in', 'Oaxaca', ',', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Jalapa', 'Language', 'codes', 'ISO', '639-3', 'maj']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Axininca', 'Spoken', 'in', 'Peru']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Axininca', 'Language', 'codes', 'ISO', '639-3', 'cpc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ndrumbea', 'Spoken', 'in', 'New', 'Caledonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ndrumbea', 'Language', 'codes', 'ISO', '639-3', 'duf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['T\xe1\xba\x9\xe1\xba\x9', 'Spoken', 'in', 'Rivers', 'State', ',', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'S', '.', 'Type: ANY']], 7, True)

['T\xe1\xba\x9\xe1\xba\x9', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['T\xe1\xba\x9\xe1\xba\x9', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['T\xe1\xba\x9\xe1\xba\x9', 'Language', 'codes', 'ISO', '639-3', 'tkq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Avatime', 'Spoken', 'in', 'Ghana']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Avatime', 'Region', 'Volta']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Avatime', 'Language', 'codes', 'ISO', '639-3', 'avn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yan-nha\xc5\x8bu', 'Spoken', 'in', 'Crocodile', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Yan-nha\xc5\x8bu', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Yan-nha\xc5\x8bu', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yan-nha\xc5\x8bu', 'Language', 'codes', 'ISO', '639-3', 'jay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kairiru', 'Language', 'codes', 'ISO', '639-3', 'kxa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jiamao', 'Region', 'Hainan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Jiamao', 'Language', 'codes', 'ISO', '639-3', 'jio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sabino', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Sabino', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sabino', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Harappan', 'Spoken', 'in', 'Indus', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Harappan', 'Language', 'codes', 'ISO', '639-3', 'xiv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romano-Greek', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Romano-Greek', 'Language', 'codes', 'ISO', '639-3', 'rge']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Djeoromitxi', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Djeoromitxi', 'Region', 'Rond', '\xc3', '\xb4', 'nia', ',', 'at', 'the',  
'headwaters', 'of', 'the', 'Rio', 'Branco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rond', '\xc3',  
'\xb4', 'nia', ',', 'at', 'the', 'headwaters', 'of', 'the', 'Rio',  
'Branco']], 14, True)

['Djeoromitxi', 'Language', 'codes', 'ISO', '639-3', 'jbt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karir\xc3\xad', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Karir\xc3\xad', 'Region', 'between', 'Bahia', 'and', 'Maranh', '\xc3',  
'\xa3', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'between',  
'Bahia', 'and', 'Maranh', '\xc3', '\xa3', 'o']], 9, True)

['Karir\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'kzw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Achuar', 'Spoken', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Achuar', 'Language', 'codes', 'ISO', '639-3', 'acu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngambay', 'Spoken', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ngambay', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ngambay', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ngambay', 'Total', 'speakers', '900', ',', '000', 'in', 'Chad']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '900'], ['NUM', '000']], 8, True)

['Ngambay', 'Writing', 'system', 'Roman']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ngambay', 'Language', 'codes', 'ISO', '639-3', 'sba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Idoma', 'Language', 'codes', 'ISO', '639-3', 'idu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngbandi', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Ngbandi', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nimadi', 'Spoken', 'in', 'India', '(', 'Nimar', 'region', 'of', 'Madhya', 'Pradesh', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', '\\(', 'Nimar', 'region', 'of', 'Madhya', 'Pradesh', '\\)']], 11, True)

['Nimadi', 'Language', 'codes', 'ISO', '639-2', 'raj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Nimadi', 'Language', 'codes', 'ISO', '639-3', 'noe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lambadi', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lambadi', 'Language', 'codes', 'ISO', '639-2', 'raj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lambadi', 'Language', 'codes', 'ISO', '639-3', 'lmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bajjika', 'Spoken', 'in', 'Bihar', 'in', 'India', ',', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Bihar',  
'in', 'India', ',', 'Type: ANY']]], 8, True)

['Bajjika', 'Writing', 'system', 'Devanagari', ',', 'Kaithi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bajjika', 'Official', 'language', 'in', 'Bihar', 'state', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Bihar', 'state', 'in', 'India']]], 8, True)

['Bajjika', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Bajjika', 'Language', 'codes', 'ISO', '639-3', 'mai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bijago', 'Spoken', 'in', 'Guinea-Bissau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bijago', 'Region', 'Offshore', 'Bissagos', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

```
['Bijago', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Bijago', 'Language', 'codes', 'ISO', '639-3', 'bjg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pyu', 'Spoken', 'in', 'Burma']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Pyu', 'Language', 'codes', 'ISO', '639-3', 'pyx']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Maastrichtian', 'Pronunciation', 'm', '\xc9', '\x99', '\xcb', '\x88',  
'st', '\xca', '\x80', 'e', '\xcb', '\x90', '\xc3', '\xa7', 's']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm',  
'\xc9', '\x99', '\xcb', '\x88', 'st', '\xca', '\x80', 'e', '\xcb', '\x90',  
'\xc3', '\xa7', 's']], 16, True)
```

```
['Maastrichtian', 'Spoken', 'in', 'City', 'of', 'Maastricht', ',', 'the',  
'Netherlands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'City',  
'of', 'Maastricht', ',', 'the', 'Netherlands']], 9, True)
```

```
['Maastrichtian', 'Official', 'language', 'in', 'Limburg', ',', 'The',  
'Netherlands', ':', 'Recognised', 'as', 'regional', 'language', 'as', 'a',  
'variant', 'of', 'Limburgish']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Limburg', ',', 'The', 'Netherlands', ':', 'Recognised', 'as',  
'regional', 'language', 'as', 'a', 'variant', 'of', 'Type: ANY']], 18,  
True)
```

```
['Maastrichtian', 'Regulated', 'by', 'Veldeke-Krink', 'Mestreech']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Maastrichtian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Cori', 'Spoken', 'in', 'Kaduna', 'State', ',', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Cori', 'Language', 'codes', 'ISO', '639-3', 'cry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khunsari', 'Spoken', 'in', 'Iran']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Khunsari', 'Language', 'codes', 'ISO', '639-3', 'kfm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dusun', 'Total', 'speakers', 'second', 'language', ':']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'second', 'language', ':']]], 6, True)

['Dusun', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dusun', 'Region', 'Sabah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dusun', 'Language', 'codes', 'ISO', '639-3', 'dtp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaingang', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kaingang', 'Region', 'S', '\xc3', '\xa3', 'o', 'Paulo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'S', '\xc3',  
'\xa3', 'o', 'Paulo']]], 7, True)

['Kaingang', 'Region', 'Santa', 'Catarina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kaingang', 'Region', 'Paran', '\xc3', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

```

['Kaingang', 'Region', 'Rio', 'Grande', 'do', 'Sul']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rio',
'Grande', 'do', 'Sul']]], 6, True)

['Kaingang', 'Language', 'codes', 'ISO', '639-3', 'kgp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bambassi', 'Spoken', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bambassi', 'Region', 'in', 'Benishangul-Gumuz', ',', 'east', 'of',
'Asosa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'in',
'Benishangul-Gumuz', ',', 'east', 'of', 'Asosa']]], 8, True)

['Bambassi', 'Language', 'codes', 'ISO', '639-3', 'myf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Izi', 'Spoken', 'in', 'Ebonyi', 'State', ',', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'S', '.', 'Type: ANY']]], 7, True)

['Izi', 'Language', 'codes', 'ISO', '639-3', 'izi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khengkha', 'Spoken', 'in', 'Bhutan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Khengkha', 'Writing', 'system', 'Tibetan', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Khengkha', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Khengkha', 'Language', 'codes', 'ISO', '639-3', 'xkf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anuak', 'Spoken', 'in', 'Ethiopia']:

```



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Anuak', 'Region', 'Gambela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Anuak', 'Language', 'codes', 'ISO', '639-3', 'anu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hozo', 'Total', 'speakers', '3000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hozo', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hozo', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Hozo', 'Language', 'codes', 'ISO', '639-3', 'hoz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kachama-Ganjule', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kachama-Ganjule', 'Region', 'on', 'islands', 'in', 'Lake', 'Chamo', 'and', 'Lake', 'Abaya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'on', 'islands', 'in', 'Lake', 'Chamo', 'and', 'Lake', 'Abaya']]], 10, True)

['Kachama-Ganjule', 'Language', 'codes', 'ISO', '639-3', 'kcx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kacipo-Balesi', 'Total', 'speakers', '10', ',', '000', 'in', 'Sudan', '2', ',', '364', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Sudan', 'NUM', ',', 'NUM', 'in', 'Ethiopia']]], ['NUM', '10'], ['NUM', '000'], ['NUM', '2'], ['NUM', '364']], 13, True)

```
['Kacipo-Balesi', 'Spoken', 'in', 'Sudan', ',', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Kacipo-Balesi', 'Language', 'codes', 'ISO', '639-3', 'koe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kafa', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kafa', 'Region', 'in', 'Keficho', 'Shekicho', 'Zone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'in',  
'Keficho', 'Shekicho', 'Zone']]], 6, True)
```

```
['Kafa', 'Language', 'codes', 'ISO', '639-3', 'kbr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'Ancient', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Old', 'Total', 'speakers', '-']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', 'oge']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Komo', 'Total', 'speakers', '10', ',', '000', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '10'], ['NUM', '000']], 8,  
True)
```

```
['Komo', 'Total', 'speakers', '8', ',', '534', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '8'], ['NUM', '534']], 8,  
True)
```

```
['Komo', 'Spoken', 'in', 'Sudan', ',', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```

['Komo', 'Language', 'codes', 'ISO', '639-3', 'xom']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Loarki', 'Spoken', 'in', 'Pakistan', '(', 'some', 'rural', 'areas', 'of',
'Sindh', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Pakistan', '\\\\(', 'some', 'rural', 'areas', 'of', 'Sindh', '\\\\)']], 11,
True)

['Loarki', 'Language', 'codes', 'ISO', '639-2', 'raj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Loarki', 'Language', 'codes', 'ISO', '639-3', 'lrk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngarla', 'Total', 'speakers', '4']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Ngarla', 'Language', 'codes', 'ISO', '639-3', 'nlr']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warumungu', 'Spoken', 'in', 'Northern', 'Territory', ',', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', ',', 'Australia']], 7, True)

['Warumungu', 'Language', 'codes', 'ISO', '639-3', 'wrm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ogba', 'Spoken', 'in', 'Ebonyi', 'and', 'Benue', 'States', ',',
'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ebonyi',
'and', 'Benue', 'States', ',', 'Nigeria']], 9, True)

['Ogba', 'Language', 'codes', 'ISO', '639-3', 'ogc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kilit', 'Spoken', 'in', 'Azerbaijan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

```

['Kilit', 'Spoken', 'in', 'Nakhjavan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kilit', 'Spoken', 'in', 'Ordubad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kilit', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Balkan', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Spoken', 'in', 'Kosovo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Spoken', 'in', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Balkan', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Balkan', 'Language', 'codes', 'ISO', '639-2', 'rom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Balkan', 'Language', 'codes', 'ISO', '639-3', 'rmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Waibuk', 'Region', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Waibuk', 'Language', 'codes', 'ISO', '639-3', 'tmd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pemon', 'Spoken', 'in', 'Pemon', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Pemon', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pemon', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Pemon', 'Spoken', 'in', 'Guyana', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Pemon', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Pemon', 'Language', 'codes', 'ISO', '639-3', 'aoc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Singlish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Changzhou', 'Region', 'Changzhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Changzhou', 'Pronunciation', 's', '\xc9', '\x91', '\xc5', '\x8b', '.', 'z', '\xcc', '\xa5', 'ei', '.', 'wu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 's', '\xc9', '\x91', '\xc5', '\x8b', '.', 'z', '\xcc', '\xa5', 'ei', '.', 'wu']]], 14, True)

['Changzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']]], ['NUM', '639-3']], 10, True)

['Changzhou', 'Language', 'codes', 'ISO', '639-3', 'wu-pil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Xuzhou', 'dialect', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Xuzhou', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Xuzhou', 'Region', 'Xuzhou', ',', 'Jiangsu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Xuzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Coatl\xc3\xa1n', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Coatl\xc3\xa1n', 'Region', 'East', 'central', 'Oaxaca']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Coatl\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'mco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Isthmus', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Isthmus', 'Region', 'Northeastern', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Isthmus', 'Language', 'codes', 'ISO', '639-3', 'mir']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Totontepec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Totontepec', 'Region', 'Northeastern', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Totontepec', 'Language', 'codes', 'ISO', '639-3', 'mto']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tlahuitoltepec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tlahuitoltepec', 'Region', 'Northeastern', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tlahuitoltepec', 'Language', 'codes', 'ISO', '639-3', 'mxp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Juquila', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Juquila', 'Region', 'East', 'central', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Juquila', 'Language', 'codes', 'ISO', '639-3', 'mxq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mazatl\nc3\xa1n', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mazatl\nc3\xa1n', 'Region', 'Eastern', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mazatl\nc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'mzl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['North', 'Region', 'Northeastern', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['North', 'Language', 'codes', 'ISO', '639-3', 'neq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quetzaltepec', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Quetzaltepec', 'Region', 'Northeastern', 'Oaxaca']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Quetzaltepec', 'Language', 'codes', 'ISO', '639-3', 'pxm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Newspeak', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Newspeak', 'Created', 'by', 'George', 'Orwell']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Newspeak', 'Category', '(', 'purpose', ')', 'constructed', 'language']:



[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Newspeak', 'Category', '(', 'purpose', ')', 'Newspeak']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Newspeak', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Forlivese', 'Spoken', 'in', 'Italy']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Forlivese', 'Total', 'speakers', 'n', '/', 'a']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'n',  
'/', 'a']]], 6, True)

['Forlivese', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

["Xo'ini", '(', 'Di', ')', 'Total', 'speakers', '?']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', '\\(', 'Di', '\\)',  
'Total', 'speakers', '?']]], 7, True)

["Xo'ini", 'Region', 'Ejarud', ',', 'Zanjan', 'County']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ejarud', ',',  
'Zanjan', 'County']]], 6, True)

["Xo'ini", 'Spoken', 'in', 'Iran']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Xo'ini", 'Pronunciation', 'xo', '.', 'i', '\\xcb', '\\x90', 'ni', '\\xcb',  
'\\x90']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'xo',  
'.', 'i', '\\xcb', '\\x90', 'ni', '\\xcb', '\\x90']]], 10, True)

["Xo'ini", 'Language', 'codes', 'ISO', '639-2', 'ira']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

["Xo'ini", 'Language', 'codes', 'ISO', '639-3', 'xkc']:  
[[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Idoma', 'Language', 'codes', 'ISO', '639-3', 'ige']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Medieval', 'Spoken', 'in', 'eastern', 'Mediterranean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Medieval', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Medieval', 'Language', 'codes', 'ISO', '639-2', 'gre']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Medieval', 'Language', 'codes', 'ISO', '639-3', 'gkm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ossetian', 'Spoken', 'in', 'Georgia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ossetian', 'Spoken', 'in', 'Turkey']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ossetian', 'Dialects', 'Digor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ossetian', 'Dialects', 'Iron']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Ossetian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Ossetian', 'Language', 'codes', 'ISO', '639-2', 'oss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Ossetian', 'Language', 'codes', 'ISO', '639-3', 'oss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ossetian', 'Language', 'codes', 'Linguasphere', '58-ABB-a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']]], 5, True)

['Ossetian', 'Language', 'codes', 'ISO', '639-1', 'os']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sikule', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sikule', 'Region', 'Aceh', ',', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sikule', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Sikule', 'Language', 'codes', 'ISO', '639-3', 'skh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Simeulue', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Simeulue', 'Region', 'Aceh', ',', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Simeulue', 'Language', 'codes', 'ISO', '639-3', 'smr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hod\xc3\xaf', 'Spoken', 'in', 'central', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Hod\xc3\xaf', 'Total', 'speakers', '300', 'out', 'of', 'an', 'ethnic', 'population', 'of', '767', 'in', 'the', '1970', 'census']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'out', 'of', 'an', 'ethnic', 'population', 'of', 'NUM', 'in', 'the', 'NUM', 'census']], ['NUM', '300'], ['NUM', '767'], ['NUM', '1970']], 15, True)

['Hod\xc3\xaf', 'Language', 'codes', 'ISO', '639-3', 'yau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bagvalal', 'Spoken', 'in', 'Southwestern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Bagvalal', 'Total', 'speakers', '6500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bagvalal', 'Language', 'codes', 'ISO', '639-3', 'kva']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cac\xc3\xa1n', 'Spoken', 'in', 'Diaguita', ',', 'Calchaqu', '\xc3', '\xad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Diaguita', ',', 'Calchaqu', '\xc3', '\xad']], 8, True)

['Cac\xc3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karbi', 'Spoken', 'in', 'Assam', ',', 'Meghalaya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Karbi', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Karbi', 'Language', 'codes', 'ISO', '639-3', 'mju']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Opuu', 'Total', 'speakers', '1', ',', '751', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '1'], ['NUM', '751']], 8, True)

['Opuuo', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Opuuo', 'Region', 'Along', 'the', 'Ethiopia-Sudan', 'border']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Along', 'the', 'Ethiopia-Sudan', 'border']], 6, True)

['Opuuo', 'Language', 'codes', 'ISO', '639-3', 'lgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amdo', 'Spoken', 'in', 'Qinghai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Amdo', 'Spoken', 'in', 'Gansu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Amdo', 'Spoken', 'in', 'Tibet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Amdo', 'Spoken', 'in', 'Sichuan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Amdo', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Amdo', 'Language', 'codes', 'ISO', '639-3', 'adx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bengali', 'Total', 'speakers', '300', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']], ['NUM', '300']], 5, True)

['Bengali', 'Region', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Bengali', 'Region', 'Nepal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bengali', 'Region', 'West', 'Bengal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Region', 'Tripura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bengali', 'Region', 'Andaman', 'and', 'Nicobar', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Andaman', 'and', 'Nicobar', 'Islands']]], 6, True)

['Bengali', 'Region', 'Assam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bengali', 'Region', 'Jharkhand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bengali', 'Region', 'Orissa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bengali', 'Region', 'Mumbai', 'and', 'Bangalore', '.', 'Prominent', 'speakers', 'in', 'Pune']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mumbai', 'and', 'Bangalore', '.', 'Prominent', 'speakers', 'in', 'Pune']]], 10, True)

['Bengali', 'Region', 'Hyderabad', 'and', 'Chennai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bengali', 'Spoken', 'in', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'India', 'and', 'significant', 'communities', 'in', 'UK']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India', 'and', 'significant', 'communities', 'in', 'UK']]], 9, True)

['Bengali', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'Saudi', 'Arabia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bengali', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'United', 'Arab', 'Emirates']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'United', 'Arab', 'Emirates']]], 6, True)

['Bengali', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bengali', 'Writing', 'system', 'Bangla', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Bengali', 'Official', 'language', 'in', 'Bangladesh', ',']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Bengali', 'Regulated', 'by', 'Bangla', 'Academy', '(', 'Bangladesh', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Bangla', 'Academy', '\\(', 'Bangladesh', '\\)']], 8, True)

['Bengali', 'Language', 'codes', 'ISO', '639-2', 'ben']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bengali', 'Language', 'codes', 'ISO', '639-3', 'ben']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bengali', 'Language', 'codes', 'Linguasphere', '59-AAF-uk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Type: ANY']], 5, True)

['Bengali', 'Language', 'codes', 'Linguasphere', 'Global', 'extent', 'of', 'Bengali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'Linguasphere', 'Global', 'extent', 'of', 'Type: ANY']], 8, True)

['Bengali', 'Language', 'codes', 'ISO', '639-1', 'bn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['British', 'Spoken', 'in', 'Roman', 'Britain', ',', 'Anglo-Saxon', 'England']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Roman', 'Britain', ',', 'Anglo-Saxon', 'England']], 8, True)

['British', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Shao-Jiang', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Shao-Jiang', 'Spoken', 'in', 'Southern', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Shao-Jiang', 'Region', 'eastern', 'Nanping', 'Prefecture', ',', 'Fujian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'eastern', 'Nanping', 'Prefecture', ',', 'Fujian']], 7, True)



```
['Shao-Jiang', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Danish', 'Signed', 'in', 'Denmark']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Mota', 'Total', 'speakers', '750']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Mota', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Mota', 'Region', 'Mota', 'island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Mota', 'Language', 'codes', 'ISO', '639-3', 'mtt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Biloxi', 'Language', 'extinction', '1930s']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Biloxi', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Biloxi', 'Region', 'Louisiana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Biloxi', 'Language', 'codes', 'ISO', '639-3', 'bll']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bangi', 'Spoken', 'in', 'Republic', 'of', 'Congo', ',', 'Democratic',  
'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Congo', ',', 'Democratic', 'Republic', 'of', 'the',  
'Congo']], 12, True)
```

```
['Bangi', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Bangi', 'Language', 'codes', 'ISO', '639-3', 'bni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nias', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Nias', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Nias', 'Language', 'codes', 'ISO', '639-2', 'nia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Nias', 'Language', 'codes', 'ISO', '639-3', 'nia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Warao', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Warao', 'Spoken', 'in', 'Guyana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Warao', 'Spoken', 'in', 'Suriname']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Warao', 'Language', 'codes', 'ISO', '639-3', 'wba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Leti', 'Total', 'speakers', '7500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Leti', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Leti', 'Region', 'Leti', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Leti', 'Language', 'codes', 'ISO', '639-3', 'lti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kembra', 'Region', 'Jayawijaya', 'Kabupaten']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kembra', 'Region', 'Okbibab', 'Kecamatan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kembra', 'Region', 'east', 'of', 'the', 'Sogber', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'east', 'of', 'the', 'Sogber', 'River']]], 7, True)

['Kembra', 'Language', 'codes', 'ISO', '639-3', 'xkw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Olu'bo", 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Olu'bo", 'Language', 'codes', 'ISO', '639-3', 'lul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dumpu', 'Region', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dumpu', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dumpu', 'Language', 'codes', 'ISO', '639-3', 'wtf', 'External', 'link']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wtf', 'External', 'link']], ['NUM', '639-3']], 8, True)

```
['Folopa', 'Spoken', 'in', 'Baimuru', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Folopa', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Folopa', 'Language', 'codes', 'ISO', '639-3', 'ppo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Conambo', 'Language', 'extinction', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Conambo', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Conambo', 'Language', 'codes', 'ISO', '639-3', 'avs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Arhuaco', 'Spoken', 'in', 'Colombia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Arhuaco', 'Spoken', 'in', 'Arhuacos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Arhuaco', 'Language', 'codes', 'ISO', '639-3', 'arh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Erromintxela', 'Spoken', 'in', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Erromintxela', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Erromintxela', 'Region', 'Basque', 'Country']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Erromintxela', 'Language', 'codes', 'ISO', '639-3', 'emx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Budu', 'Spoken', 'in', 'Orientale', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Budu', 'Language', 'codes', 'ISO', '639-3', 'buu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["S'gaw", 'Spoken', 'in', 'Burma', ',', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

["S'gaw", 'Region', 'Throughout', 'the', 'Far', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Throughout',  
'the', 'Far', 'Type: ANY']]], 6, True)

["S'gaw", 'Language', 'codes', 'ISO', '639-3', 'ksw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Makonde', 'Spoken', 'in', 'Tanzania', ',', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Makonde', 'Dialects', 'Matembwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Makonde', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Makonde', 'Language', 'codes', 'ISO', '639-3', 'yao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karamay', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

```
['Nimbari', 'Spoken', 'in', 'northern', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Nimbari', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
\xe2', '\x80', '\x93']]], 6, True)
```

```
['Nimbari', 'Language', 'codes', 'ISO', '639-3', 'nmr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kam', 'Spoken', 'in', 'eastern', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Kam', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
\xe2', '\x80', '\x93']]], 6, True)
```

```
['Kam', 'Language', 'codes', 'ISO', '639-3', 'kdx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Suki', 'Region', 'Western', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Suki', 'Language', 'codes', 'ISO', '639-3', 'sui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mednyj', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mednyj', 'Region', 'Commander', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mednyj', 'Language', 'codes', 'ISO', '639-3', 'mud']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kumbainggar', 'Spoken', 'in', 'New', 'South', 'Wales', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'Type: ANY', 'Wales', ',', 'Australia']]], 8, True)

['Kumbainggar', 'Language', 'codes', 'ISO', '639-3', 'kgs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kelabit', 'Spoken', 'in', 'Borneo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kelabit', 'Language', 'codes', 'ISO', '639-3', 'kzi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['K\xc4\x93len', 'Category', '(', 'sources', ')', 'a', 'priori',  
'language', ',', 'consciously', 'rejecting', 'linguistic', 'universals',  
'to', 'create', 'an', 'alien', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'a', 'priori', 'language', ',', 'consciously',  
'rejecting', 'linguistic', 'universals', 'to', 'create', 'an', 'alien',  
'language']]], 18, True)

['K\xc4\x93len', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['K\xc4\x93len', 'Setting', 'and', 'usage', 'alien', 'species', '(', 'the',  
'K', '\xc4', '\x93', 'le', '\xc3', '\xb1', 'i', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'alien', 'species', '\\(', 'the', 'K', '\xc4', '\x93', 'le',  
'\xc3', '\xb1', 'i', '\\)']]], 16, True)

['K\xc4\x93len', 'Created', 'by', 'Sylvia', 'Sotomayor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['K\xc4\x93len', 'Category', '(', 'purpose', ')', 'constructed',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['K\xc4\x93len', 'Category', '(', 'purpose', ')', 'artistic', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['K\xc4\x93len', 'Category', '(', 'purpose', ')', 'K', '\xc4', '\x93',

```

'len']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'K', '\\xc4', '\\x93', 'len']]], 9, True)

['K\\xc4\\x93len', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',
'\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Malagasy', 'Spoken', 'in', 'Madagascar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Malagasy', 'Spoken', 'in', 'Comoros']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Malagasy', 'Spoken', 'in', 'Mayotte']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-2', 'mlg', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'mlg', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-2', 'mlg', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'mlg', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'mlg', '\\xe2', '\\x80',
'\\x93', 'Macrolanguage']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'mlg', '\\xe2', '\\x80', '\\x93', 'Macrolanguage']], ['NUM',
'639-3']], 10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'individual', 'codes']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'individual', 'code']], ['NUM', '639-3']], 7, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', ':', 'xmv', '\\xe2',
'\\x80', '\\x93', 'Antankarana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'xmv', '\\xe2', '\\x80', '\\x93', 'Antankarana']], ['NUM',
'639-3']], 11, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'bhr', '\\xe2', '\\x80',

```



```

'\x93', 'Bara']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'bhr', '\xe2', '\x80', '\x93', 'Bara']], ['NUM', '639-3']],
10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'msh', '\xe2', '\x80',
'\x93', 'Masikoro']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'msh', '\xe2', '\x80', '\x93', 'Masikoro']], ['NUM',
'639-3']], 10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'skg', '\xe2', '\x80',
'\x93', 'Sakalava']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'skg', '\xe2', '\x80', '\x93', 'Sakalava']], ['NUM',
'639-3']], 10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'tdx', '\xe2', '\x80',
'\x93', 'Tandroy-Mafahaly']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'tdx', '\xe2', '\x80', '\x93', 'Tandroy-Mafahaly']], ['NUM',
'639-3']], 10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'txy', '\xe2', '\x80',
'\x93', 'Tanosy']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'txy', '\xe2', '\x80', '\x93', 'Tanosy']], ['NUM', '639-3']],
10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'xmw', '\xe2', '\x80',
'\x93', 'Tsimihety']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'xmw', '\xe2', '\x80', '\x93', 'Tsimihety']], ['NUM',
'639-3']], 10, True)

['Malagasy', 'Language', 'codes', 'ISO', '639-1', 'mg']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Zhongshan', 'Spoken', 'in', 'China', ',', 'United', 'States', '(',
'California', ')']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'China',
',', 'United', 'States', '\\(', 'California', '\\)']]], 10, True)

['Zhongshan', 'Region', 'Guangdong', 'Province']]:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Zhongshan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Zhongshan', 'Language', 'codes', 'ISO', '639-3', 'nan-lon', '(', 'Longdu', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'nan-lon', '\\(', 'Longdu', '\\)']], ['NUM', '639-3']], 9, True)

['Zhenan', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhenan', 'Region', 'Zhejiang', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhenan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Zhenan', 'Language', 'codes', 'ISO', '639-3', 'nan-zhe']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Longyan', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Longyan', 'Region', 'Fujian', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Longyan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Quanzhou', 'dialect', 'Total', 'speakers', 'over', '7', 'million']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',

'speakers', 'over', 'NUM', 'million']], ['NUM', '7']], 7, True)

['Quanzhou', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quanzhou', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quanzhou', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quanzhou', 'Spoken', 'in', 'Burma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quanzhou', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quanzhou', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Quanzhou', 'Spoken', 'in', 'Hong', 'Kong', 'and', 'Macau']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Hong', 'Kong', 'and', 'Macau']]], 7, True)

['Quanzhou', 'Region', 'Southern', 'Fujian', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Quanzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Zhangzhou', 'dialect', 'Total', 'speakers', 'Over', '4', 'million', '(', 'no', 'recent', 'data', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total', 'speakers', 'Over', 'NUM', 'million', '\\(', 'no', 'recent', 'data', '\\)']]], ['NUM', '4']], 12, True)

['Zhangzhou', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Zhangzhou', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhangzhou', 'Spoken', 'in', 'Singapore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhangzhou', 'Spoken', 'in', 'Taiwan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhangzhou', 'Spoken', 'in', 'and', 'Hong', 'Kong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'Hong', 'Kong']]], 6, True)

['Zhangzhou', 'Region', 'Southern', 'Fujian', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Zhangzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Zhanjiang', 'dialect', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhanjiang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zhanjiang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Shantou', 'dialect', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shantou', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Shantou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Shantou', 'Language', 'codes', 'ISO', '639-3', 'nan-sha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Medan', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Medan', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Medan', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Medan', 'Language', 'codes', 'ISO', '639-3', 'nan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Medan', 'Language', 'codes', 'ISO', '639-1', 'zh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Sanxiang', 'dialect', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sanxiang', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sanxiang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Nyangatom', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Nyangatom', 'Region', 'Omo', 'River', 'region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)

['Nyangatom', 'Writing', 'system', 'none']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nyangatom', 'Language', 'codes', 'ISO', '639-3', 'nnj']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Hui'an", 'dialect', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Hui'an", 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

["Hui'an", 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Nanlang', 'dialect', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nanlang', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nanlang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Telefol', 'Region', 'Sandaun', 'Province', ',', 'Telefomin', 'District']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sandaun',
'Province', ',', 'Telefomin', 'District']]], 7, True)

['Telefol', 'Language', 'codes', 'ISO', '639-3', 'tlf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uropi', 'Total', 'speakers', '?', '?', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '?',

```

'?', '?']]], 6, True)

['Uropi', 'Category', '(', 'purpose', ')', 'constructed', 'language']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Uropi', 'Category', '(', 'purpose', ')', 'Uropi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY']]], 6, True)

['Uropi', 'Created', 'by', 'Jo', '\\xc3', '\\xab', 'l', 'Landais']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Jo',  
'\\xc3', '\\xab', 'l', 'Landais']]], 8, True)

['Uropi', 'Category', '(', 'sources', ')', 'based', 'on', 'Indo-European',  
'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)'], 'based', 'on', 'Indo-European', 'languages']]], 9, True)

['Uropi', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Fali', 'Total', 'speakers', 'few']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fali', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fali', 'Region', 'Taraba', 'State']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fali', 'Language', 'codes', 'ISO', '639-3', 'fah']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mbire', 'Total', 'speakers', '200']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mbire', 'Spoken', 'in', 'Ivory', 'Coast']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

```

['Mbire', 'Region', 'Marabadiassa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Mbire', 'Language', 'codes', 'ISO', '639-3', 'mka']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ega', 'Total', 'speakers', '2500']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ega', 'Spoken', 'in', 'Ivory', 'Coast']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ega', 'Region', 'near', 'Gly']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ega', 'Language', 'codes', 'ISO', '639-3', 'ega']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ukaan', 'Total', 'speakers', '18', ',', '000', 'in', '1973']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '18'], ['NUM', '000']], 8, True)

['Ukaan', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ukaan', 'Region', 'Ondo', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ukaan', 'Language', 'codes', 'ISO', '639-3', 'kcf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fali', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```



```

['Fali', 'Region', 'North']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Fali', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Akpes', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Akpes', 'Region', 'Ondo', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Akpes', 'Language', 'codes', 'ISO', '639-3', 'ibe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ayere\xe2\x80\x93Ahan', 'Total', 'speakers', '3000', '(', 'Ayere', ')', ',', '300', '(', 'Ahan', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\\\(', 'Ayere', '\\\\)', ',', 'NUM', '\\\\(', 'Ahan', '\\\\)']], ['NUM', '3000'], ['NUM', '300']], 12, True)

['Ayere\xe2\x80\x93Ahan', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ayere\xe2\x80\x93Ahan', 'Region', 'Ondo', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ayere\xe2\x80\x93Ahan', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ayere\xe2\x80\x93Ahan', 'Language', 'codes', 'ISO', '639-3', ':', 'ahn', '\xe2', '\x80', '\x93', 'Ahaan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ahn', '\xe2', '\x80', '\x93', 'Ahaan']], ['NUM', '639-3']], 11, True)

['Ayere\xe2\x80\x93Ahan', 'Language', 'codes', 'ISO', '639-3', 'aye', '\xe2', '\x80', '\x93', 'Ayere']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'aye', '\xe2', '\x80', '\x93', 'Ayere']], ['NUM', '639-3']], 10, True)

['(North)', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(North)', 'Region', 'Ondo', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(North)', 'Language', 'codes', 'ISO', '639-3', 'aqq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oko', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oko', 'Region', 'Kogi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Oko', 'Language', 'codes', 'ISO', '639-3', 'aqq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Spoken', 'in', 'Paraguay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Western', 'Language', 'codes', 'ISO', '639-2', 'grn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'gnw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sukuma', 'Spoken', 'in', 'Tanzania', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Sukuma', 'Spoken', 'in', 'Sukuma', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Sukuma', 'Language', 'codes', 'ISO', '639-3', 'suk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyamwezi', 'Spoken', 'in', 'Tanzania', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Nyamwezi', 'Spoken', 'in', 'Nyamwezi', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Nyamwezi', 'Language', 'codes', 'ISO', '639-3', 'nym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Datooga', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Datooga', 'Region', 'Great', 'Rift', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Datooga', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Datooga', 'Language', 'codes', 'ISO', '639-3', 'tcc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cajamarca', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Cajamarca', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Cajamarca', 'Language', 'codes', 'ISO', '639-3', 'qvc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inkawasi-Ka\xc3\xblaris', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']], 6, True)

['Inkawasi-Ka\xc3\xblaris', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inkawasi-Ka\xc3\xblaris', 'Language', 'codes', 'ISO', '639-3', 'quf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gwandara', 'Spoken', 'in', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Gwandara', 'Total', 'speakers', '30', 'thousand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'thousand']], ['NUM', '30']], 5, True)

['Gwandara', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Gwandara', 'Language', 'codes', 'ISO', '639-3', 'gwn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chachapoyas', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']], 6, True)

['Chachapoyas', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Chachapoyas', 'Language', 'codes', 'ISO', '639-3', 'quk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lamas', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per',  
'\xc3', '\xba']]], 6, True)

['Lamas', 'Language', 'codes', 'ISO', '639-2', 'que']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lamas', 'Language', 'codes', 'ISO', '639-3', 'qvs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kele', 'Spoken', 'in', 'Republic', 'of', 'Zaire', ',', 'Democratic',  
'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Zaire', ',', 'Democratic', 'Republic', 'of', 'the',  
'Congo']]], 12, True)

['Kele', 'Region', 'Kisangani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kele', 'Language', 'codes', 'ISO', '639-3', 'khy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Noanam\xc3\xa1', 'Spoken', 'in', 'Panama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Noanam\xc3\xa1', 'Total', 'speakers', '6000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Noanam\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'noa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mattole', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mattole', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Mattole', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Mattole', 'Language', 'codes', 'ISO', '639-3', 'mvp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wailaki', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wailaki', 'Spoken', 'in', 'USA']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wailaki', 'Region', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wailaki', 'Language', 'codes', 'ISO', '639-3', 'wlk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kabalai', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kabalai', 'Region', 'central']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kabalai', 'Language', 'codes', 'ISO', '639-3', 'kvf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Emiliano', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Emiliano', 'Total', 'speakers', '2', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', 'NUM', 'million']]], ['NUM', '2']], 5, True)

```

['Emiliano', 'Language', 'codes', 'ISO', '639-3', 'egl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Romagnolo', 'Spoken', 'in', 'Italy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Romagnolo', 'Spoken', 'in', 'San', 'Marino']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Romagnolo', 'Total', 'speakers', '1', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'NUM', 'million']], ['NUM', '1']], 5, True)

['Romagnolo', 'Language', 'codes', 'ISO', '639-3', 'rgn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gorgotoqui', 'Spoken', 'in', 'eastern', 'Bolivia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Gorgotoqui', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inga', 'Spoken', 'in', 'Colombia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Inga', 'Language', 'codes', 'ISO', '639-2', 'que']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Inga', 'Language', 'codes', 'ISO', '639-3', 'either']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Inga', 'Language', 'codes', 'ISO', '639-3', 'inj', '\xe2', '\x80',
'\x93', 'Jungle', 'Inga']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'inj', '\xe2', '\x80', '\x93', 'Jungle', 'Type: ANY']],
['NUM', '639-3']], 11, True)

```

['Gutian', 'Spoken', 'in', 'Zagros', 'Mountains', '?', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Zagros',  
'Mountains', '?', 'Ethnicity']]], 7, True)

['Gutian', 'Spoken', 'in', 'Gutian', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Gutian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Mbugu', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mbugu', 'Region', 'Usambara', 'Mountains']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mbugu', 'Language', 'codes', 'ISO', '639-3', 'mhd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2',  
'\x80', '\x93']]], 6, True)

['Old', 'Region', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Official', 'language', 'in', 'dead']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Old', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-2', 'pro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'pro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Urapmin', 'Region', 'Sandaun', 'Province', ',', 'Telefomin', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sandaun',  
'Province', ',', 'Telefomin', 'District']]]], 7, True)

['Urapmin', 'Language', 'codes', 'ISO', '639-3', 'urm']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tifal', 'Region', 'Sandaun', 'Province', ',', 'Telefomin', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sandaun',  
'Province', ',', 'Telefomin', 'District']]]], 7, True)

['Tifal', 'Language', 'codes', 'ISO', '639-3', 'tif']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Swardspeak', 'Spoken', 'in', 'Philippines']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['Swardspeak', 'Total', 'speakers', 'no', 'estimate', 'available']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'no', 'estimate', 'available']]]], 6, True)

['Swardspeak', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Esimbi', 'Spoken', 'in', 'Southwest', 'Cameroon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Esimbi', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]]], 6, True)

['Esimbi', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Esimbi', 'Language', 'codes', 'ISO', '639-3', 'ags']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Borbecksch', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Borbecksch', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Borbecksch', 'Region', 'Essen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Borbecksch', 'Region', 'Oberhausen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Borbecksch', 'Region', 'Bottrop', '(', 'Ruhr', 'Area', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Bottrop', '\\(', 'Ruhr', 'Area', '\\)']], 7, True)

['Borbecksch', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Quinigua', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Quinigua', 'Spoken', 'in', 'NE', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Quinigua', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Quinigua', 'Language', 'codes', 'ISO', '639-3', 'qr4']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kh\\xc3\\xbc\\n', 'Language', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Kh\\xc3\\xbc\\n', 'Region', 'Kengtung']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kh\\xc3\\xbc\\n', 'Spoken', 'in', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kh\\xc3\\xbc\\n', 'Spoken', 'in', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kh\\xc3\\xbc\\n', 'Spoken', 'in', 'Laos']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kh\\xc3\\xbc\\n', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kh\\xc3\\xbc\\n', 'Pronunciation', '/', 'tai', 'k', '\\xca', '\\xb0', '\\xc9', '\\xaf', 'n', '/', 'missing', 'tone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', 'tai', 'k', '\\xca', '\\xb0', '\\xc9', '\\xaf', 'n', '/', 'missing', 'tone']]], 13, True)

['Kh\\xc3\\xbc\\n', 'Language', 'codes', 'ISO', '639-3', 'kkh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Odu\\xa1', 'Spoken', 'in', 'Rivers', 'State', '(', 'Nigeria', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Rivers', 'State', '\\\\(', 'Nigeria', '\\\\)']]], 8, True)

['Odu\\xa1', 'Total', 'speakers', '18', ',', '000', '(', 'Ethnologue', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']]], ['NUM', '18'], ['NUM', '000']], 9, True)

['Odu\\xa1', 'Language', 'codes', 'ISO', '639-3', 'odu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yem\\xa1', 'Spoken', 'in', 'Ethiopia']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Yemsa', 'Writing', 'system', 'Ethiopic', '/', 'Latin']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Yemsa', 'Language', 'codes', 'ISO', '639-3', 'jnj']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Arabana', 'Spoken', 'in', 'South', 'Australia']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Arabana', 'Total', 'speakers', '6']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Arabana', 'Language', 'codes', 'ISO', '639-3', 'ard']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tangale', 'Spoken', 'in', 'Nigeria']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tangale', 'Language', 'codes', 'ISO', '639-3', 'tan']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tangale', 'Language', 'codes', 'ISO', '639-3', 'Ethnic', 'territories', 'of', 'the', 'Tangale-speaking', 'people', 'in', 'Nigeria', 'in', 'green']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Ethnic', 'territories', 'of', 'the', 'Tangale-speaking', 'people', 'in', 'Nigeria', 'in', 'green']]], ['NUM', '639-3']], 15, True)
```

```
['Tima', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']]], 7, True)
```

```
['Tima', 'Total', 'speakers', '(', 'unknown', ')']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\\\(', 'unknown', '\\\\)']]], 6, True)
```

```
['Tima', 'Language', 'codes', 'ISO', '639-3', 'tms']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nimanbur', 'Spoken', 'in', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nimanbur', 'Total', 'speakers', '0']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nimanbur', 'Language', 'codes', 'ISO', '639-3', 'nmp']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Narangga', 'Language', 'extinction', '1936']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Narangga', 'Spoken', 'in', 'South', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Narangga', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Narangga', 'Language', 'codes', 'ISO', '639-3', 'nnr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luwo', 'Spoken', 'in', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Luwo', 'Region', 'Bahr', 'el', 'Ghazal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Luwo', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Luwo', 'Language', 'codes', 'ISO', '639-3', 'lwo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nukunu', 'Spoken', 'in', 'South', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Nukunu', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nukunu', 'Language', 'codes', 'ISO', '639-3', 'nnv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oroch', 'Total', 'speakers', '260']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Oroch', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Oroch', 'Region', 'Russian', 'Far', 'East']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Oroch', 'Language', 'codes', 'ISO', '639-3', 'oac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alorese', 'Spoken', 'in', 'Indonesia', ',', '']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', '']], 5, True)

['Alorese', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Alorese', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Alorese', 'Language', 'codes', 'ISO', '639-3', 'aol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Birhor', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Birhor', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Birhor', 'Language', 'codes', 'ISO', '639-3', 'biy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wulguru', 'Language', 'extinction', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wulguru', 'Spoken', 'in', 'Townsville', ',', 'Queensland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Wulguru', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Wulguru', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Wulguru', 'Language', 'codes', 'ISO', '639-3', 'agu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saliba', 'Spoken', 'in', 'Colombia', 'and', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Saliba', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Saliba', 'Language', 'codes', 'ISO', '639-3', 'slc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piaroa', 'Spoken', 'in', 'Colombia', 'and', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Piaroa', 'Dialects', 'Maco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Piaroa', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Piaroa', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Piaroa', 'Language', 'codes', 'ISO', '639-3', ':', 'pid', '\xe2', '\x80', '\x93', 'Piaroa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'pid', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 11, True)

['Piaroa', 'Language', 'codes', 'ISO', '639-3', 'wpc', '\xe2', '\x80', '\x93', 'Maco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wpc', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Maco', 'Spoken', 'in', 'Colombia', 'and', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'and', 'Type: ANY']]], 6, True)

['Maco', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Maco', 'Language', 'codes', 'ISO', '639-3', 'wpc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Enawene', 'Spoken', 'in', 'Mato', 'Grosso', ',', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mato', 'Grosso', ',', 'Brazil']]], 7, True)

['Enawene', 'Total', 'speakers', '350']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



['Enawene', 'Language', 'codes', 'ISO', '639-3', 'unk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fergana', 'Total', 'speakers', 'no', 'information', ',', 'possibly',  
'extinct', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'no', 'information', ',', 'possibly', 'extinct', 'language']]], 9, True)

['Fergana', 'Language', 'codes', 'ISO', '639-3', 'qwm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hidatsa', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Hidatsa', 'Region', 'North', 'Dakota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hidatsa', 'Region', 'Montana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Hidatsa', 'Region', 'South', 'Dakota']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hidatsa', 'Language', 'codes', 'ISO', '639-3', 'hid']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Olukonjo', 'Spoken', 'in', 'Uganda', 'and', 'Democratic', 'Republic',  
'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Uganda',  
'and', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 10, True)

['Olukonjo', 'Language', 'codes', 'ISO', '639-3', 'koo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Quenya', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death', 'of',  
'Tolkien']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',

'None', ',', 'after', 'the', 'death', 'of', 'Tolkien']]], 10, True)

['Quenya', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Quenya', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Quenya', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Quenya', 'Category', '(', 'purpose', ')', 'Quenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY']]]], 6, True)

['Quenya', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',  
'.', 'R', '.', 'R', '.', 'Tolkien']]]], 10, True)

['Quenya', 'Setting', 'and', 'usage', 'The', 'fictional', 'world', 'of',  
'Middle-earth']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'The', 'fictional', 'world', 'of', 'Middle-earth']]]], 9, True)

['Quenya', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Sindarin', 'Total', 'speakers', 'None', ',', 'after', 'the', 'death',  
'of', 'Tolkien']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'None', ',', 'after', 'the', 'death', 'of', 'Tolkien']]]], 10, True)

['Sindarin', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Sindarin', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'purpose', '\\\\)', 'Type: ANY', 'language']]]], 7, True)

['Sindarin', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',

'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Sindarin', 'Category', '(', 'purpose', ')', 'Sindarin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Sindarin', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',  
'.', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Sindarin', 'Setting', 'and', 'usage', 'The', 'fictional', 'world', 'of',  
'Middle-earth']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'The', 'fictional', 'world', 'of', 'Middle-earth']]], 9, True)

['Sindarin', 'Language', 'codes', 'ISO', '639-3', 'sjn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Meadow', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\\xe2', '\\x80', '\\x93']]], 6, True)

['Meadow', 'Writing', 'system', 'Cyrillic', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Meadow', 'Language', 'codes', 'ISO', '639-3', 'mhr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pannonian', 'Spoken', 'in', 'Pannonia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Pannonian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Pannonian', 'Language', 'codes', 'ISO', '639-3', 'lat-pan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Urdu', 'Spoken', 'in', 'Fiji']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Urdu', 'Spoken', 'in', 'with', 'significant', 'minorities', 'in', 'the', 'Australia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'with', 'significant', 'minorities', 'in', 'the', 'Australia']]], 9, True)

['Urdu', 'Spoken', 'in', 'New', 'Zealand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Urdu', 'Spoken', 'in', 'USA']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Urdu', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Urdu', 'Total', 'speakers', 'Unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Urdu', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Taihu', 'dialect', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taihu', 'Region', 'South', 'Jiangsu', 'province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Taihu', 'Region', 'North', 'Zhejiang', 'province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Taihu', 'Region', 'southeastern', 'Anhui']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Taihu', 'Region', 'and', 'Shanghai', '.', 'Linguistic', 'exclave', 'in', 'Cangnan', 'county', 'in', 'southern', 'Zhejiang', 'province']:

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Shanghai', '.', 'Linguistic', 'exclave', 'in', 'Cangnan', 'county', 'in',  
'southern', 'Zhejiang', 'province']]], 14, True)
```

```
['Taihu', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Taihu', 'Language', 'codes', 'ISO', '639-3', 'wu-tai']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jiangshan', 'Region', 'Jiangshan', 'city']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jiangshan', 'Region', 'Quzhou', 'prefecture']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jiangshan', 'Region', 'Zhejiang', 'province']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jiangshan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Jiangshan', 'Language', 'codes', 'ISO', '639-3', 'wu-chh']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Chuqu', 'Wu', 'Total', 'speakers']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Chuqu', 'Region', 'Zhejiang']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Chuqu', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Chuqu', 'Language', 'codes', 'ISO', '639-3', 'wu-chu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kota', 'Spoken', 'in', 'Gabon', ',', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Gabon', ',', 'Republic', 'of', 'the', 'Congo']]], 9, True)

['Kota', 'Language', 'codes', 'ISO', '639-3', 'koq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Usarufa', 'Spoken', 'in', 'Okapa', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Usarufa', 'Language', 'codes', 'ISO', '639-3', 'usa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sami', 'Spoken', 'in', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sami', 'Spoken', 'in', 'Norway']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sami', 'Spoken', 'in', 'Sweden']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sami', 'Spoken', 'in', 'and', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sami', 'Region', 'S', '\xc3', '\xa1', 'pmi', '(', 'Lapland', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'S', '\xc3', '\xa1', 'pmi', '\\(', 'Lapland', '\\)']]], 9, True)

['Sami', 'Official', 'language', 'in', 'in', 'some', 'parts', 'of', 'Norway', ';', 'recognized', 'as', 'a', 'minority', 'language', 'in',

'several', 'municipalities', 'of', 'Sweden', 'and', 'Finland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'in', 'some', 'parts', 'of', 'Norway', ';', 'recognized', 'as', 'a',  
'minority', 'language', 'in', 'several', 'municipalities', 'of', 'Sweden',  
'and', 'Finland']]], 22, True)

['Sami', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-2', 'sma']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-2', 'sme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-2', 'smi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-2', 'smj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-2', 'smn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-2', 'sms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sami', 'Language', 'codes', 'ISO', '639-3', ':', 'sia', '\xe2', '\x80',  
'\x93', 'Akkala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'sia', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 11, True)

['Sami', 'Language', 'codes', 'ISO', '639-3', 'sjd', '\xe2', '\x80',  
'\x93', 'Kildin']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sjd', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'sjk', '\xe2', '\x80', '\x93', 'Kemi']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sjk', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'sjt', '\xe2', '\x80', '\x93', 'Ter']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sjt', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'smn', '\xe2', '\x80', '\x93', 'Inari']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'smn', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'sms', '\xe2', '\x80', '\x93', 'Skolt']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sms', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'sju', '\xe2', '\x80', '\x93', 'Ume']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sju', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'sje', '\xe2', '\x80', '\x93', 'Pite']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sje', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```

```
['Sami', 'Language', 'codes', 'ISO', '639-3', 'sme', '\xe2', '\x80', '\x93', 'North']:
```

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'sme', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)
```



```

['Sami', 'Language', 'codes', 'ISO', '639-3', 'smj', '\xe2', '\x80',
'\x93', 'Lule']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'smj', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Sami', 'Language', 'codes', 'ISO', '639-3', 'sma', '\xe2', '\x80',
'\x93', 'South']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'sma', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Bolak', 'Total', 'speakers', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bolak', 'Date', 'founded', '1899']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bolak', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Bolak', 'Category', '(', 'purpose', ')', 'Bolak']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Bolak', 'Created', 'by', 'L', '\xc3', '\xa9', 'on', 'Bollack']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'L',
'\xc3', '\xa9', 'on', 'Bollack']]], 8, True)

['Bolak', 'Language', 'codes', 'ISO', '639-2', 'art']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bolak', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Idu', 'Language', 'codes', 'ISO', '639-3', 'clk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yameo', 'Language', 'extinction', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Yameo', 'Spoken', 'in', 'western', 'Amazon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Yameo', 'Language', 'codes', 'ISO', '639-3', 'yme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tequiraca', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tequiraca', 'Total', 'speakers', 'a', 'few']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a',  
'few']]], 5, True)

['Tequiraca', 'Language', 'codes', 'ISO', '639-3', 'ash']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hupd\xc3\xab', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hupd\xc3\xab', 'Total', 'speakers', '1', ',', '400', 'in', '1995']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '1'], ['NUM', '400']], 8,  
True)

['Hupd\xc3\xab', 'Language', 'codes', 'ISO', '639-3', 'jup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nung', 'Spoken', 'in', 'Fugong', 'County', ',', 'Myanmar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Fugong',  
'County', ',', 'Type: ANY']]], 7, True)

['Nung', 'Language', 'codes', 'ISO', '639-3', 'nun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sakha', 'Region', 'Sakha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Sakha', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Sakha', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'Japan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'Korea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sakha', 'Spoken', 'in', 'United', 'Kingdom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Sakha', 'Spoken', 'in', 'Sweden']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sakha', 'Spoken', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sakha', 'Writing', 'system', 'Cyrillic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Sakha', 'Official', 'language', 'in', 'Sakha', 'Republic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Sakha', 'Regulated', 'by', 'No', 'official', 'regulation']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Sakha', 'Language', 'codes', 'ISO', '639-2', 'sah']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Sakha', 'Language', 'codes', 'ISO', '639-3', 'sah']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['High', 'Speech', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['High', 'Setting', 'and', 'usage', 'Fictional', 'Mid-World', ':', 'The', 'Dark', 'Tower', 'Series']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Fictional', 'Mid-World', ':', 'The', 'Dark', 'Tower', 'Series']]], 11, True)

['High', 'Created', 'by', 'Stephen', 'King']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['High', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['High', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['High', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['High', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Aromanian', 'Region', 'Southeastern', 'Europe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aromanian', 'Spoken', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aromanian', 'Spoken', 'in', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aromanian', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aromanian', 'Spoken', 'in', 'Republic', 'of', 'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Republic', 'of', 'Macedonia']]], 6, True)

['Aromanian', 'Spoken', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aromanian', 'Spoken', 'in', 'Bulgaria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Aromanian', 'Official', 'language', 'in', ':', 'Republic', 'of',  
'Macedonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', ':', 'Republic', 'of', 'Macedonia']]], 8, True)

['Aromanian', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:

ANY', 'of', 'Type: ANY']]], 6, True)

['Aromanian', 'Language', 'codes', 'ISO', '639-2', 'rup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Aromanian', 'Language', 'codes', 'ISO', '639-3', 'rup']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aromanian', 'Language', 'codes', 'Linguasphere', '51-AAD-ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Nicaraguan', 'Spoken', 'in', 'Nicaragua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nicaraguan', 'Total', 'speakers', '6', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '6']], 5, True)

['Nicaraguan', 'Writing', 'system', 'Latin', '(', 'Spanish', 'variant',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']]], 8, True)

['Nicaraguan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Bafanji', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bafanji', 'Language', 'codes', 'ISO', '639-3', 'bfj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Akw\xc3\xa1wa', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Akw\xc3\xa1wa', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Akw\xc3\xa1wa', 'Language', 'codes', 'ISO', '639-3', 'asu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zeem', 'Spoken', 'in', 'Bauchi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Zeem', 'Language', 'codes', 'ISO', '639-3', 'zua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mwani', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mwani', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mwani', 'Language', 'codes', 'ISO', '639-3', 'wmw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mwani', 'Language', 'codes', 'ISO', '639-3', 'The', 'area', 'where',  
'Kimwani', 'is', 'spoken', 'is', 'in', 'gray']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'The', 'area', 'where', 'Kimwani', 'is', 'spoken', 'is',  
'in', 'gray']], ['NUM', '639-3']], 14, True)

['Kemihua', 'Spoken', 'in', 'Indochina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kemihua', 'Language', 'codes', 'ISO', '639-3', 'kfj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chilisso', 'Spoken', 'in', 'Eastern', 'Kohistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Chilisso', 'Language', 'codes', 'ISO', '639-3', 'clh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Takuu', 'Language', 'codes', 'ISO', '639-3', 'nho']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Djerbi', 'Spoken', 'in', 'Tunisia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Djerbi', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Djerbi', 'Language', 'codes', 'ISO', '639-3', 'jbn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abip\\xc3\\xb3n', 'Spoken', 'in', 'Argentina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Abip\\xc3\\xb3n', 'Language', 'codes', 'ISO', '639-3', 'axb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cia-Cia', 'Spoken', 'in', 'Buton', 'Island', ',', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Buton',  
'Island', ',', 'Sulawesi']], 7, True)

['Cia-Cia', 'Writing', 'system', 'Hangul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Cia-Cia', 'Writing', 'system', 'Gundul', '(', 'former', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']], 7, True)

['Cia-Cia', 'Language', 'codes', 'ISO', '639-3', 'cia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lakon', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)



```

['Lakon', 'Region', 'Gaua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Lakon', 'Language', 'codes', 'ISO', '639-2', 'map']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Lakon', 'Language', 'codes', 'ISO', '639-3', 'lkn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pnar', 'Spoken', 'in', 'India', ',', 'Bangladesh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Pnar', 'Language', 'codes', 'ISO', '639-3', 'pbv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Highland', 'Spoken', 'in', 'Mexico', ':', 'Puebla']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', ':', 'Puebla']], 6, True)

['Highland', 'Spoken', 'in', 'Veracruz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Highland', 'Spoken', 'in', 'Hidalgo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Highland', 'Spoken', 'in']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']], 3, True)

['Highland', 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Highland', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']], 16, True)

```

```

['Highland', 'Language', 'codes', 'ISO', '639-2', 'oto']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Highland', 'Language', 'codes', 'ISO', '639-3', 'otm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Temoaya', 'Spoken', 'in', 'Mexico', ':', 'Mexico', 'state']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico',
':', 'Mexico', 'state']]], 7, True)

['Temoaya', 'Official', 'language', 'in', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Temoaya', 'Regulated', 'by', 'Secretar', '\xc3', '\xad', 'a', 'de',
'Educaci', '\xc3', '\xb3', 'n', 'P', '\xc3', '\xba', 'blica']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',
'Secretar', '\xc3', '\xad', 'a', 'de', 'Educaci', '\xc3', '\xb3', 'n', 'P',
'\xc3', '\xba', 'blica']]], 16, True)

['Temoaya', 'Language', 'codes', 'ISO', '639-2', 'oto']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Temoaya', 'Language', 'codes', 'ISO', '639-3', 'ott']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yodz', 'Spoken', 'in', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yodz', 'Region', 'Perm', 'Krai']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Yodz', 'Language', 'codes', 'ISO', '639-3', 'kpv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Raga', 'Total', 'speakers', '6500']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

['Raga', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Raga', 'Region', 'Pentecost', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Raga', 'Language', 'codes', 'ISO', '639-3', 'lml']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Anuki', 'Total', 'speakers', '891']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Anuki', 'Language', 'codes', 'ISO', '639-3', 'mwc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Doga', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Doga', 'Language', 'codes', 'ISO', '639-3', 'dgg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fuliiru', 'Spoken', 'in', 'Uganda', 'and', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Uganda', 'and', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 10, True)

['Fuliiru', 'Language', 'codes', 'ISO', '639-3', 'flr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sampang', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Sampang', 'Language', 'codes', 'ISO', '639-3', 'rav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ga'dang", 'Spoken', 'in', 'Philippines']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
["Ga'dang", 'Region', 'Luzon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
["Ga'dang", 'Language', 'codes', 'ISO', '639-3', 'gdg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Pyu', 'Region', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Pyu', 'Language', 'codes', 'ISO', '639-3', 'pby']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tinigua', 'Total', 'speakers', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tinigua', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tinigua', 'Region', 'Meta', 'Department']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Tinigua', 'Region', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Tinigua', 'Language', 'codes', 'ISO', '639-3', 'tit']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Beigo', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Beigo', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```

['Beigo', 'Region', 'Darfur']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Beigo', 'Language', 'codes', 'ISO', '639-3', 'byg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Narom', 'Spoken', 'in', 'Sarawak', ',', 'Malaysia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Narom', 'Language', 'codes', 'ISO', '639-3', 'nrm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kiai', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kiai', 'Total', 'speakers', '450']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kiai', 'Official', 'language', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Kiai', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Kiai', 'Language', 'codes', 'ISO', '639-3', 'frt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jinxiang', 'dialect', 'Total', 'speakers', '15', ',', '000', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total', 'speakers', 'NUM', ',', 'NUM', 'speakers']], ['NUM', '15'], ['NUM', '000']], 8, True)

['Jinxiang', 'Region', 'Jinxiang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

```

```
['Jinxiang', 'Region', 'Cangnan', 'county']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jinxiang', 'Region', 'Wenzhou', 'prefecture-level', 'city']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Jinxiang', 'Region', 'Zhejiang', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jinxiang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Kurt\xc3\xb6p', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kurt\xc3\xb6p', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Kurt\xc3\xb6p', 'Language', 'codes', 'ISO', '639-3', 'xkz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Betanure', 'Spoken', 'in', 'Israel', ',', 'previously', 'Betanure']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Israel',  
,',', 'previously', 'Type: ANY']]], 7, True)
```

```
['Betanure', 'Total', 'speakers', 'at', 'most', '3', 'dozen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'at', 'most', 'NUM', 'dozen']], ['NUM', '3']], 7, True)
```

```
['Betanure', 'Language', 'codes', 'ISO', '639-3', 'lsd-bet', '\xe2',  
'\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'lsd-bet', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,  
True)
```

```
['Lou', 'Region', 'Manus', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Lou', 'Language', 'codes', 'ISO', '639-3', 'loj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sinauna', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Sinauna', 'Region', 'Tanay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Sinauna', 'Region', 'Rizal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Sinauna', 'Region', 'General', 'Nakar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Sinauna', 'Region', 'Quezon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Sinauna', 'Region', 'Montalban']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Sinauna', 'Region', 'Rizal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Sinauna', 'Region', 'Antipolo', 'City']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Sinauna', 'Language', 'codes', 'ISO', '639-3', 'agv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jek', 'Spoken', 'in', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Jek', 'Region', 'Quba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Jek', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Colville-Okanagan', 'Total', 'speakers', '150']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Colville-Okanagan', 'Spoken', 'in', 'Canada', ',', 'United', 'States',  
'of', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Canada',  
,',', 'United', 'States', 'of', 'America']], 9, True)

['Colville-Okanagan', 'Language', 'codes', 'ISO', '639-3', 'oka']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alekano', 'Spoken', 'in', 'Goroka', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Alekano', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Alekano', 'Language', 'codes', 'ISO', '639-3', 'gah']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Shilluk', 'Spoken', 'in', 'Southern', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Shilluk', 'Total', 'speakers', '200', ',', '000', 'in', '1980s']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '200'], ['NUM', '000']], 8,  
True)

['Shilluk', 'Writing', 'system', 'Latin', 'alphabet']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Shilluk', 'Official', 'language', 'in', 'Shilluk', 'Kingdom']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',



'in', 'Type: ANY', 'Type: ANY']]], 6, True)

['Shilluk', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Shilluk', 'Language', 'codes', 'ISO', '639-3', 'shk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Fwe', 'Spoken', 'in', 'Okavango', 'River', ',', 'Namibia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'River', ',', 'Type: ANY']]], 7, True)

['Fwe', 'Language', 'codes', 'ISO', '639-3', 'fwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kuhane', 'Spoken', 'in', 'Okavango', 'River']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kuhane', 'Language', 'codes', 'ISO', '639-3', 'sbs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kalto', 'Spoken', 'in', 'Nandurbar', 'and', 'Jalgaon', 'Districts', ',', 'on', 'the', 'border', 'of', 'Maharashtra', 'and', 'Madhya', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nandurbar', 'and', 'Jalgaon', 'Districts', ',', 'on', 'the', 'border', 'of', 'Maharashtra', 'and', 'Madhya', 'Pradesh']]], 16, True)

['Kalto', 'Language', 'codes', 'ISO', '639-3', 'nlx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nihali', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nihali', 'Region', 'Jalgaon', 'Jamod', ',', 'on', 'the', 'border', 'of', 'Maharashtra', 'and', 'Madhya', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jalgaon', 'Jamod', ',', 'on', 'the', 'border', 'of', 'Maharashtra', 'and', 'Madhya', 'Pradesh']]], 13, True)

['Nihali', 'Language', 'codes', 'ISO', '639-3', 'nll']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Navalese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Navalese', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Navalese', 'Region', 'Somontano', 'de', 'Barbastro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Navalese', 'Region', 'El', 'Grado']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Navalese', 'Region', 'Naval']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Navalese', 'Region', 'Hoz', 'y', 'Costean']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Navalese', 'Region', 'and', 'Coscojuela', 'de', 'Fantova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Coscojuela', 'de', 'Fantova']]], 6, True)

['Navalese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Fonz', 'Aragonese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Fonz', 'Region', 'Ribagor', '\xc3', '\xa7', 'a', ',', 'Fonz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ribagor',  
'\xc3', '\xa7', 'a', ',', 'Type: ANY']]], 8, True)

['Fonz', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', ',', 'Type: ANY']]], 6, True)

['Fonz', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Panticutian', 'Aragonese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Panticutian', 'Dialects', 'Tensinian', 'Aragonese']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Panticutian', 'Region', 'Alto', 'G', '\xc3', '\xa1', 'llego', ',',  
'Panticosa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Alto', 'G',  
'\xc3', '\xa1', 'llego', ',', 'Panticosa']]], 9, True)

['Panticutian', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Panticutian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Damakawa', 'Total', 'speakers', '10', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '?']], ['NUM', '10']], 5, True)

['Damakawa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Damakawa', 'Region', 'Kebbi', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Damakawa', 'Language', 'codes', 'ISO', '639-3', 'dam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dolakhae', 'Spoken', 'in', 'Nepal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Dolakhae', 'Total', 'speakers', '6000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dolakhae', 'Language', 'codes', 'ISO', '639-3', 'new']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Folkspraak', 'Total', 'speakers', 'Probably', 'a', 'few', 'dozen', 'on', 'the', 'Internet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Probably', 'a', 'few', 'dozen', 'on', 'the', 'Internet']]], 10, True)

['Folkspraak', 'Date', 'founded', '1995']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Folkspraak', 'Category', '(', 'purpose', ')', 'constructed', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY', 'language']]], 7, True)

['Folkspraak', 'Category', '(', 'purpose', ')', 'Folkspraak']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)'], 'Type: ANY']]], 6, True)

['Folkspraak', 'Created', 'by', 'Collaborative', 'work']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Folkspraak', 'Category', '(', 'sources', ')', 'Vocabulary', ',', 'phonology', 'and', 'grammar', 'from', 'Germanic', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'Vocabulary', ',', 'phonology', 'and', 'grammar', 'from', 'Germanic', 'languages']]], 13, True)

['Folkspraak', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Tepecano', 'Spoken', 'in', 'Mexico', ':', 'Jalisco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Mexico', ':', 'Jalisco']]], 6, True)

```
['Tepecano', 'Total', 'speakers', 'extinct']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tepecano', 'Language', 'codes', 'ISO', '639-3', 'tep']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mongolian', 'Signed', 'in', 'Mongolia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Mongolian', 'Total', 'signers', '(', 'unknown', ')']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\
(', 'unknown', '\\)']]], 6, True)
```

```
['Cilentan', 'Spoken', 'in', 'Italy']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Cilentan', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Waxianghua', 'Spoken', 'in', 'western', 'Hunan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Waxianghua', 'Language', 'codes', 'ISO', '639-3', 'wxa']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngwe', 'Spoken', 'in', 'Cameroon']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Ngwe', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)
```

```
['Venezuelan', 'Signed', 'in', 'Venezuela']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Venezuelan', 'Total', 'signers', 'unknown']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
```

'Type: ANY']]], 4, True)

['Bakumpai', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bakumpai', 'Region', 'Kalimantan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bakumpai', 'Language', 'codes', 'ISO', '639-3', 'bkr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ma'anyan", 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

["Ma'anyan", 'Region', 'Kalimantan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

["Ma'anyan", 'Language', 'codes', 'ISO', '639-3', 'mhy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lawangan', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lawangan', 'Region', 'Kalimantan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lawangan', 'Language', 'codes', 'ISO', '639-3', 'lhx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Benga', 'Spoken', 'in', 'Equatorial', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Benga', 'Spoken', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Benga', 'Language', 'codes', 'ISO', '639-3', 'bng']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kombe', 'Spoken', 'in', 'Equatorial', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Kombe', 'Language', 'codes', 'ISO', '639-3', 'nui']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Seki', 'Spoken', 'in', 'Equatorial', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Seki', 'Spoken', 'in', 'Gabon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Seki', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Seki', 'Language', 'codes', 'ISO', '639-3', 'syi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gongduk', 'Spoken', 'in', 'Bhutan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Gongduk', 'Writing', 'system', 'Tibetan', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Gongduk', 'Language', 'codes', 'ISO', '639-3', 'goe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Central', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Central', 'Region', 'Northern', 'California']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Central', 'Language', 'codes', 'ISO', '639-3', 'poo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northeastern', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Northeastern', 'Region', 'Northern', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northeastern', 'Language', 'codes', 'ISO', '639-3', 'pef']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southeastern', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Southeastern', 'Region', 'Northern', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Southeastern', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Southeastern', 'Language', 'codes', 'ISO', '639-3', 'pom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Slovianski', 'Regulated', 'by', 'no', 'regulating', 'body']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'no',  
'regulating', 'body']]], 6, True)

['Slovianski', 'Total', 'speakers', '50-100']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Slovianski', 'Date', 'founded', '2006']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



['Slovianski', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Cyrillic', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Slovianski', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Slovianski', 'Category', '(', 'purpose', ')', 'international', 'auxiliary', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'international', 'auxiliary', 'language']], 8, True)

['Slovianski', 'Category', '(', 'purpose', ')', 'zonal', 'constructed', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'zonal', 'constructed', 'languages']], 8, True)

['Slovianski', 'Category', '(', 'purpose', ')', 'Slovianski']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)

['Slovianski', 'Created', 'by', 'Ondrej', 'Re', '\\xc4', '\\x8d', 'nik']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Ondrej', 'Re', '\\xc4', '\\x8d', 'nik']], 8, True)

['Slovianski', 'Created', 'by', 'Gabriel', 'Svoboda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Slovianski', 'Created', 'by', 'Jan', 'van', 'Steenbergen']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Jan', 'van', 'Steenbergen']], 6, True)

['Slovianski', 'Created', 'by', 'Igor', 'Polyakov']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Slovianski', 'Setting', 'and', 'usage', 'Auxiliary', 'language', ',', 'intended', 'for', 'communication', 'between', 'or', 'with', 'speakers', 'of', 'different', 'Slavic', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Auxiliary', 'language', ',', 'intended', 'for', 'communication',

'between', 'or', 'with', 'speakers', 'of', 'different', 'Slavic',  
'languages']]], 18, True)

['Slovianski', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Slovianski', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Eastern', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Eastern', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'hae']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kyakhta', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Musi', 'Total', 'speakers', '4', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '4']], 5, True)

['Musi', 'Spoken', 'in', 'Palembang', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Musi', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Musi', 'Language', 'codes', 'ISO', '639-2', 'msa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Musi', 'Language', 'codes', 'ISO', '639-3', 'mui']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tangwang', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tangwang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tangwang', 'Region', 'Gansu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tangwang', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Botlikh', 'Spoken', 'in', 'Southwestern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Botlikh', 'Language', 'codes', 'ISO', '639-3', 'bph']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chamalal', 'Spoken', 'in', 'Southwestern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Chamalal', 'Language', 'codes', 'ISO', '639-3', 'cji']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karata', 'Spoken', 'in', 'Southern', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Karata', 'Language', 'codes', 'ISO', '639-3', 'kpt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Varhadi', 'Spoken', 'in', 'Vidarbha', 'region', 'of', 'Maharashtra', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Vidarbha', 'region', 'of', 'Maharashtra', ',', 'India']], 9, True)

['Varhadi', 'Writing', 'system', 'Devanagari', '(', 'official', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Varhadi', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Varhadi', 'Language', 'codes', 'ISO', '639-3', 'vah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kryts', 'Spoken', 'in', 'Quba', 'Rayon', ',', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Quba',  
'Rayon', ',', 'Azerbaijan']]], 7, True)

['Kryts', 'Language', 'codes', 'ISO', '639-3', 'kry']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Budukh', 'Spoken', 'in', 'Quba', 'Rayon', ',', 'Azerbaijan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Quba',  
'Rayon', ',', 'Azerbaijan']]], 7, True)

['Budukh', 'Language', 'codes', 'ISO', '639-3', 'bdk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kajtak', 'Spoken', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kajtak', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kajtak', 'Language', 'codes', 'ISO', '639-3', 'dar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kubachi', 'Spoken', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kubachi', 'Total', 'speakers', '?']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kubachi', 'Language', 'codes', 'ISO', '639-3', 'dar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Itsari', 'Spoken', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Itsari', 'Language', 'codes', 'ISO', '639-3', 'dar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Chirag', 'Spoken', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chirag', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chirag', 'Language', 'codes', 'ISO', '639-3', 'dar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kanikkaran', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kanikkaran', 'Writing', 'system', 'Roman', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kanikkaran', 'Writing', 'system', 'Malayalam', 'Script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kanikkaran', 'Writing', 'system', 'Tamil', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kanikkaran', 'Language', 'codes', 'ISO', '639-3', 'kev']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Kgalagadi', 'Spoken', 'in', 'Botswana']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kgalagadi', 'Total', 'speakers', '4000']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kgalagadi', 'Language', 'codes', 'ISO', '639-3', 'xkv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tuyuca', 'Spoken', 'in', 'Colombia', ',', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Tuyuca', 'Total', 'speakers', '570', '(', 'Etnias', 'de', 'Colombia',
')', ';', '810', '(', 'SIL', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', '\\(', 'Etnias', 'de', 'Colombia', '\\)', ';', 'NUM', '\\(', 'SIL',
'\\)'']], ['NUM', '570'], ['NUM', '810']], 14, True)
```

```
['Tuyuca', 'Language', 'codes', 'ISO', '639-3', 'tue']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Musasa', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],
3, True)
```

```
['Musasa', 'Spoken', 'in']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3,
True)
```

```
['Musasa', 'Region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']]], 2, True)
```

```
['Musasa', 'Language', 'codes', 'ISO', '639-2', 'bih']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Musasa', 'Language', 'codes', 'ISO', '639-3', 'smm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

['Musasa', 'Language', 'codes', 'ISO', '639-1', 'bh', '(', 'Bihari', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'bh', '\\\\(', 'Bihari', '\\\\)']], ['NUM', '639-1']], 9, True)

['Himariote', 'Spoken', 'in', 'Himar', '\\xc3', '\\xab', ',', 'Albania', ';',  
'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Himar',  
'\\xc3', '\\xab', ',', 'Albania', ';', 'Greece']]], 10, True)

['Himariote', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Gwahatike', 'Total', 'speakers', '1570']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gwahatike', 'Language', 'codes', 'ISO', '639-3', 'dah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nkore-Kiga', 'Standard', 'forms', 'Kitara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nkore-Kiga', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nkore-Kiga', 'Dialects', 'Nkore']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Nkore-Kiga', 'Dialects', 'Kiga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Nkore-Kiga', 'Language', 'codes', 'ISO', '639-3', 'nyn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Plains', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Plains', 'Language', 'codes', 'ISO', '639-3', 'pmw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Na\xe2\x80\x99vi', 'Total', 'speakers', '-']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Na\xe2\x80\x99vi', 'Date', 'founded', '2005']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Na\xe2\x80\x99vi', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Na\xe2\x80\x99vi', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Na\xe2\x80\x99vi', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Na\xe2\x80\x99vi', 'Category', '(', 'purpose', ')', 'Na', '\xe2', '\x80', '\x99', 'vi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Na', '\xe2', '\x80', '\x99', 'vi']], 10, True)

['Na\xe2\x80\x99vi', 'Pronunciation', '\xcb', '\x88', 'na', '\xca', '\x94', 'vi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xcb', '\x88', 'na', '\xca', '\x94', 'vi']], 8, True)

['Na\xe2\x80\x99vi', 'Created', 'by', 'Paul', 'Frommer']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Na\xe2\x80\x99vi', 'Category', '(', 'sources', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'Type: ANY', 'Type: ANY']], 7, True)

['Na\xe2\x80\x99vi', 'Category', '(', 'sources', ')', 'a', 'priori',



```

'languages']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'sources', '\\)', 'a', 'priori', 'language']]], 8, True)

['Na\\xe2\\x80\\x99vi', 'Setting', 'and', 'usage', '2009', 'film', 'Avatar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',
'usage', 'NUM', 'film', 'Avatar']], ['NUM', '2009']], 7, True)

['Na\\xe2\\x80\\x99vi', 'Language', 'codes', 'ISO', '639-2', 'art']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Na\\xe2\\x80\\x99vi', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kasem', 'Spoken', 'in', 'Ghana', ',', 'Burkina', 'Faso']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ghana',
',', 'Burkina', 'Faso']]], 7, True)

['Kasem', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Kasem', 'Language', 'codes', 'ISO', '639-3', 'xsm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cahuilla', 'Spoken', 'in', 'USA']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cahuilla', 'Region', 'Southern', 'California']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Cahuilla', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Cahuilla', 'Language', 'codes', 'ISO', '639-3', 'chl']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Siyi', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],

```

3, True)

```
['Siyi', 'Region', 'Sze', 'Yup']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Siyi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Siyi', 'Language', 'codes', 'ISO', '639-3', 'yue-siy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Vao', 'Language', 'codes', 'ISO', '639-3', 'vao']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Siamou', 'Spoken', 'in', 'Burkina', 'Faso']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Siamou', 'Language', 'codes', 'ISO', '639-3', 'sif']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Epie', 'Spoken', 'in', 'Bayelsa', 'state', ',', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Bayelsa', 'state', ',', 'Nigeria']]], 7, True)
```

```
['Epie', 'Total', 'speakers', '12', ',', '000', 'in', '1973']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '12'], ['NUM', '000']], 8,  
True)
```

```
['Epie', 'Language', 'codes', 'ISO', '639-3', 'epi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mian', 'Region', 'Sandaun', 'province', ',', '']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mian', 'Region', 'Telefomin', 'district']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mian', 'Language', 'codes', 'ISO', '639-3', 'mpt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Intemelio', 'Spoken', 'in', 'Monaco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Intemelio', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Intemelio', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Intemelio', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Kuuk', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kuuk', 'Language', 'codes', 'ISO', '639-3', 'thd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yirrk-Thangalkl', 'Total', 'speakers', 'Unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yirrk-Thangalkl', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yirrk-Thangalkl', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Yirrk-Thangalkl', 'Language', 'codes', 'ISO', '639-3', 'yiy-jir']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dhivehi', 'Spoken', 'in', 'Maldives', 'and', 'Minicoy', '(', 'India', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Maldives', 'and', 'Minicoy', '\\(', 'India', '\\)']], 9, True)

['Dhivehi', 'Region', 'South', 'Asia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dhivehi', 'Writing', 'system', 'T', '\\xc4', '\\x81', 'na']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'T', '\\xc4', '\\x81', 'na']], 7, True)

['Dhivehi', 'Official', 'language', 'in', 'Maldives']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']], 5, True)

['Dhivehi', 'Language', 'codes', 'ISO', '639-2', 'div']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dhivehi', 'Language', 'codes', 'ISO', '639-3', 'div']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dhivehi', 'Language', 'codes', 'ISO', '639-1', 'dv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Benasquese', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Benasquese', 'Region', 'Ribagorza']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Benasquese', 'Region', 'Campo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Benasquese', 'Region', 'Estadilla']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

ANY']]], 3, True)

['Benasquese', 'Region', 'Graus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Benasquese', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Benasquese', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Kuuk', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kuuk', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bankon', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bankon', 'Region', 'Littoral', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bankon', 'Language', 'codes', 'ISO', '639-3', 'abb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ske', 'Total', 'speakers', '600']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ske', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ske', 'Region', 'Pentecost', 'island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Ske', 'Language', 'codes', 'ISO', '639-3', 'ske']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lombi', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Lombi', 'Region', 'Southwest', 'region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Lombi', 'Language', 'codes', 'ISO', '639-3', 'bbi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['U', 'Spoken', 'in', 'China']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['U', 'Language', 'codes', 'ISO', '639-3', 'uuu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Algerian', 'Signed', 'in', 'Algeria', ',', 'Morocco']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in',
'Algeria', ',', 'Morocco']]], 6, True)

['Algerian', 'Total', 'signers', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sursurunga', 'Spoken', 'in', 'New', 'Ireland']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Sursurunga', 'Language', 'codes', 'ISO', '639-3', 'sgz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['IsiNgqumo', 'Total', 'speakers', 'unknown']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['IsiNgqumo', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Koraga', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Koraga', 'Writing', 'system', 'Kannada', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Koraga', 'Language', 'codes', 'ISO', '639-3', 'kfd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Beni', 'Spoken', 'in', 'Algeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beni', 'Total', 'speakers', 'unknown']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Beni', 'Language', 'codes', 'ISO', '639-3', 'rif']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karamojong', 'Spoken', 'in', 'Uganda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Karamojong', 'Region', 'Karamoja', ',', 'Moroto', 'District']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Karamoja', ',', 'Moroto', 'District']], 6, True)

['Karamojong', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Karamojong', 'Language', 'codes', 'ISO', '639-3', 'kdj']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Duruwa', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Duruwa', 'Writing', 'system', 'Devanagari', 'script', ',', 'Oriya', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Duruwa', 'Language', 'codes', 'ISO', '639-3', 'pci']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Globish', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Globish', 'Date', 'founded', '1998']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Globish', 'Category', '(', 'purpose', ')', 'constructed', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Globish', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY', 'language']], 7, True)

['Globish', 'Category', '(', 'purpose', ')', 'Globish']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Type: ANY']], 6, True)

['Globish', 'Created', 'by', 'Madhukar', 'Gogate']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type: ANY', 'Type: ANY']], 5, True)

['Globish', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'international', 'auxiliary', 'language']], 7, True)

['Globish', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Forest', 'Spoken', 'in', 'Northern', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)



```

['Forest', 'Language', 'codes', 'ISO', '639-3', 'yrk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tundra', 'Spoken', 'in', 'Northern', 'Russia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Tundra', 'Language', 'codes', 'ISO', '639-3', 'yrk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Globish', 'Total', 'speakers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']]], 6, True)

['Globish', 'Date', 'founded', '2004']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Globish', 'Category', '(', 'purpose', ')', 'constructed', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Globish', 'Category', '(', 'purpose', ')', 'auxiliary', 'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Globish', 'Category', '(', 'purpose', ')', 'Globish']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Globish', 'Created', 'by', 'Jean-Paul', 'Nerriere']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Globish', 'Setting', 'and', 'usage', 'international', 'auxiliary',
'language']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',
'usage', 'international', 'auxiliary', 'language']]], 7, True)

['Globish', 'Language', 'codes', 'ISO', '639-3', 'eng']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Margi', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Margi', 'Region', 'Borno', 'State', ',', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Borno',  
'State', ',', 'Adamawa', 'State']]], 7, True)

['Margi', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Margi', 'Language', 'codes', 'ISO', '639-3', 'mrt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Margi', 'Language', 'codes', 'ISO', '639-3', 'Ethnic', 'territories',  
'of', 'the', 'Marghi-speaking', 'people', 'in', 'Nigeria', '(', 'organge',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Ethnic', 'territories', 'of', 'the', 'Marghi-speaking',  
'people', 'in', 'Nigeria', '\\(', 'organge', '\\)']], ['NUM', '639-3']],  
16, True)

['Konkani', 'Region', 'Konkan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Konkani', 'Region', 'includes', 'the', 'state', 'of', 'Karnataka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'includes',  
'the', 'state', 'of', 'Karnataka']]], 7, True)

['Konkani', 'Region', 'Maharashtra', 'and', 'some', 'parts', 'of',  
'Kerala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maharashtra',  
'and', 'some', 'parts', 'of', 'Kerala']]], 8, True)

['Konkani', 'Region', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Konkani', 'Region', 'Pakistan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Konkani', 'Region', 'Persian', 'Gulf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Region', 'Lisbon', 'in', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Konkani', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Writing', 'system', 'Devanagari', '(', 'official', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Konkani', 'Writing', 'system', 'Roman']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Writing', 'system', 'Kannada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Konkani', 'Writing', 'system', 'Malayalam', 'and', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Konkani', 'Pronunciation', 'k', '\\xc3', '\\xb5', 'k', '\\xc9', '\\xb5', '\\xc9', '\\xb3', 'i', '(', 'standard', ')', ',', 'k', '\\xc3', '\\xb5', 'k', '\\xc9', '\\xb3', 'i', '(', 'popular', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'k', '\\xc3', '\\xb5', 'k', '\\xc9', '\\xb5', '\\xc9', '\\xb3', 'i', '\\(', 'standard', '\\)', ',', 'k', '\\xc3', '\\xb5', 'k', '\\xc9', '\\xb3', 'i', '\\(', 'popular', '\\)']]], 25, True)

['Konkani', 'Official', 'language', 'in', 'Goa', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', ',', 'Type: ANY']]], 7, True)

['Konkani', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Goa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Various', 'academies', 'and', 'the', 'Government', 'of', 'Goa']]], 10, True)

['Konkani', 'Language', 'codes', 'ISO', '639-2', 'kok']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Konkani', 'Language', 'codes', 'ISO', '639-3', 'kex']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Isconahua', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba', ',', 'Brazil']]], 8, True)

['Isconahua', 'Language', 'codes', 'ISO', '639-3', 'isc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tanzanian', 'Signed', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tanzanian', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Makolkol', 'Spoken', 'in', 'New', 'Britain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Makolkol', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Makolkol', 'Language', 'codes', 'ISO', '639-3', 'zmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koyraboro', 'Spoken', 'in', 'Mali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koyraboro', 'Language', 'codes', 'ISO', '639-3', 'ses']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wawa', 'Total', 'speakers', '3000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wawa', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wawa', 'Region', 'Adamawa', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wawa', 'Region', 'Mayo-Banyo', 'Division']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wawa', 'Region', 'Bankim', 'Subdivision']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wawa', 'Region', 'west', 'of', 'Banyo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Wawa', 'Region', 'thirteen', 'villages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Wawa', 'Dialects', 'Gandua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Wawa', 'Language', 'codes', 'ISO', '639-3', 'www']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['(Batak)', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['(Batak)', 'Region', 'Sumatra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['(Batak)', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Batak', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

```

['(Batak)', 'Language', 'codes', 'ISO', '639-3', 'btm']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Itsekiri', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Itsekiri', 'Language', 'codes', 'ISO', '639-3', 'its']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Croatian', 'Dialects', 'Chakavian', '(', '\xc4', '\x8c', 'akavian', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
'\\(', '\xc4', '\x8c', 'akavian', '\\)']]], 8, True)

['Croatian', 'Dialects', 'Kajkavian']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Croatian', 'Dialects', 'Shtokavian', '(', '\xc5', '\xa0', 'tokavian',
')', '(', 'standard', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
'\\(', '\xc5', '\xa0', 'tokavian', '\\)', '\\(', 'standard', '\\)']]], 11,
True)

['Croatian', 'Dialects', 'Torlakian', '(', 'Krashovani', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
'\\(', 'Type: ANY', '\\)']]], 6, True)

['Croatian', 'Spoken', 'in', 'Croatia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Croatian', 'Spoken', 'in', 'Bosnia', 'and', 'Herzegovina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'and', 'Type: ANY']]], 6, True)

['Croatian', 'Spoken', 'in', 'Serbia', '(', 'Vojvodina', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', '\\(', 'Type: ANY', '\\)']]], 7, True)

['Croatian', 'Spoken', 'in', 'Montenegro']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

['Croatian', 'Spoken', 'in', 'Slovenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Croatian', 'Spoken', 'in', 'and', 'diaspora']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Croatian', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Croatian', 'Pronunciation', 'x', '\xc5', '\x99', '\xcc', '\xa9', '\xca',  
'\x8b', 'a', '\xcb', '\x90', 'tski', '\xcb', '\x90']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'x',  
'\xc5', '\x99', '\xcc', '\xa9', '\xca', '\x8b', 'a', '\xcb', '\x90',  
'tski', '\xcb', '\x90']]], 15, True)

['Croatian', 'Official', 'language', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Croatian', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Bosnia', 'and', 'Herzegovina']]], 7, True)

['Croatian', 'Recognised', 'minority', 'language', 'in', 'Austria', '(',  
'in', 'Burgenland', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Austria', '\\(', 'in', 'Burgenland', '\\)']]], 10, True)

['Croatian', 'Recognised', 'minority', 'language', 'in', 'Italy', '(',  
'in', 'Molise', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Italy', '\\(', 'in', 'Molise', '\\)']]], 10, True)

['Croatian', 'Recognised', 'minority', 'language', 'in', 'Romania', '(',  
'in', 'Cara', '\xc5', '\x9f', 'ova', ',', 'Lupac', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Romania', '\\(', 'in', 'Cara', '\xc5', '\x9f', 'ova',  
,', 'Lupac', '\\)']]], 15, True)

['Croatian', 'Recognised', 'minority', 'language', 'in', 'Serbia', '(',  
'Vojvodina', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',  
'language', 'in', 'Serbia', '\\(', 'Vojvodina', '\\)']]], 9, True)

['Croatian', 'Language', 'codes', 'ISO', '639-2', 'hrv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Croatian', 'Language', 'codes', 'ISO', '639-3', 'hrv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Croatian', 'Language', 'codes', 'Linguasphere', 'Dialectal', 'map', 'of',  
'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Dialectal', 'map', 'of', 'Croatia']]]], 8, True)

['Croatian', 'Language', 'codes', 'ISO', '639-1', 'hr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Mandalorian', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]]], 6, True)

['Mandalorian', 'Category', '(', 'purpose', ')', 'constructed',  
'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]]], 7, True)

['Mandalorian', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]]], 7, True)

['Mandalorian', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]]], 7, True)

['Mandalorian', 'Category', '(', 'purpose', ')', 'Mandalorian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]]], 6, True)

['Mandalorian', 'Created', 'by', 'Karen', 'Traviss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Type:  
ANY', 'Type: ANY']]]], 5, True)

['Mandalorian', 'Category', '(', 'sources', ')', 'constructed',  
'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',



'sources', '\\)', 'Type: ANY', 'Type: ANY']]], 7, True)

['Mandalorian', 'Category', '(', 'sources', ')', 'a', 'priori',  
'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'sources', '\\)', 'a', 'priori', 'language']]], 8, True)

['Mandalorian', 'Setting', 'and', 'usage', 'The', 'Mandalorians', 'of',  
'Star', 'Wars']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'The', 'Mandalorians', 'of', 'Star', 'Wars']]], 9, True)

['Mandalorian', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',  
'\\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Lombardic', 'Spoken', 'in', 'Pannonia', 'and', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['Lombardic', 'Writing', 'system', 'Runic', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Lombardic', 'Language', 'codes', 'ISO', '639-3', 'lng']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Orsam\\xc3\\xa5l', 'Spoken', 'in', 'Sweden']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Orsam\\xc3\\xa5l', 'Region', 'Orsa', 'municipality']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Orsam\\xc3\\xa5l', 'Official', 'language', 'in', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Orsam\\xc3\\xa5l', 'Regulated', 'by', 'Ossmolslag']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Orsam\\xc3\\xa5l', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80',

'\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Guragone', 'Spoken', 'in', 'Northern', 'Territory']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Guragone', 'Total', 'speakers', 'a', 'few']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'a', 'few']], 5, True)

['Guragone', 'Language', 'codes', 'ISO', '639-3', 'gge']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngwo', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ngwo', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Ngwo', 'Pronunciation', '\xc5', '\x8b', 'w', '\xc9', '\x94']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc5', '\x8b', 'w', '\xc9', '\x94']], 7, True)

['Ngwo', 'Language', 'codes', 'ISO', '639-3', 'ngn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['North', 'Spoken', 'in', 'Germany']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['North', 'Region', 'Saxony']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['North', 'Region', 'Saxony-Anhalt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['North', 'Region', 'Brandenburg']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

ANY']]], 3, True)

['North', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Pokomo', 'Spoken', 'in', 'Kenya']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Pokomo', 'Region', 'Tana', 'River', 'District']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Pokomo', 'Language', 'codes', 'ISO', '639-3', 'pkb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cal\xc3\xa3o', 'Spoken', 'in', 'Portugal']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Cal\xc3\xa3o', 'Language', 'codes', 'ISO', '639-3', 'rge']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sinti-Manouche', 'Spoken', 'in', 'Germany']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Sinti-Manouche', 'Spoken', 'in', 'France']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Sinti-Manouche', 'Spoken', 'in', 'Austria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Sinti-Manouche', 'Spoken', 'in', 'northern', 'Italy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Sinti-Manouche', 'Total', 'speakers', '318', ',', '920', '(', 'SIL',  
'estimate', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', 'SIL', 'estimate', '\\)']], ['NUM', '318'],

['NUM', '920']], 10, True)

['Sinti-Manouche', 'Language', 'codes', 'ISO', '639-3', 'rmo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['French', 'Spoken', 'in', 'Saint', 'Lucia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['French', 'Spoken', 'in', 'Dominica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['French', 'Spoken', 'in', 'Grenada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['French', 'Spoken', 'in', 'Trinidad', 'and', 'Tobago']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]], 6, True)

['French', 'Language', 'codes', 'ISO', '639-3', 'acf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['French', 'Language', 'codes', 'Linguasphere', '51-AAC-ccg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Kalderash', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kalderash', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)

['Kalderash', 'Language', 'codes', 'ISO', '639-2', 'rom']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Kalderash', 'Language', 'codes', 'ISO', '639-3', 'rmy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cayuvava', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Cayuvava', 'Language', 'codes', 'ISO', '639-3', 'cyb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Northern', 'Spoken', 'in', 'Baltic', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Northern', 'Spoken', 'in', 'Belgium']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Britain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Czech', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Northern', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Germany']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Netherlands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Poland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Romania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Scandinavia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Spoken', 'in', 'Slovakia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Northern', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Northern', 'Language', 'codes', 'ISO', '639-2', 'rom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Northern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Scottish', 'Cant', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Scottish', 'Spoken', 'in', 'United', 'Kingdom']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Scottish', 'Region', 'Scotland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Scottish', 'Language', 'codes', 'ISO', '639-3', 'trl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lucumi', 'Spoken', 'in', 'Cuba']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lucumi', 'Spoken', 'in', 'Puerto', 'Rico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

```
['Lucumi', 'Spoken', 'in', 'Dominican', 'Republic']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Lucumi', 'Total', 'speakers', '---']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Lucumi', 'Language', 'codes', 'ISO', '639-3', 'luq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['High', 'Franconian', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['High', 'Spoken', 'in', 'Germany']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['High', 'Region', 'Bavaria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['High', 'Region', 'Baden-W', '\xc3', '\xbc', 'rttemberg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Baden-W',
'\xc3', '\xbc', 'rttemberg']]], 6, True)
```

```
['High', 'Region', 'Thuringia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['High', 'Region', 'Saxony']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['High', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Wiru', 'Spoken', 'in', 'Ialibu-Pangia', 'District', '', '']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ialibu-
Pangia', 'District', '', '']]], 6, True)
```

```
['Wiru', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['Wiru', 'Language', 'codes', 'ISO', '639-3', 'wiu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Greek', 'Signed', 'in', 'Greece']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Galician', 'Spoken', 'in', 'Galicia', ';', 'also', 'in', 'other',  
'parts', 'of', 'Spain', 'and', 'Portugal']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Galicia', ';', 'also', 'in', 'other', 'parts', 'of', 'Spain', 'and',  
'Portugal']]], 13, True)

['Galician', 'Pronunciation', '\xc9', '\xa1', 'a', '\xcb', '\x88', 'le',  
'\xc9', '\xa1', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\xc9',  
'\xa1', 'a', '\xcb', '\x88', 'le', '\xc9', '\xa1', 'o']]], 11, True)

['Galician', 'Regulated', 'by', 'Real', 'Academia', 'Galega']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by',  
'Real', 'Academia', 'Galega']]], 6, True)

['Galician', 'Language', 'codes', 'ISO', '639-2', 'glg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Galician', 'Language', 'codes', 'ISO', '639-3', 'glg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Galician', 'Language', 'codes', 'Linguasphere', '51-AAA-ab']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]], 5, True)

['Galician', 'Language', 'codes', 'ISO', '639-1', 'gl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Zanaki', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Zanaki', 'Region', 'Mara']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Zanaki', 'Region', 'Musoma', 'Rural', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Zanaki', 'Region', 'Makongoro', 'Division']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zanaki', 'Language', 'codes', 'ISO', '639-3', 'zak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Abaknon', 'Region', 'Eastern', 'Visayas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abaknon', 'Region', 'Capul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Abaknon', 'Region', 'Northern', 'Samar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abaknon', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abaknon', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abaknon', 'Language', 'codes', 'ISO', '639-3', 'abx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lausitzisch-Neum\xc3\xa4rkisch', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Lausitzisch-Neum\xc3\xa4rkisch', 'Spoken', 'in', 'Germany', ',', 'possibly', 'Poland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Germany', ',', 'possibly', 'Poland']]], 7, True)

['Lausitzisch-Neum\xc3\xa4rkisch', 'Region', 'Lusatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Lausitzisch-Neum\xc3\xa4rkisch', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Zialo', 'Spoken', 'in', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Zialo', 'Language', 'codes', 'ISO', '639-3', 'zil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chokwe', 'Spoken', 'in', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chokwe', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Chokwe', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chokwe', 'Official', 'language', 'in', 'Angola', '(', 'national', 'language', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Angola', '\\(', 'national', 'language', '\\)']]], 9, True)

['Chokwe', 'Regulated', 'by', 'Instituto', 'de', 'L', '\xc3', '\xad', 'nguas', 'Nacionais']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Instituto', 'de', 'L', '\xc3', '\xad', 'nguas', 'Nacionais']]], 10, True)

['Chokwe', 'Language', 'codes', 'ISO', '639-3', 'cjk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Balboa', 'Spoken', 'in', 'California']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Balboa', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Balboa', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Pacaraos', 'Spoken', 'in', 'Peru']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Pacaraos', 'Total', 'speakers', 'nearly', 'extinct']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'nearly', 'extinct']]], 5, True)
```

```
['Pacaraos', 'Language', 'codes', 'ISO', '639-2', 'que']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Pacaraos', 'Language', 'codes', 'ISO', '639-3', 'qvp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Old', 'Spoken', 'in', 'northern', 'France', 'and', 'parts', 'of',
'Belgium', '(', 'Wallonia', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'northern', 'France', 'and', 'parts', 'of', 'Belgium', '\\(', 'Wallonia',
'\\)']]], 12, True)
```

```
['Old', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
["Kwara'ae", 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
["Kwara'ae", 'Region', 'Malaita', 'Island']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Kwara'ae', 'Language', 'codes', 'ISO', '639-3', 'kwf']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tutelo', 'Language', 'extinction', 'dormant']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tutelo', 'Spoken', 'in', 'United', 'States']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Tutelo', 'Region', 'Virginia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Tutelo', 'Region', 'West', 'Virginia']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Tutelo', 'Region', 'North', 'Carolina', 'Ethnicity']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Tutelo', 'Region', 'Tutelo']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Tutelo', 'Region', 'Saponi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Tutelo', 'Region', 'Occaneechi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Tutelo', 'Region', 'Manahoac']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Tutelo', 'Region', 'Monacan']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)
```

```
['Tutelo', 'Language', 'codes', 'ISO', '639-3', 'tta']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jewish', 'Spoken', 'in', 'Palestine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jewish', 'Writing', 'system', 'Hebrew', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Jewish', 'Language', 'codes', 'ISO', '639-3', 'jpa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koalib', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']], 7, True)

['Koalib', 'Total', 'speakers', '44', ',', '000', 'in', '1984']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '44'], ['NUM', '000']], 8, True)

['Koalib', 'Language', 'codes', 'ISO', '639-3', 'kib']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Katla', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']], 7, True)

['Katla', 'Total', 'speakers', '14', ',', '000', 'in', '1984']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '14'], ['NUM', '000']], 8, True)

['Katla', 'Language', 'codes', 'ISO', '639-3', 'kcr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Saint-Barth\xc3\xa9lemy', 'Spoken', 'in', 'Saint-Barth', '\xc3', '\xa9', 'lemy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Saint-Barth', '\xc3', '\xa9', 'lemy']], 7, True)

['Saint-Barth\xc3\xa9lemy', 'Language', 'codes', 'ISO', '639-3', '\xe2',

```

'\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Guro', 'Language', 'codes', 'ISO', '639-3', 'goa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bundjalung', 'Spoken', 'in', 'New', 'South', 'Wales', ',', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY', 'Wales', ',', 'Australia']]], 8, True)

['Bundjalung', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Bundjalung', 'Language', 'codes', 'ISO', '639-3', 'bdy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mamuju', 'Spoken', 'in', 'Sulawesi', ',', 'Indonesia', 'Coordinates']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Sulawesi', ',', 'Indonesia', 'Coordinates']]], 7, True)

['Mamuju', 'Dialects', 'Mamuju', '(', 'prestigious', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
'\\(', 'Type: ANY', '\\)']]], 6, True)

['Mamuju', 'Dialects', 'Sumare-Rangas']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mamuju', 'Dialects', 'Padang']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mamuju', 'Dialects', 'Sinyonyoi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mamuju', 'Language', 'codes', 'ISO', '639-3', 'mqx']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ha', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

```

'Type: ANY']]], 4, True)

['Ha', 'Language', 'codes', 'ISO', '639-3', 'haq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyakyusa', 'Spoken', 'in', 'Tanzania', ',', 'Malawi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Nyakyusa', 'Language', 'codes', 'ISO', '639-3', 'nyy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hehe', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Hehe', 'Language', 'codes', 'ISO', '639-3', 'heh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bena', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bena', 'Language', 'codes', 'ISO', '639-3', 'bez']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luguru', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Luguru', 'Total', 'speakers', '690', ',', '000', 'in', '2001']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '690'], ['NUM', '000']], 8,  
True)

['Luguru', 'Language', 'codes', 'ISO', '639-3', 'ruf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Turu', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Turu', 'Total', 'speakers', '600', ',', '000', 'in', '2006']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '600'], ['NUM', '000']], 8,
True)
```

```
['Turu', 'Language', 'codes', 'ISO', '639-3', 'rim']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kansa', 'Spoken', 'in', 'United', 'States']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Kansa', 'Region', 'Oklahoma']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)
```

```
['Kansa', 'Language', 'codes', 'ISO', '639-3', 'ksk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Rangi', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Rangi', 'Total', 'speakers', '350', ',', '000', 'in', '1999']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '350'], ['NUM', '000']], 8,
True)
```

```
['Rangi', 'Language', 'codes', 'ISO', '639-3', 'lag']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nilamba', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Nilamba', 'Total', 'speakers', '450', ',', '000', 'in', '2006']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '450'], ['NUM', '000']], 8,
True)
```

```
['Nilamba', 'Language', 'codes', 'ISO', '639-3', 'nim']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mbugwe', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mbugwe', 'Total', 'speakers', '24', ',', '000', 'in', '1999']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '24'], ['NUM', '000']], 8, True)

['Mbugwe', 'Language', 'codes', 'ISO', '639-3', 'mgz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Ma'ya", 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Ma'ya", 'Total', 'speakers', '4000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

["Ma'ya", 'Language', 'codes', 'ISO', '639-3', 'slz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ohlone', 'Language', 'extinction', '1950s']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ohlone', 'Pronunciation', '/', 'o', '\xca', '\x8a', '\xcb', '\x88', 'lo', '\xca', '\x8a', 'ni', '\xcb', '\x90', '/']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', 'o', '\xca', '\x8a', '\xcb', '\x88', 'lo', '\xca', '\x8a', 'ni', '\xcb', '\x90', '/']], 15, True)

['Ohlone', 'Spoken', 'in', 'California', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ohlone', 'Spoken', 'in', 'Ohlone', 'people']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ohlone', 'Language', 'codes', 'ISO', '639-3', 'variously']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ohlone', 'Language', 'codes', 'ISO', '639-3', 'krb', '\xe2', '\x80', '\x93', 'Karkin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'krb', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM', '639-3']], 10, True)

['Tapachultec', 'Language', 'extinction', '1930s']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tapachultec', 'Spoken', 'in', 'Mexico']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tapachultec', 'Region', 'Chiapas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Kaikadi', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kaikadi', 'Region', 'Maharashtra', '(', 'Jalgaon', 'District', ')', ',', 'Karnataka']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maharashtra', '\\(', 'Jalgaon', 'District', '\\)', ',', 'Karnataka']], 9, True)

['Kaikadi', 'Language', 'codes', 'ISO', '639-3', 'kep']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Larantuka', 'Spoken', 'in', 'Flores']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Larantuka', 'Language', 'codes', 'ISO', '639-2', 'may', '(', 'B', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'may', '\\(', 'B', '\\)']], ['NUM', '639-2']], 9, True)

['Larantuka', 'Language', 'codes', 'ISO', '639-2', 'msa', '(', 'T', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'msa', '\\(', 'T', '\\)']], ['NUM', '639-2']], 9, True)

['Larantuka', 'Language', 'codes', 'ISO', '639-3', 'lrt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Larantuka', 'Language', 'codes', 'ISO', '639-1', 'ms']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Obispe\xc3\xb1o', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Obispe\xc3\xb1o', 'Language', 'codes', 'ISO', '639-3', 'obi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Songe', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Songe', 'Region', 'Kasai-Oriental', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Songe', 'Language', 'codes', 'ISO', '639-3', 'sop']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nande', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Nande', 'Region', 'Nord-Kivu', 'province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nande', 'Language', 'codes', 'ISO', '639-3', 'nnb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chuwabu', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chuwabu', 'Language', 'codes', 'ISO', '639-3', 'chw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chopi', 'Spoken', 'in', 'Mozambique']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chopi', 'Language', 'codes', 'ISO', '639-3', 'cce']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yaka', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo',  
,',', 'Angola']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo', ',,', 'Angola']]], 10, True)

['Yaka', 'Language', 'codes', 'ISO', '639-3', 'yaf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aka', 'Language', 'codes', 'ISO', '639-3', 'axk']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mbati', 'Spoken', 'in', 'Central', 'African', 'Republic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY', 'Republic']]], 6, True)

['Mbati', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Mbati', 'Language', 'codes', 'ISO', '639-3', 'mdn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gyele', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gyele', 'Language', 'codes', 'ISO', '639-3', 'gyi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baka', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Baka', 'Spoken', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Baka', 'Language', 'codes', 'ISO', '639-3', 'bkc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Basa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Basa', 'Language', 'codes', 'ISO', '639-3', 'bzw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luo', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Luo', 'Region', 'Adamawa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Luo', 'Language', 'codes', 'ISO', '639-3', 'luw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Asoa', 'Spoken', 'in', 'Congo', '(', 'DRC', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Asoa', 'Region', 'Ituri', 'forest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Asoa', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Asoa', 'Language', 'codes', 'ISO', '639-3', 'asv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kresh', 'Spoken', 'in', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kresh', 'Language', 'codes', 'ISO', '639-3', 'krs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kango', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Kango', 'Region', 'Ituri', 'forest']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kango', 'Language', 'codes', 'ISO', '639-3', 'kzy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kara', 'Language', 'codes', 'ISO', '639-3', 'kcm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Doba', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Doba', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Doba', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Doba', 'Language', 'codes', 'ISO', '639-3', ':', 'bjv', '\xe2', '\x80', '\x93', 'Bedjond']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'bjv', '\xe2', '\x80', '\x93', 'Bedjond']], ['NUM', '639-3']], 11, True)

['Doba', 'Language', 'codes', 'ISO', '639-3', 'mge', '\xe2', '\x80', '\x93', 'Mango']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'mge', '\xe2', '\x80', '\x93', 'Mango']], ['NUM', '639-3']], 10, True)

['Doba', 'Language', 'codes', 'ISO', '639-3', 'gqr', '\xe2', '\x80',  
\x93', 'Gor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'gqr', '\xe2', '\x80', '\x93', 'Gor']], ['NUM', '639-3']],  
10, True)

['Moru', 'Spoken', 'in', 'Southern', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Moru', 'Language', 'codes', 'ISO', '639-3', 'mgd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Jur', 'Spoken', 'in', 'South', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Jur', 'Language', 'codes', 'ISO', '639-3', 'bex']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bongo', 'Spoken', 'in', 'South', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Bongo', 'Language', 'codes', 'ISO', '639-3', 'bot']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bongo', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Bongo', 'Language', 'codes', 'ISO', '639-3', 'sys']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gaam', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gaam', 'Language', 'codes', 'ISO', '639-3', 'tbi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nyala', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nyala', 'Region', 'Daju', 'Hills']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nyala', 'Language', 'codes', 'ISO', '639-3', 'daj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sila', 'Spoken', 'in', 'Chad', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Sila', 'Region', 'Daju', 'Hills']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sila', 'Language', 'codes', 'ISO', '639-3', 'daj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Daju', 'Spoken', 'in', 'Chad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Daju', 'Region', 'Daju', 'Hills']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Daju', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Daju', 'Language', 'codes', 'ISO', '639-3', 'djç']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mamaind\xc3\xaa', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mamaind\xc3\xaa', 'Region', 'Mato', 'Grosso']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mamaind\xc3\xaa', 'Language', 'codes', 'ISO', '639-3', 'wmd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ofay\xc3\xa9', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ofay\xc3\xa9', 'Region', 'Mato', 'Grosso', 'do', 'Sul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mato', 'Grosso', 'do', 'Sul']]], 6, True)

['Ofay\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'opy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nahuatl', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nahuatl', 'Spoken', 'in', 'and', 'immigrants', 'in', 'United', 'States', 'and', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'and', 'immigrants', 'in', 'United', 'States', 'and', 'Canada']]], 10, True)

['Nahuatl', 'Official', 'language', 'in', 'In', 'Mexico', 'through', 'the', 'General', 'Law', 'of', 'Linguistic', 'Rights', 'of', 'Indigenous', 'Peoples']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'In', 'Mexico', 'through', 'the', 'General', 'Law', 'of', 'Linguistic', 'Rights', 'of', 'Indigenous', 'Peoples']]], 16, True)

['Nahuatl', 'Regulated', 'by', 'Instituto', 'Nacional', 'de', 'Lenguas', 'Ind', '\xc3', '\xad', 'genas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Instituto', 'Nacional', 'de', 'Lenguas', 'Ind', '\xc3', '\xad', 'genas']]], 11, True)

['Nahuatl', 'Language', 'codes', 'ISO', '639-2', 'nah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'For', 'modern',

'varieties', ',', 'see', 'List', 'of', 'Nahuan', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'For', 'modern', 'varieties', ',', 'see', 'List', 'of',  
'Nahuan', 'languages']], ['NUM', '639-3']], 14, True)

['Ngbaka', 'Spoken', 'in', 'CAR', ',', 'Republic', 'of', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'CAR',  
,',', 'Republic', 'of', 'Congo']]], 8, True)

['Ngbaka', 'Language', 'codes', 'ISO', '639-3', 'nbm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Krongo', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Krongo', 'Region', 'Kordofan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Krongo', 'Language', 'codes', 'ISO', '639-3', 'kgo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kadugli', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kadugli', 'Region', 'Kordofan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kadugli', 'Dialects', 'Kadugli']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kadugli', 'Dialects', 'Katcha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kadugli', 'Dialects', 'Miri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Kadugli', 'Language', 'codes', 'ISO', '639-3', 'xtc']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Moro', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']]), 7, True)
```

```
['Moro', 'Total', 'speakers', '30', ',', '000', 'in', '1982']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '30'], ['NUM', '000']], 8, True)
```

```
['Moro', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Moro', 'Language', 'codes', 'ISO', '639-3', 'mor']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tiro', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']]), 7, True)
```

```
['Tiro', 'Total', 'speakers', '40', ',', '000', 'in', '1982']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '40'], ['NUM', '000']], 8, True)
```

```
['Tiro', 'Language', 'codes', 'ISO', '639-3', 'tic']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Laro', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']]), 7, True)
```

```
['Laro', 'Total', 'speakers', '40', ',', '000', 'in', '1998']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '40'], ['NUM', '000']], 8, True)
```

```
['Laro', 'Language', 'codes', 'ISO', '639-3', 'lro']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Dengebu', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba',  
'Hills', ',', 'Sudan']]], 7, True)
```

```
['Dengebu', 'Language', 'codes', 'ISO', '639-3', 'dec']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lumun', 'Language', 'codes', 'ISO', '639-3', 'lmd']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tegali', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba',  
'Hills', ',', 'Sudan']]], 7, True)
```

```
['Tegali', 'Language', 'codes', 'ISO', '639-3', 'ras']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tingal', 'Spoken', 'in', 'Tengali', 'Hills', ',', 'Sudan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Tengali', 'Hills', ',', 'Sudan']]], 7, True)
```

```
['Tingal', 'Language', 'codes', 'ISO', '639-3', 'ras']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ngbaka', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'Congo']]], 7, True)
```

```
['Ngbaka', 'Total', 'speakers', '1', 'million']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '1']], 5, True)
```

```
['Ngbaka', 'Language', 'codes', 'ISO', '639-3', 'nga']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mayogo', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'Congo']]], 7, True)
```

```
['Mayogo', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Mayogo', 'Language', 'codes', 'ISO', '639-3', 'mdm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mba', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'Congo']]], 7, True)

['Mba', 'Language', 'codes', 'ISO', '639-3', 'mfc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wahgi', 'Language', 'codes', 'ISO', '639-3', 'wgi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chrau', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chrau', 'Region', '\xc4', '\x90', '\xe1', '\xbb', '\x93', 'ng', 'Nai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', '\xc4', '\x90',  
'\xe1', '\xbb', '\x93', 'ng', 'Type: ANY']]], 9, True)

['Chrau', 'Region', 'T', '\xc3', '\xa2', 'y', 'Ninh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'T', '\xc3',  
'\xa2', 'y', 'Ninh']]], 7, True)

['Chrau', 'Region', 'and', 'B', '\xc3', '\xac', 'nh', 'D', '\xc6', '\xb0',  
'\xc6', '\xa1', 'ng', 'provinces']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and', 'B',  
'\xc3', '\xac', 'nh', 'D', '\xc6', '\xb0', '\xc6', '\xa1', 'ng',  
'provinces']]], 14, True)

['Chrau', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Chrau', 'Language', 'codes', 'ISO', '639-3', 'crw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amba', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Amba', 'Spoken', 'in', 'DR', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Amba', 'Language', 'codes', 'ISO', '639-3', 'rwm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaki', 'Spoken', 'in', 'New', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Kaki', 'Total', 'speakers', '510']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kaki', 'Language', 'codes', 'ISO', '639-3', 'tbd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Deori', 'Spoken', 'in', 'India', ':', 'Assam', ':', 'Lakhimpur']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'India',

['Deori', 'Spoken', 'in', 'Assam', ';', 'Dhemaji', ';', 'Tinsukia', ';',  
'Jorhat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Assam',  
';', 'Dhemaji', ';', 'Tinsukia', ';', 'Jorhat']]], 10, True)

['Deori', 'Spoken', 'in', 'Arunachal', 'Pradesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Deori', 'Language', 'codes', 'ISO', '639-3', 'der']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyala', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nyala', 'Region', 'Busia', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```

['Nyala', 'Language', 'codes', 'ISO', '639-3', 'nle']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyole', 'Spoken', 'in', 'Kenya']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nyole', 'Region', 'Kakamega', 'District']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nyole', 'Language', 'codes', 'ISO', '639-3', 'nyd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bishuo', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bishuo', 'Total', 'speakers', '1']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Bishuo', 'Language', 'codes', 'ISO', '639-3', 'bwh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Noy', 'Spoken', 'in', 'Chad']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Noy', 'Language', 'codes', 'ISO', '639-3', 'noy']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Noone', 'Spoken', 'in', 'Cameroon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Noone', 'Language', 'codes', 'ISO', '639-3', 'nhu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bidyara', 'Spoken', 'in', 'Australia']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bidyara', 'Language', 'codes', 'ISO', '639-3', 'bym']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kurmalì', 'Spoken', 'in', 'India', ',', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Kurmalì', 'Region', 'Jharkhand', 'and', 'surrounding', 'states']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jharkhand', 'and', 'surrounding', 'states']]], 6, True)

['Kurmalì', 'Language', 'codes', 'ISO', '639-2', 'bih']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kurmalì', 'Language', 'codes', 'ISO', '639-3', 'kyw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Frainc-Comtou', 'Spoken', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Frainc-Comtou', 'Spoken', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Frainc-Comtou', 'Region', 'Franche-Comt', '\xc3', '\xa9']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Frainc-Comtou', 'Region', 'Canton', 'of', 'Jura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Frainc-Comtou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Sichuanese', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',



'Type: ANY']]], 4, True)

['Sichuanese', 'Region', 'Sichuan', 'and', 'Chongqing']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Sichuanese', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Sichuanese', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Sichuanese', 'Language', 'codes', 'ISO', '639-3', 'cmn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Sichuanese', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)

['Kongu', 'Tamil', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',  
'speakers', '\\xe2', '\\x80', '\\x93']]], 7, True)

['Kongu', 'Region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY']]], 2, True)

['Kongu', 'Writing', 'system', 'Tamil', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Kongu', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Kongu', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Yangbye', 'Region', 'Ramree', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yangbye', 'Region', 'South', 'Arakan', 'Coast']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Yangbye', 'Region', 'Irrawaddy', 'Division']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yangbye', 'Region', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Yangbye', 'Spoken', 'in', 'Arakan', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Yangbye', 'Pronunciation', 'IPA', ':', 'r', '\xc9', '\x99', 'mbr', '\xc3', '\xa8', '\xcb', '\x90', 'b', '\xc3', '\xa0', '\xc3', '\xb0', '\xc3', '\xa0', 'z', '\xc9', '\x99', '\xc9', '\xa1', '\xc3', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'IPA', ':', 'r', '\xc9', '\x99', 'mbr', '\xc3', '\xa8', '\xcb', '\x90', 'b', '\xc3', '\xa0', '\xc3', '\xb0', '\xc3', '\xa0', 'z', '\xc9', '\x99', '\xc9', '\xa1', '\xc3', '\xa1']], 26, True)

['Yangbye', 'Language', 'codes', 'ISO', '639-3', 'ybd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dothraki', 'Total', 'speakers', 'Fictional', 'language', ',', 'zero', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Fictional', 'language', ',', 'zero', 'speakers']], 8, True)

['Dothraki', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\(', 'purpose', '\)'], 'Type: ANY', 'language']], 7, True)

['Dothraki', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\(', 'purpose', '\)'], 'Type: ANY', 'language']], 7, True)

['Dothraki', 'Category', '(', 'purpose', ')', 'fictional', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\(', 'purpose', '\)'], 'Type: ANY', 'language']], 7, True)

['Dothraki', 'Category', '(', 'purpose', ')', 'Dothraki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\(',

'purpose', '\\)', 'Type: ANY']]], 6, True)

['Dothraki', 'Pronunciation', '\\xcb', '\\x88', 'd', '\\xcc', '\\xaa', 'o',  
'.', '\\xce', '\\xb8', '\\xc9', '\\xbe', 'a', '.', 'ki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb',  
'\\x88', 'd', '\\xcc', '\\xaa', 'o', '.', '\\xce', '\\xb8', '\\xc9', '\\xbe', 'a',  
'.', 'ki']]]], 16, True)

['Dothraki', 'Created', 'by', 'George', 'R', '.', 'R', '.', 'Martin', ',',  
'David', 'J', '.', 'Peterson']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by',  
'George', 'R', '.', 'R', '.', 'Martin', ',', 'David', 'J', '.',  
'Peterson']]]], 14, True)

['Dothraki', 'Category', '(', 'sources', ')', 'constructed', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'Type: ANY', 'Type: ANY']]]], 7, True)

['Dothraki', 'Category', '(', 'sources', ')', 'a', 'priori', 'languages']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\\\(',  
'sources', '\\\\)', 'a', 'priori', 'language']]]], 8, True)

['Dothraki', 'Setting', 'and', 'usage', '2011', 'Forthcoming', 'Series',  
'Game', 'of', 'Thrones', ',', 'A', 'Song', 'of', 'Ice', 'and', 'Fire']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'NUM', 'Forthcoming', 'Series', 'Game', 'of', 'Thrones', ',', 'A',  
'Song', 'of', 'Ice', 'and', 'Fire']], ['NUM', '2011']], 17, True)

['Dothraki', 'Language', 'codes', 'ISO', '639-2', 'art']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dothraki', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Kwambi', 'Spoken', 'in', 'Namibia', 'and', 'Angola']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']]]], 6, True)

['Kwambi', 'Region', 'Ovamboland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['Kwambi', 'Language', 'codes', 'ISO', '639-3', 'kwm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Embu', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Embu', 'Language', 'codes', 'ISO', '639-3', 'ebu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['D\\xc3\\xa2w', 'Total', 'speakers', '94']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['D\\xc3\\xa2w', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['D\\xc3\\xa2w', 'Region', 'Amazonas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['D\\xc3\\xa2w', 'Language', 'codes', 'ISO', '639-3', 'kwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Gurindji', 'Spoken', 'in', 'Victoria', 'River', 'and', 'Wave', 'Hill']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Victoria', 'River', 'and', 'Wave', 'Type: ANY']]], 8, True)

['Gurindji', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Gurindji', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Gurindji', 'Language', 'codes', 'ISO', '639-3', 'gue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Seimat', 'Region', 'Ninigo', 'and', 'Anchorite', 'island', 'groups', ',',  
'Manus', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ninigo',  
'and', 'Anchorite', 'island', 'groups', ',', 'Manus', 'Province']]], 10,

True)

```
['Seimat', 'Language', 'codes', 'ISO', '639-3', 'ssg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kaniat', 'Language', 'extinction', '1950']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kaniat', 'Language', 'codes', 'ISO', '639-3', 'ktk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bipi', 'Language', 'codes', 'ISO', '639-3', 'biq']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Nyindrou', 'Region', 'westernmost', 'Manus', 'Island', ',', 'Manus',  
'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'westernmost',  
'Manus', 'Island', ',', 'Manus', 'Province']], 8, True)
```

```
['Nyindrou', 'Language', 'codes', 'ISO', '639-3', 'lid']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Likum', 'Total', 'speakers', '80']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Likum', 'Region', 'westernmost', 'Manus', 'Island', ',', 'Manus',  
'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'westernmost',  
'Manus', 'Island', ',', 'Manus', 'Province']], 8, True)
```

```
['Likum', 'Language', 'codes', 'ISO', '639-3', 'lib']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Hermit', 'Language', 'extinction', 'extinct']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Hermit', 'Region', 'Hermit']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```

ANY']]], 3, True)

['Hermit', 'Region', 'Luf', 'and', 'Maron', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Luf', 'and',  
'Maron', 'Islands']]], 6, True)

['Hermit', 'Region', 'western', 'Manus', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Hermit', 'Language', 'codes', 'ISO', '639-3', 'llf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sori-Harengan', 'Total', 'speakers', '570']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sori-Harengan', 'Region', 'Sori', 'and', 'Harengan', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sori', 'and',  
'Harengan', 'Islands']]], 6, True)

['Sori-Harengan', 'Region', 'and', 'northwestern', 'coast', 'of', 'Manus',  
'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'northwestern', 'coast', 'of', 'Manus', 'Island']]], 8, True)

['Sori-Harengan', 'Region', 'Manus', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Sori-Harengan', 'Language', 'codes', 'ISO', '639-3', 'sbh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khehek', 'Region', 'West-central', 'Manus', 'Island', ',', 'Manus',  
'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'West-central',  
'Manus', 'Island', ',', 'Manus', 'Province']]], 8, True)

['Khehek', 'Dialects', 'Drehet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Khehek', 'Dialects', 'Levei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY']]], 3, True)

['Khehek', 'Language', 'codes', 'ISO', '639-3', 'tlx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tulu-Bohuai', 'Total', 'speakers', '1400']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tulu-Bohuai', 'Region', 'central', 'Manus', 'Island', 'and', 'Peli',  
'Island', ',', 'Manus', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'central',  
'Manus', 'Island', 'and', 'Peli', 'Island', ',', 'Manus', 'Province']]],  
11, True)

['Tulu-Bohuai', 'Language', 'codes', 'ISO', '639-3', 'rak']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tangale', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tangale', 'Language', 'codes', 'ISO', '639-3', 'ksq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Minjiang', 'dialect', 'Total', 'speakers', 'About', '30', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'About', 'NUM', 'million']], ['NUM', '30']], 7, True)

['Minjiang', 'Region', 'Sichuan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Minjiang', 'Region', 'Chongqing']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Minjiang', 'Region', 'Guizhou', 'and', 'Yunnan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Minjiang', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Minjiang', 'Pronunciation', 'min', '\xcb', '\xa8', '\xcb', '\xa9', 't', '\xc9', '\x95', 'ia', '\xc5', '\x8b', '\xcb', '\xa5', 'xa', '\xcb', '\xa8', '\xcb', '\xa8', '\xcb', '\xa6']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'min', '\xcb', '\xa8', '\xcb', '\xa9', 't', '\xc9', '\x95', 'ia', '\xc5', '\x8b', '\xcb', '\xa5', 'xa', '\xcb', '\xa8', '\xcb', '\xa8', '\xcb', '\xa6']]], 22, True)

['Minjiang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Yat\xc3\xaa-Fulni\xc3\xb4', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yat\xc3\xaa-Fulni\xc3\xb4', 'Region', 'Pernambuco']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yat\xc3\xaa-Fulni\xc3\xb4', 'Dialects', 'Fulni', '\xc3', '\xb4']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Fulni', '\xc3', '\xb4']]], 5, True)

['Yat\xc3\xaa-Fulni\xc3\xb4', 'Dialects', 'Yat', '\xc3', '\xaa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Yat', '\xc3', '\xaa']]], 5, True)

['Yat\xc3\xaa-Fulni\xc3\xb4', 'Language', 'codes', 'ISO', '639-3', 'fun']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Western', 'Aragonese', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Western', 'Dialects', 'Ans', '\xc3', '\xb3', 'Aragonese']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Ans', '\xc3', '\xb3', 'Type: ANY']]], 6, True)

['Western', 'Dialects', 'Cheso']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)



```
['Western', 'Dialects', 'Arag', '\xc3', '\xbc', '\xc3', '\xa9', 's',  
'Aragonese']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Arag',  
'\xc3', '\xbc', '\xc3', '\xa9', 's', 'Type: ANY']]], 9, True)
```

```
['Western', 'Region', 'Hecho']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Western', 'Region', 'Ans', '\xc3', '\xb3']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Western', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Western', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Ans\xc3\xb3', 'Aragonese', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Ans\xc3\xb3', 'Region', 'Ans', '\xc3', '\xb3', 'Valley']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Ans', '\xc3',  
'\xb3', 'Valley']]], 6, True)
```

```
['Ans\xc3\xb3', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Ans\xc3\xb3', 'Writing', 'system', 'Latin', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Ans\xc3\xb3', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['A\xc3\xadsa', 'Aragonese', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['A\\xc3\\xada', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['A\\xc3\\xada', 'Region', 'A', '\\xc3', '\\xad', 'sa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'A', '\\xc3', '\\xad', 'sa']]], 6, True)

['A\\xc3\\xada', 'Region', 'Esposa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['A\\xc3\\xada', 'Region', 'Sinu', '\\xc3', '\\xa9', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sinu', '\\xc3', '\\xa9', 's']]], 6, True)

['A\\xc3\\xada', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

['Arag\\xc3\\xbc\\xc3\\xa9s', 'Aragonese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Arag\\xc3\\xbc\\xc3\\xa9s', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Arag\\xc3\\xbc\\xc3\\xa9s', 'Region', 'A', '\\xc3', '\\xad', 'sa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'A', '\\xc3', '\\xad', 'sa']]], 6, True)

['Arag\\xc3\\xbc\\xc3\\xa9s', 'Region', 'Esposa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Arag\\xc3\\xbc\\xc3\\xa9s', 'Region', 'Sinu', '\\xc3', '\\xa9', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Sinu', '\\xc3', '\\xa9', 's']]], 6, True)

['Arag\\xc3\\xbc\\xc3\\xa9s', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-3']], 8, True)

```
['Jaca', 'Aragonese', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Jaca', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Jaca', 'Region', 'Jacetania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Jaca', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Cheso', 'Spoken', 'in', 'Aragon', ',', 'Spain']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)
```

```
['Cheso', 'Region', 'Hecho', 'Valley']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Cheso', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Rakhine', 'Region', 'Bandarban']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Rakhine', 'Region', 'Khagrachari']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Rakhine', 'Region', 'Patuakhali']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Rakhine', 'Region', 'and', 'Barguna', 'districts', 'of', 'Bangladesh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'and',  
'Barguna', 'districts', 'of', 'Bangladesh']]], 7, True)
```

```
['Rakhine', 'Region', 'Tripura', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',
```

'Type: ANY', 'Type: ANY']]], 5, True)

['Rakhine', 'Spoken', 'in', 'Burma']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rakhine', 'Spoken', 'in', 'Bangladesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rakhine', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rakhine', 'Pronunciation', 'IPA', ':', '\xc9', '\xb9', '\xc9', '\x99', 'k', '\xca', '\xb0', '\xc3', '\xa0', 'i', '\xc9', '\xb4', 'z', '\xc3', '\xa0', 'or', 'IPA', ':', '\xc9', '\xb9', '\xc9', '\x99', 'k', '\xca', '\xb0', '\xc3', '\xa0', 'i', '\xc9', '\xb4', 'z', '\xc9', '\x99', '\xc9', '\xa1', '\xc3', '\xa1']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'IPA', ':', '\xc9', '\xb9', '\xc9', '\x99', 'k', '\xca', '\xb0', '\xc3', '\xa0', 'i', '\xc9', '\xb4', 'z', '\xc3', '\xa0', 'or', 'IPA', ':', '\xc9', '\xb9', '\xc9', '\x99', 'k', '\xca', '\xb0', '\xc3', '\xa0', 'i', '\xc9', '\xb4', 'z', '\xc9', '\x99', '\xc9', '\xa1', '\xc3', '\xa1']]], 41, True)

['Rakhine', 'Language', 'codes', 'ISO', '639-3', 'rki']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Purubor\xc3\xa1', 'Total', 'speakers', '2']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Purubor\xc3\xa1', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Purubor\xc3\xa1', 'Region', 'Rond', '\xc3', '\xb4', 'nia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rond', '\xc3', '\xb4', 'nia']]], 6, True)

['Purubor\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'pur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Southern', 'Spoken', 'in', 'Malaysia']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Southern', 'Region', 'Melaka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Southern', 'Region', 'Muar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Southern', 'Region', 'Batu', 'Pahat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Southern', 'Region', 'Pontian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Southern', 'Region', 'Segamat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Southern', 'Region', 'Johor', 'Bahru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']]], ['NUM', '639-2']], 9, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']]], ['NUM', '639-2']], 9, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', 'nan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-1']], 6, True)
```

```
['Kei', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

['Kei', 'Region', 'Maluku', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kei', 'Language', 'codes', 'ISO', '639-3', 'kei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Damana', 'Total', 'speakers', '2000']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Damana', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Damana', 'Language', 'codes', 'ISO', '639-3', 'mbp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yugoslav', 'Signed', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yugoslav', 'Signed', 'in', 'Serbia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yugoslav', 'Signed', 'in', 'Bosnia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yugoslav', 'Signed', 'in', 'Slovenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yugoslav', 'Dialects', 'Slovenian', 'Sign']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yugoslav', 'Dialects', 'Croatian', 'Sign']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yugoslav', 'Language', 'codes', 'ISO', '639-3', 'either']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kenati', 'Language', 'codes', 'ISO', '639-3', 'gat']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngombe', 'Spoken', 'in', 'Central', 'African', 'Republic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', 'Republic']], 6, True)

['Ngombe', 'Language', 'codes', 'ISO', '639-3', 'nmj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngombe', 'Spoken', 'in', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ngombe', 'Language', 'codes', 'ISO', '639-3', 'ngc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Missouri', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Missouri', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Marra', 'Spoken', 'in', 'Arnhem', 'Land', ',', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY', ',', 'Australia']], 7, True)

['Marra', 'Language', 'codes', 'ISO', '639-3', 'mec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sangu', 'Spoken', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Sangu', 'Region', 'Central', 'Africa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```

['Sangu', 'Language', 'codes', 'ISO', '639-3', 'snq']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sangu', 'Spoken', 'in', 'Tanzania']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Sangu', 'Language', 'codes', 'ISO', '639-3', 'sbp']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bwisi', 'Spoken', 'in', 'Congo', ',', 'Gabon']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Bwisi', 'Language', 'codes', 'ISO', '639-3', 'bwz']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Luvale', 'Spoken', 'in', 'Angola']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Luvale', 'Spoken', 'in', 'Zambia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Luvale', 'Official', 'language', 'in', 'Zambia', '(', 'regional',
'language', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Zambia', '\\(', 'regional', 'language', '\\)']]], 9, True)

['Luvale', 'Regulated', 'by', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Luvale', 'Language', 'codes', 'ISO', '639-3', 'lue']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Midob', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\\xe2', '\\x80', '\\x93']]], 6, True)

```



['Midob', 'Spoken', 'in', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Midob', 'Region', 'West']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Midob', 'Language', 'codes', 'ISO', '639-3', 'mei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yavitero', 'Language', 'extinction', '1984']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yavitero', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yavitero', 'Region', 'Orinoco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Yavitero', 'Language', 'codes', 'ISO', '639-3', 'yvt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yaminahua', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yaminahua', 'Spoken', 'in', 'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yaminahua', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yaminahua', 'Total', 'speakers', '1400']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Yaminahua', 'Language', 'codes', 'ISO', '639-3', 'yaa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yaruro', 'Spoken', 'in', 'Venezuela']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yaruro', 'Language', 'codes', 'ISO', '639-3', 'yae']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aymaru', 'Spoken', 'in', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aymaru', 'Language', 'codes', 'ISO', '639-3', 'ayz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bangka', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bangka', 'Region', 'Sumatra']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Bangka', 'Language', 'codes', 'ISO', '639-2', 'msa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Bangka', 'Language', 'codes', 'ISO', '639-3', 'mfb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jakun', 'Spoken', 'in', 'Malaysia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Jakun', 'Region', 'Johor']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Jakun', 'Region', 'Pahang']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

```

['Jakun', 'Language', 'codes', 'ISO', '639-2', 'msa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Jakun', 'Language', 'codes', 'ISO', '639-3', 'jak']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eastern', 'Spoken', 'in', 'Guinea']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Eastern', 'Spoken', 'in', 'Liberia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Eastern', 'Spoken', 'in', 'Sierra', 'Leone']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)

['Eastern', 'Writing', 'system', 'Latin', 'script', ',', 'N\'Ko', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Eastern', 'Language', 'codes', 'ISO', '639-3', 'emk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mandara', 'Dialects', 'Simberi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Mandara', 'Dialects', 'Tatau']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Mandara', 'Dialects', 'Tabar']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Mandara', 'Writing', 'system', 'Latin', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']], 5, True)

['Mandara', 'Language', 'codes', 'ISO', '639-3', 'tbf']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bariba', 'Spoken', 'in', 'Benin', ',', 'Niger']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Bariba', 'Region', 'Borgou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bariba', 'Language', 'codes', 'ISO', '639-3', 'bba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Talinga', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Talinga', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the',  
'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Democratic', 'Republic', 'of', 'the', 'Congo']]], 8, True)

['Talinga', 'Language', 'codes', 'ISO', '639-3', 'tlj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['South', 'Region', 'Ksour', 'Mountains']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Region', 'Saoura', 'basin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Region', 'Figuig', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Spoken', 'in', 'Algeria', ',', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['South', 'Writing', 'system', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['South', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['South', 'Writing', 'system', 'Tifinagh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['South', 'Language', 'codes', 'ISO', '639-3', 'tzm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taushiro', 'Spoken', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Taushiro', 'Language', 'codes', 'ISO', '639-3', 'trr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jarawa', 'Spoken', 'in', 'eastern', 'Nigeria', ',', 'near', 'Bauchi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'eastern', 'Nigeria', ',', 'near', 'Bauchi']], 8, True)

['Jarawa', 'Language', 'codes', 'ISO', '639-3', 'jar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nsenga', 'Spoken', 'in', 'Zambia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nsenga', 'Spoken', 'in', 'Mozambique']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nsenga', 'Spoken', 'in', 'Zimbabwe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nsenga', 'Language', 'codes', 'ISO', '639-3', 'nse']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ternate', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ternate', 'Language', 'codes', 'ISO', '639-3', 'tft']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paya', 'Spoken', 'in', 'Honduras']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Paya', 'Language', 'codes', 'ISO', '639-3', 'pay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ombamba', 'Spoken', 'in', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ombamba', 'Spoken', 'in', 'Gabon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ombamba', 'Language', 'codes', 'ISO', '639-3', 'mbm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mondropolon', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mondropolon', 'Region', 'north-central', 'Manus', 'Island', ',', 'Manus', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'north-central', 'Manus', 'Island', ',', 'Manus', 'Province']], 8, True)

['Mondropolon', 'Language', 'codes', 'ISO', '639-3', 'nnp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['British', 'Language', 'extinction', 'Cumbric']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['British', 'Spoken', 'in', 'Iron', 'Age', 'Britain', ',', 'south', 'of', 'the', 'Firth', 'of', 'Forth']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Iron', 'Age', 'Britain', ',', 'south', 'of', 'the', 'Firth', 'of', 'Forth']]], 13, True)

['British', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bandi', 'Spoken', 'in', 'Guinea', ',', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Bandi', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bandi', 'Dialects', 'Tahamba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bandi', 'Dialects', 'Wawana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bandi', 'Dialects', 'Wulukoha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bandi', 'Dialects', 'Hasala']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bandi', 'Dialects', 'Lukasa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bandi', 'Dialects', 'Hembeh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bandi', 'Language', 'codes', 'ISO', '639-3', 'bza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mann', 'Spoken', 'in', 'Guinea', 'and', 'Liberia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'and', 'Type: ANY']]], 6, True)

['Mann', 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Mann', 'Language', 'codes', 'ISO', '639-3', 'mev']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oued', 'Spoken', 'in', 'Algeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oued', 'Language', 'codes', 'ISO', '639-3', 'tjo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oued', 'Language', 'codes', 'ISO', '639-3', 'Berber-speaking', 'areas', 'of', 'the', 'Mzab']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Berber-speaking', 'areas', 'of', 'the', 'Mzab']], ['NUM', '639-3']], 10, True)

['Oued', 'Language', 'codes', 'ISO', '639-3', 'Ouargla']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bayot', 'Spoken', 'in', 'Senegal', ',', 'Guinea-Bissau']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Bayot', 'Dialects', 'Essin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bayot', 'Language', 'codes', 'ISO', '639-3', 'bda']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kukuya', 'Language', 'codes', 'ISO', '639-3', 'kkw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koro', 'Spoken', 'in', 'Arunachal', 'Pradesh', ',', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',



'Arunachal', 'Pradesh', ',', 'India']]], 7, True)

['Awtuw', 'Language', 'codes', 'ISO', '639-3', 'kmn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lhokpu', 'Spoken', 'in', 'SW', 'Bhutan', '(', 'Samtse', ',', 'Chukha',  
)']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'SW',  
'Bhutan', '\\\\(', 'Samtse', ',', 'Chukha', '\\\\)']], 10, True)

['Lhokpu', 'Language', 'codes', 'ISO', '639-3', 'lhp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koro', 'Total', 'speakers', '280']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Koro', 'Spoken', 'in', 'Vanuatu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Koro', 'Region', 'Gaua']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Koro', 'Language', 'codes', 'ISO', '639-3', 'krf']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lo-Toga', 'Total', 'speakers', '580']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lo-Toga', 'Spoken', 'in', 'Vanuatu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lo-Toga', 'Region', 'Torres', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lo-Toga', 'Language', 'codes', 'ISO', '639-3', 'lht']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Numanggang', 'Language', 'codes', 'ISO', '639-3', 'nop']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mwerlap', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Mwerlap', 'Region', 'Merelava', ',', 'Gaua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Mwerlap', 'Language', 'codes', 'ISO', '639-3', 'mrm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taznatit', 'Spoken', 'in', 'Algeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Taznatit', 'Region', 'Gourara', '(', 'wilaya', 'of', 'Adrar', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Gourara', '\\  
(', 'wilaya', 'of', 'Adrar', '\\)']]], 8, True)
```

```
['Taznatit', 'Language', 'codes', 'ISO', '639-3', 'grr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zuara', 'Total', 'speakers', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zuara', 'Spoken', 'in', 'Libya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zuara', 'Region', 'Zuwarah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Zuara', 'Language', 'codes', 'ISO', '639-3', 'jbn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ouargli', 'Spoken', 'in', 'Algeria']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ouargli', 'Region', 'Ouargla', ',', 'N'Goussa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ouargli', 'Language', 'codes', 'ISO', '639-3', 'oua', 'Berber-speaking', 'areas', 'of', 'the', 'Mzab']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'oua', 'Berber-speaking', 'areas', 'of', 'the', 'Mzab']], ['NUM', '639-3']], 11, True)

['Ouargli', 'Language', 'codes', 'ISO', '639-3', 'Ouargla']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Macedonian', 'Signed', 'in', 'Macedonia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Macedonian', 'Language', 'codes', 'ISO', '639-2', 'sgn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Tanchangya', 'Total', 'speakers', '150000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tanchangya', 'Spoken', 'in', 'Bangladesh', ',', 'India', 'and', 'Myanmar']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Bangladesh', ',', 'India', 'and', 'Type: ANY']]], 8, True)

['Tanchangya', 'Region', 'Chittagong', 'Hill', 'Tracts']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Tanchangya', 'Region', 'Mizoram']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tanchangya', 'Region', 'Arunachal', 'Pradesh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Tanchangya', 'Region', 'Tripura']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tanchangya', 'Region', 'Rakhine', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tanchangya', 'Language', 'codes', 'ISO', '639-3', 'tnv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Matmata', 'Spoken', 'in', 'Tunisia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Matmata', 'Region', 'Matm', '\xc3', '\xa2', 'ta']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Matm', '\xc3', '\xa2', 'ta']], 6, True)

['Matmata', 'Language', 'codes', 'ISO', '639-3', 'jbn']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Uduk', 'Total', 'speakers', '20', ',', '000', 'in', 'Ethiopia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '20'], ['NUM', '000']], 8, True)

['Uduk', 'Spoken', 'in', 'Ethiopia', ',', 'Sudan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Uduk', 'Region', 'Gambela', 'region']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uduk', 'Writing', 'system', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Uduk', 'Language', 'codes', 'ISO', '639-3', 'udu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hiw', 'Total', 'speakers', '280']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hiw', 'Spoken', 'in', 'Vanuatu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Hiw', 'Region', 'Hiw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hiw', 'Language', 'codes', 'ISO', '639-3', 'hiw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ong', 'Region', 'Hainan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Ong', 'Language', 'codes', 'ISO', '639-3', 'onb']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Melayu', 'Spoken', 'in', 'Brunei']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Melayu', 'Spoken', 'in', 'Malaysia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Melayu', 'Total', 'speakers', '266000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Melayu', 'Language', 'codes', 'ISO', '639-2', 'msa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Melayu', 'Language', 'codes', 'ISO', '639-3', 'kxd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Asu', 'Spoken', 'in', 'Nigeria']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Asu', 'Language', 'codes', 'ISO', '639-3', 'aum']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Ilongot', 'Language', 'codes', 'ISO', '639-3', 'ilk']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Proto-Mon\xe2\x80\x93Khmer', 'Spoken', 'in', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']], 5, True)
```

```
['Proto-Mon\xe2\x80\x93Khmer', 'Total', 'speakers', '\xe2', '\x80',
'\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',
'\xe2', '\x80', '\x93']], 6, True)
```

```
['Proto-Mon\xe2\x80\x93Khmer', 'Language', 'codes', 'ISO', '639-3', '\xe2',
'\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Ambelau', 'Region', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Ambelau', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Ambelau', 'Writing', 'system', 'none']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Ambelau', 'Language', 'codes', 'ISO', '639-3', 'amv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Buru', 'Region', 'Southeast', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

```
['Buru', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)
```

['Buru', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Buru', 'Language', 'codes', 'ISO', '639-3', 'mhs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kayeli', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kayeli', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kayeli', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kayeli', 'Language', 'codes', 'ISO', '639-3', 'kzl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lisela', 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lisela', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lisela', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lisela', 'Language', 'codes', 'ISO', '639-3', 'lcl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cab\x3\xa9car', 'Spoken', 'in', 'Costa', 'Rica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Cab\x3\xa9car', 'Region', 'Turrialba', '(', 'Cartago', 'Province', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Turrialba', '\(\', 'Cartago', 'Province', '\\\)']]], 7, True)

['Cab\xc3\xa9car', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Cab\xc3\xa9car', 'Language', 'codes', 'ISO', '639-3', 'cjp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Amuzgo', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Amuzgo', 'Region', 'Guerrero']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Amuzgo', 'Language', 'codes', 'ISO', '639-3', 'amu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kibajuni', 'Spoken', 'in', 'Kenya']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kibajuni', 'Spoken', 'in', 'Somalia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kibajuni', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Kibajuni', 'Language', 'codes', 'ISO', '639-3', 'swl', '\xe2', '\x80', '\x93', 'Kibajuni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'swl', '\xe2', '\x80', '\x93', 'Type: ANY']]], ['NUM', '639-3']], 10, True)

['Kibajuni', 'Language', 'codes', 'ISO', '639-3', '(', 'generic', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\(, 'generic', '\\\)']]], ['NUM', '639-3']], 8, True)

['Kensiu', 'Total', 'speakers', '600']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kensiu', 'Spoken', 'in', 'Malaysia', ',', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Kensiu', 'Region', 'Northeast', 'Kedah']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kensiu', 'Region', 'Thailand']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kensiu', 'Dialects', 'Ijoh', '(', 'Ijok', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 6, True)

['Kensiu', 'Dialects', 'Jarum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kensiu', 'Dialects', 'Jeher', '(', 'Sakai', 'Tanjong', 'of', 'Temongoh', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Jeher', '\\\\(', 'Sakai', 'Tanjong', 'of', 'Temongoh', '\\\\)']]], 9, True)

['Kensiu', 'Dialects', 'Kedah', '(', 'Quedah', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 6, True)

['Kensiu', 'Dialects', 'Plus']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Kensiu', 'Dialects', 'Ulu', 'Selama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kensiu', 'Dialects', 'Kensiu', 'Batu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kensiu', 'Dialects', 'Kensiu', 'Siong']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Kensiu', 'Dialects', 'Kentaq', 'Nakil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kensiu', 'Language', 'codes', 'ISO', '639-3', 'kns']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paha', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Paha', 'Total', 'speakers', '600']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Paha', 'Language', 'codes', 'ISO', '639-3', 'yha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laha', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Laha', 'Language', 'codes', 'ISO', '639-3', 'lha']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lachi', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lachi', 'Language', 'codes', 'ISO', '639-3', 'lbt']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['En', 'Spoken', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['En', 'Total', 'speakers', '250']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['En', 'Language', 'codes', 'ISO', '639-3', 'enc']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kannada', 'Spoken', 'in']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in']]], 3,
True)

['Kannada', 'Language', 'codes', 'ISO', '639-3', 'kfi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nam', 'Language', 'extinction', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Nam', 'Spoken', 'in', 'Central', 'Asia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)

['Nam', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dorig', 'Total', 'speakers', '300']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Dorig', 'Spoken', 'in', 'Vanuatu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Dorig', 'Region', 'Gaua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Dorig', 'Language', 'codes', 'ISO', '639-3', 'wwo']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Toposa', 'Total', 'speakers', '100', ',', '000', '(', 'Ethnologue', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']], ['NUM', '100'], ['NUM',
'000']], 9, True)

['Toposa', 'Spoken', 'in', 'Sudan']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Toposa', 'Region', 'Eastern', 'Africa']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Toposa', 'Writing', 'system', 'none']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Toposa', 'Language', 'codes', 'ISO', '639-3', 'toq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Koro', 'Total', 'speakers', '400']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Koro', 'Region', 'northeastern', 'Manus', 'Island', ',', 'Manus', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northeastern', 'Manus', 'Island', ',', 'Manus', 'Province']]], 8, True)

['Koro', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Koro', 'Language', 'codes', 'ISO', '639-3', 'kxr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Proto-Tai', 'Spoken', 'in', 'Southern', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Proto-Tai', 'Spoken', 'in', 'Thailand']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proto-Tai', 'Spoken', 'in', 'Laos']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proto-Tai', 'Spoken', 'in', 'Burma']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Proto-Tai', 'Spoken', 'in', 'India']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Proto-Tai', 'Total', 'speakers', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Proto-Tai', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Maonan', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maonan', 'Region', 'Huanjiang', 'County']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maonan', 'Region', 'Hechi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Maonan', 'Region', 'northern', 'Guangxi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Maonan', 'Language', 'codes', 'ISO', '639-3', 'mmd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chadong', 'Spoken', 'in', 'China']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadong', 'Region', 'Northern', 'Guangxi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Chadong', 'Language', 'codes', 'ISO', '639-3', 'cdy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Then', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Then', 'Region', 'Pingtang', 'County', ',', 'southern', 'Guizhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Pingtang',  
'County', ',', 'southern', 'Guizhou']]], 7, True)

['Then', 'Language', 'codes', 'ISO', '639-3', 'tct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mulam', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mulam', 'Region', 'Luocheng', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mulam', 'Region', 'Hechi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mulam', 'Region', 'northern', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mulam', 'Language', 'codes', 'ISO', '639-3', 'mlm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mak', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mak', 'Region', 'Libo', 'County', ',', 'southern', 'Guizhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Libo',  
'County', ',', 'southern', 'Guizhou']]], 7, True)

['Mak', 'Language', 'codes', 'ISO', '639-3', 'mkg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lakkia', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Lakkia', 'Language', 'codes', 'ISO', '639-3', 'lbc']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Biao', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Biao', 'Language', 'codes', 'ISO', '639-3', 'byk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['ut-Ma\xe2\x80\x99in', 'Spoken', 'in', 'Fakai', 'district', ',',  
'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Fakai',  
'district', ',', 'Nigeria']]], 7, True)

['ut-Ma\xe2\x80\x99in', 'Language', 'codes', 'ISO', '639-3', 'gel']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warrwa', 'Total', 'speakers', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Warrwa', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Warrwa', 'Language', 'codes', 'ISO', '639-3', 'wwr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mamanwa', 'Spoken', 'in', 'Philippines']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mamanwa', 'Region', 'Agusan', 'del', 'Norte', 'and', 'Surigao',  
'provinces', ',', 'Mindanao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Agusan',  
'del', 'Norte', 'and', 'Surigao', 'provinces', ',', 'Mindanao']]], 10,  
True)

['Mamanwa', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Mamanwa', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Regional', 'language', 'in', 'the', 'Philippines']], 9, True)

['Mamanwa', 'Language', 'codes', 'ISO', '639-3', 'mmn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fipa', 'Spoken', 'in', 'Tanzania']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fipa', 'Language', 'codes', 'ISO', '639-3', 'fip']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ganza', 'Total', 'speakers', '5', ',', '400', '(', 'Ethnologue', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']], ['NUM', '5'], ['NUM', '400']], 9, True)

['Ganza', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ganza', 'Region', 'Asosa', 'Zone', 'of', 'Benishangul-Gumuz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

['Ganza', 'Writing', 'system', 'none']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ganza', 'Language', 'codes', 'ISO', '639-3', 'gza']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kwama', 'Spoken', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kwama', 'Region', 'Benishangul-Gumuz']:



```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Kwama', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kwama', 'Language', 'codes', 'ISO', '639-3', 'kmq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kwegu', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kwegu', 'Region', 'Southwest', ',', 'Omo', 'River', 'west', 'bank']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Southwest', ',', 'Omo', 'River', 'west', 'bank']], 8, True)
```

```
['Kwegu', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Kwegu', 'Language', 'codes', 'ISO', '639-3', 'xwg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Mesqan', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Mesqan', 'Region', 'West', 'Gurage', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Mesqan', 'Writing', 'system', 'none']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Mesqan', 'Language', 'codes', 'ISO', '639-3', 'mvz']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Murle', 'Spoken', 'in', 'Sudan', ',', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)
```

['Murle', 'Total', 'speakers', '60', ',', '000', 'in', 'Sudan', ',', '470', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Sudan', ',', 'NUM', 'in', 'Ethiopia']], ['NUM', '60'], ['NUM', '000'], ['NUM', '470']], 12, True)

['Murle', 'Writing', 'system', 'Latin', '(', 'in', 'Sudan', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', '.', 'Type: ANY', '.', 'Type: ANY']], 8, True)

['Murle', 'Language', 'codes', 'ISO', '639-3', 'mur']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Suri', 'Total', 'speakers', '27', ',', '886', 'in', 'Ethiopia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '27'], ['NUM', '886']], 8, True)

['Suri', 'Total', 'speakers', '1', ',', '000', 'in', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']], ['NUM', '1'], ['NUM', '000']], 8, True)

['Suri', 'Spoken', 'in', 'Ethiopia', ',', 'Sudan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Suri', 'Region', 'Bench', 'Maji', 'Zone']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Suri', 'Writing', 'system', 'Ethiopic']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Suri', 'Language', 'codes', 'ISO', '639-3', 'suq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Seze', 'Total', 'speakers', '3', ',', '000', '(', 'Ethnologue', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', '\\\\(', 'Ethnologue', '\\\\)']], ['NUM', '3'], ['NUM', '000']], 9, True)

```
['Seze', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Seze', 'Region', 'West', 'Oromo', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Seze', 'Language', 'codes', 'ISO', '639-3', 'sze']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shekkacho', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Shekkacho', 'Region', 'North', 'Kafa', 'region']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Shekkacho', 'Writing', 'system', 'Ethiopic', '/', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)
```

```
['Shekkacho', 'Language', 'codes', 'ISO', '639-3', 'moy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Zayse-Zergulla', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zayse-Zergulla', 'Writing', 'system', 'Ethiopic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zayse-Zergulla', 'Language', 'codes', 'ISO', '639-3', 'zay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Koorete', 'Spoken', 'in', 'Ethiopia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Koorete', 'Writing', 'system', 'Ethiopic', '/', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
```

'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Koorete', 'Language', 'codes', 'ISO', '639-3', 'kqy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ait', 'Seghrouchen', 'Berber', 'Total', 'speakers', '\xe2', '\x80',  
\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Seghrouchen', 'Berber',  
'Total', 'speakers', '\xe2', '\x80', '\x93']], 8, True)

['Ait', 'Spoken', 'in', 'Morocco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ait', 'Writing', 'system', 'Tifinagh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ait', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ait', 'Writing', 'system', 'Arabic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ait', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Ait', 'Language', 'codes', 'ISO', '639-3', 'tzm-cen']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

["Makuv'a", 'Spoken', 'in', 'East', 'Timor']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

["Makuv'a", 'Region', 'Southeast', 'Asia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

["Makuv'a", 'Language', 'codes', 'ISO', '639-3', 'lva']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Fiuman', 'Spoken', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fiuman', 'Spoken', 'in', 'Italy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fiuman', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fiuman', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fiuman', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Fiuman', 'Region', 'Kvarner', 'Gulf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Fiuman', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Proto-Hlai', 'Spoken', 'in', 'Hainan', ',', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']], 6, True)

['Proto-Hlai', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']], 6, True)

['Proto-Hlai', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Proto-Kra', 'Spoken', 'in', 'Northern', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:

ANY', 'Type: ANY']]], 5, True)

['Proto-Kra', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Proto-Kra', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Teanu', 'Spoken', 'in', 'Solomon', 'Islands']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Teanu', 'Language', 'codes', 'ISO', '639-3', 'tkw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Khuzdul', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)

['Khuzdul', 'Category', '(', 'purpose', ')', 'constructed', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Khuzdul', 'Category', '(', 'purpose', ')', 'artistic', 'languages']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY', 'language']]], 7, True)

['Khuzdul', 'Category', '(', 'purpose', ')', 'Khuzdul']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(',  
'purpose', '\\)', 'Type: ANY']]], 6, True)

['Khuzdul', 'Created', 'by', 'J', '.', 'R', '.', 'R', '.', 'Tolkien']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'J',  
'.', 'R', '.', 'R', '.', 'Tolkien']]], 10, True)

['Khuzdul', 'Setting', 'and', 'usage', 'Middle-earth', ',', 'the',  
'setting', 'of', 'the', 'novel', 'The', 'Lord', 'of', 'the', 'Rings']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and',  
'usage', 'Middle-earth', ',', 'the', 'setting', 'of', 'the', 'novel',  
'The', 'Lord', 'of', 'the', 'Rings']]], 16, True)

['Khuzdul', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Balanta', 'Total', 'speakers', '423', ',', '000', 'for', 'Balanta-Kentohe', 'and', '86', ',', '000', 'for', 'Balanta-Ganja']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'for', 'Balanta-Kentohe', 'and', 'NUM', ',', 'NUM', 'for', 'Balanta-Ganja']], ['NUM', '423'], ['NUM', '000'], ['NUM', '86'], ['NUM', '000']], 14, True)

['Balanta', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Balanta', 'Language', 'codes', 'ISO', '639-3', ':', 'ble', '\xe2', '\x80', '\x93', 'Balanta-Kentohe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'ble', '\xe2', '\x80', '\x93', 'Balanta-Kentohe']], ['NUM', '639-3']], 11, True)

['Balanta', 'Language', 'codes', 'ISO', '639-3', 'bjt', '\xe2', '\x80', '\x93', 'Balanta-Ganja']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'bjt', '\xe2', '\x80', '\x93', 'Balanta-Ganja']], ['NUM', '639-3']], 10, True)

['Enggano', 'Total', 'speakers', '1500', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '?']], ['NUM', '1500']], 5, True)

['Enggano', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Enggano', 'Language', 'codes', 'ISO', '639-3', 'eno']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Roglai', 'Spoken', 'in', 'Vietnam']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Roglai', 'Region', 'Southeast', 'Asia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Roglai', 'Official', 'language', 'in', 'Vietnam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY']]], 5, True)

['Roglai', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Roglai', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Pa-Hng', 'Spoken', 'in', 'China', ',', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Pa-Hng', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Pa-Hng', 'Language', 'codes', 'ISO', '639-3', 'pha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Krenak', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Krenak', 'Region', 'Mato', 'Grosso', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Krenak', 'Region', 'Botocudo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Krenak', 'Language', 'codes', 'ISO', '639-3', 'kqq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Oti', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Oti', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)



['Oti', 'Region', 'S', '\xc3', '\xa3', 'o', 'Paulo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'S', '\xc3',  
'\xa3', 'o', 'Paulo']]], 7, True)

['Oti', 'Language', 'codes', 'ISO', '639-3', 'oti']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aghwan', 'Language', 'extinction', 'developed', 'into', 'Udi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'developed', 'into', 'Type: ANY']]], 6, True)

['Aghwan', 'Spoken', 'in', 'Caucasian', 'Albania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Aghwan', 'Writing', 'system', 'Caucasian', 'Albanian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Aghwan', 'Language', 'codes', 'ISO', '639-3', 'xag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kui', 'Spoken', 'in', 'Alor', 'Island']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Kui', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kui', 'Language', 'codes', 'ISO', '639-3', 'kvd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pakanha', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Pakanha', 'Language', 'codes', 'ISO', '639-3', 'pkn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kuku', 'Spoken', 'in', 'Annan', 'River', ',', 'Queensland']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'River', ',', 'Type: ANY']]], 7, True)

['Kuku', 'Total', 'speakers', '5']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kuku', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kuku', 'Language', 'codes', 'ISO', '639-3', 'gvn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Phom', 'Spoken', 'in', 'Nagaland', ',', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Phom', 'Language', 'codes', 'ISO', '639-3', 'nph']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Whitesands', 'Total', 'speakers', '7500']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Whitesands', 'Spoken', 'in', 'Vanuatu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Whitesands', 'Region', 'Tanna', 'Island', ',', 'east', 'coast']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Tanna', 'Island', ',', 'east', 'coast']]], 7, True)

['Whitesands', 'Language', 'codes', 'ISO', '639-3', 'tnp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Truma\x3\xad', 'Total', 'speakers', '120']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Truma\x3\xad', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Truma\xc3\xad', 'Region', 'upper', 'Xingu', 'River', 'Ethnicity']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'upper',  
'Xingu', 'River', 'Ethnicity']]]], 6, True)
```

```
['Truma\xc3\xad', 'Region', 'Trumai', 'people']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Truma\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'tpy']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Mushunguli', 'Spoken', 'in', 'Somalia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Mushunguli', 'Language', 'codes', 'ISO', '639-3', 'xma']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Kposo', 'Spoken', 'in', 'Ghana']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Kposo', 'Region', 'Volta']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)
```

```
['Kposo', 'Language', 'codes', 'ISO', '639-3', 'kpo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Akebu', 'Spoken', 'in', 'Ghana', ',', 'Togo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)
```

```
['Akebu', 'Language', 'codes', 'ISO', '639-3', 'keu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Graus', 'Aragonese', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)
```

```
['Graus', 'Spoken', 'in', 'Aragon', ',', 'Spain']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Graus', 'Region', 'Ribagorza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Graus', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Bunu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bunu', 'Region', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bunu', 'Region', 'Yunnan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Bunu', 'Region', 'and', 'Guizhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bunu', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Bunu', 'Language', 'codes', 'ISO', '639-3', 'bwx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jiongnai', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jiongnai', 'Region', 'Dahua', 'County', ',', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Dahua', 'County', ',', 'Guangxi']]], 6, True)

['Jiongnai', 'Language', 'codes', 'ISO', '639-3', 'pnu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Younuo', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Younuo', 'Region', 'Longsheng', 'County']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Younuo', 'Region', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Younuo', 'Region', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Younuo', 'Language', 'codes', 'ISO', '639-3', 'buh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Bunu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bunu', 'Region', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Bunu', 'Region', 'Yunnan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Bunu', 'Region', 'and', 'Guizhou']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Bunu', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Bunu', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Dzala', 'Spoken', 'in', 'Bhutan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Dzala', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Dzala', 'Language', 'codes', 'ISO', '639-3', 'dzl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ba-Shu', 'Spoken', 'in', 'China']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ba-Shu', 'Region', 'Sichuan', 'and', 'Chongqing']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ba-Shu', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Akose', 'Spoken', 'in', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Akose', 'Language', 'codes', 'ISO', '639-3', 'bss']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mbule', 'Spoken', 'in', 'southwestern', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Mbule', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Mbule', 'Language', 'codes', 'ISO', '639-3', 'mlb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Olekha', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

["'Olekha", 'Writing', 'system', 'Tibetan', 'script']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

["'Olekha", 'Language', 'codes', 'ISO', '639-3', 'ole']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Pannonian', 'Spoken', 'in', 'Serbia', ',', 'Croatia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', ',', 'Type: ANY']]], 6, True)

['Pannonian', 'Total', 'speakers', '20', ',', '000', 'approx']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', ',', 'NUM', 'approx']]], ['NUM', '20'], ['NUM', '000']], 7, True)

['Pannonian', 'Official', 'language', 'in', 'Vojvodina', 'province', '(',
'Serbia', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Vojvodina', 'province', '\\(', 'Serbia', '\\)']]], 9, True)

['Pannonian', 'Regulated', 'by', 'Statute', 'of', 'Vojvodina']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:
ANY', 'of', 'Type: ANY']]], 6, True)

['Pannonian', 'Language', 'codes', 'ISO', '639-2', 'sla', '(', 'B', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'sla', '\\(', 'B', '\\)']]], ['NUM', '639-2']], 9, True)

['Pannonian', 'Language', 'codes', 'ISO', '639-2', 'rue', '(', 'T', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'rue', '\\(', 'T', '\\)']]], ['NUM', '639-2']], 9, True)

['Pannonian', 'Language', 'codes', 'ISO', '639-3', 'rue']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Moroccan', 'Spoken', 'in', 'Morocco']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Moroccan', 'Writing', 'system', 'Arabic', 'alphabet']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

```

```
['Moroccan', 'Pronunciation', 'm', '\xc9', '\x91', '\xc9', '\xa3', 'ribi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'm',
'\xc9', '\x91', '\xc9', '\xa3', 'ribi']]], 8, True)
```

```
['Moroccan', 'Language', 'codes', 'ISO', '639-2', 'ara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Moroccan', 'Language', 'codes', 'ISO', '639-3', 'ary']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pijin', 'Spoken', 'in', 'Solomon', 'Islands']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
ANY', 'Type: ANY']]], 5, True)
```

```
['Pijin', 'Language', 'codes', 'ISO', '639-2', 'cpe']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Pijin', 'Language', 'codes', 'ISO', '639-3', 'pis']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Pijin', 'Language', 'codes', 'Linguasphere', '52-ABB-cd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'Linguasphere', 'Type: ANY']]], 5, True)
```

```
['Banyumasan', 'Total', 'speakers', '12', '\xe2', '\x80', '\x93', '15',
'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',
'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']]], ['NUM', '12'], ['NUM',
'15']], 9, True)
```

```
['Banyumasan', 'Language', 'codes', 'ISO', '639-2', 'jav']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)
```

```
['Banyumasan', 'Language', 'codes', 'ISO', '639-3', 'jav', '\xe2', '\x80',
'\x93', 'Javanese']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'jav', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)
```

```
['Banyumasan', 'Language', 'codes', 'ISO', '639-1', 'jv']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```



'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Min', 'Language', 'codes', 'ISO', '639-2', 'chi', '(', 'B', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'chi', '\\\\(', 'B', '\\\\)']], ['NUM', '639-2']], 9, True)

['Min', 'Language', 'codes', 'ISO', '639-2', 'zho', '(', 'T', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zho', '\\\\(', 'T', '\\\\)']], ['NUM', '639-2']], 9, True)

['Min', 'Language', 'codes', 'ISO', '639-3', 'mnp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Min', 'Language', 'codes', 'ISO', '639-1', 'zh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-1']], 6, True)

['Chocangacakha', 'Spoken', 'in', 'Bhutan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Chocangacakha', 'Writing', 'system', 'Tibetan', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Chocangacakha', 'Language', 'codes', 'ISO', '639-3', 'cgk']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lakha', 'Spoken', 'in', 'Bhutan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Lakha', 'Writing', 'system', 'Tibetan', 'script']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Lakha', 'Language', 'codes', 'ISO', '639-3', 'lkh']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Brokkat', 'Spoken', 'in', 'Bhutan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

```
['Brokkat', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Brokkat', 'Language', 'codes', 'ISO', '639-3', 'bro']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Brokpa', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Brokpa', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Brokpa', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)
```

```
['Brokpa', 'Language', 'codes', 'ISO', '639-3', 'sgt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Bumthang', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bumthang', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Bumthang', 'Language', 'codes', 'ISO', '639-3', 'kjj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Dakpa', 'Spoken', 'in', 'Bhutan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Dakpa', 'Writing', 'system', 'Tibetan', 'script']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Dakpa', 'Language', 'codes', 'ISO', '639-3', 'dka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chali', 'Spoken', 'in', 'Bhutan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chali', 'Writing', 'system', 'Tibetan', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Chali', 'Language', 'codes', 'ISO', '639-3', 'tgf']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyenkha', 'Spoken', 'in', 'Bhutan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Nyenkha', 'Writing', 'system', 'Tibetan', 'script']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Nyenkha', 'Language', 'codes', 'ISO', '639-3', 'neh']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Russian', 'Region', 'Commonwealth', 'of', 'Independent', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Commonwealth', 'of', 'Independent', 'States']], 6, True)

['Russian', 'Signed', 'in', 'Russia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Signed', 'in', 'Ukraine']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Signed', 'in', 'Belorussia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Signed', 'in', 'Kazakhstan']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Russian', 'Signed', 'in', 'Moldova']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Russian', 'Signed', 'in', 'Bulgaria', ';', 'partly', 'in', 'Latvia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in',  
'Bulgaria', ';', 'partly', 'in', 'Latvia']]], 8, True)

['Russian', 'Signed', 'in', 'Estonia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Russian', 'Signed', 'in', 'Lithuania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Russian', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dargwa', 'Spoken', 'in', 'Russia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Dargwa', 'Region', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Dargwa', 'Official', 'language', 'in', 'Dagestan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Dargwa', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Dargwa', 'Language', 'codes', 'ISO', '639-2', 'dar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Dargwa', 'Language', 'codes', 'ISO', '639-3', 'dar']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gaulish', 'Spoken', 'in', 'Gaul']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Gaulish', 'Writing', 'system', 'Old', 'Italic', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

['Gaulish', 'Writing', 'system', 'Greek', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Gaulish', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Gaulish', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gaulish', 'Language', 'codes', 'ISO', '639-3', ':', 'xtg', '\xe2',  
'\x80', '\x93', 'Transalpine', 'Gaulish']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'xtg', '\xe2', '\x80', '\x93', 'Transalpine', 'Type:  
ANY']], ['NUM', '639-3']], 12, True)

['Gaulish', 'Language', 'codes', 'ISO', '639-3', 'xlp', '\xe2', '\x80',  
'\x93', 'Leponitic']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'xlp', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Gaulish', 'Language', 'codes', 'ISO', '639-3', 'xga', '\xe2', '\x80',  
'\x93', 'Galatian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'xga', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',  
'639-3']], 10, True)

['Terengganuan', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Terengganuan', 'Region', 'Terengganu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Terengganuan', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Terengganuan', 'Language', 'codes', 'ISO', '639-3', ':', 'zlm-inl',  
'\xe2', '\x80', '\x93', '(', 'inland', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'zlm-inl', '\xe2', '\x80', '\x93', '\\(', 'inland', '\\  
)']], ['NUM', '639-3']], 13, True)

['Terengganuan', 'Language', 'codes', 'ISO', '639-3', 'zlm-coa', '\xe2',  
'\x80', '\x93', '(', 'coastal', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'zlm-coa', '\xe2', '\x80', '\x93', '\\(', 'coastal', '\\  
)']], ['NUM', '639-3']], 12, True)

['Tepes', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tepes', 'Total', 'speakers', 'few']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tepes', 'Language', 'codes', 'ISO', '639-3', 'teu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ik', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ik', 'Language', 'codes', 'ISO', '639-3', 'ikx']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nyangia', 'Spoken', 'in', 'Uganda']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nyangia', 'Total', 'speakers', '0']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Nyangia', 'Language', 'codes', 'ISO', '639-3', 'nyp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Lower', 'Spoken', 'in', 'Huai', 'and', 'Yangzi', 'Rivers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Huai',  
'and', 'Yangzi', 'Rivers']]], 7, True)
```

```
['Lower', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Lower', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Lower', 'Language', 'codes', 'ISO', '639-3', 'cmn-jin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Lower', 'Language', 'codes', 'ISO', '639-3', 'Jianghuai', 'is', 'in',  
'pale', 'blue']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Jianghuai', 'is', 'in', 'pale', 'blue']], ['NUM', '639-3']],  
10, True)
```

```
['Lan\xe2\x80\x93Yin', 'Spoken', 'in', 'Silk', 'Road']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Lan\xe2\x80\x93Yin', 'Total', 'speakers', '10', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'NUM', 'million']], ['NUM', '10']], 5, True)
```

```
['Lan\xe2\x80\x93Yin', 'Language', 'codes', 'ISO', '639-3', 'cmn-lan',  
'\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'cmn-lan', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9,  
True)
```

```
['Rwanda-Rundi', 'Spoken', 'in', 'Rwanda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Rwanda-Rundi', 'Spoken', 'in', 'Burundi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

['Rwanda-Rundi', 'Spoken', 'in', 'Uganda']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'DR', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Tanzania', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Hutu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Tutsi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Twa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Ha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Shubi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Hangaza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Spoken', 'in', 'Vinza']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rwanda-Rundi', 'Total', 'speakers', '14', '\xe2', '\x80', '\x93', '18', 'million']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\xe2', '\x80', '\x93', 'NUM', 'million']]], ['NUM', '14'], ['NUM', '18']], 9, True)



```

['Rwanda-Rundi', 'Dialects', 'Kinyarwanda']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Rwanda-Rundi', 'Dialects', 'Kirundi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Rwanda-Rundi', 'Dialects', 'Ha']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Rwanda-Rundi', 'Dialects', 'Shubi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Rwanda-Rundi', 'Dialects', 'Hangaza']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Rwanda-Rundi', 'Dialects', 'Vinza']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Rwanda-Rundi', 'Official', 'language', 'in', 'Rwanda', ',', 'Burundi']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Type: ANY', ',', 'Type: ANY']]], 7, True)

['Rwanda-Rundi', 'Regulated', 'by', 'No', 'official', 'regulation']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']]], 6, True)

['Rwanda-Rundi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 8, True)

['Apinay\xc3\xa9', 'Spoken', 'in', 'Brazil']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Apinay\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'apn']:
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Cubeo', 'Spoken', 'in', 'Brazil', ',', 'Colombia']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Cubeco', 'Language', 'codes', 'ISO', '639-3', 'cub']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lyons', 'Signed', 'in', 'France']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lyons', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lyons', 'Language', 'codes', 'ISO', '639-2', 'sgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Armenian', 'Signed', 'in', 'Armenia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Armenian', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ecuadorian', 'Signed', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ecuadorian', 'Language', 'codes', 'ISO', '639-2', 'sgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Nepali', 'Signed', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rennellese', 'Signed', 'in', 'Solomon', 'Islands']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'Solomon', 'Islands']]], 5, True)

['Rennellese', 'Total', 'signers', 'few']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Rennellese', 'Language', 'codes', 'ISO', '639-2', 'sgn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Saudi', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Peruvian', 'Signed', 'in', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Peruvian', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tijuana', 'Signed', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Tijuana', 'Total', 'signers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Croatian', 'Signed', 'in', 'Croatia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Swiss', 'Signed', 'in', 'Switzerland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Swiss', 'Writing', 'system', 'SignWriting']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Swiss', 'Dialects', 'Swiss', 'German']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Swiss', 'Dialects', 'Swiss', 'French']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Swiss', 'Dialects', 'Swiss', 'Italian']:

```
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Swiss', 'Language', 'codes', 'ISO', '639-3', 'variously']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Bugan', 'Spoken', 'in', 'China']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Bugan', 'Region', 'Longlin', 'County', ',', 'Guangxi']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Longlin', 'County', ',', 'Guangxi']]], 6, True)
```

```
['Bugan', 'Language', 'codes', 'ISO', '639-3', 'ply']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Darling', 'Spoken', 'in', 'Australia']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Darling', 'Total', 'speakers', '5', '(', 'Ethnologue', ')']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\(', 'Ethnologue', '\\)']], ['NUM', '5']], 7, True)
```

```
['Darling', 'Language', 'codes', 'ISO', '639-3', 'drl']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Plateau', 'Region', 'Columbia', 'Plateau']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```
['Plateau', 'Signed', 'in', 'Canada', ',', 'USA']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Signed', 'in', 'Canada', ',', 'USA']]], 6, True)
```

```
['Inuit', 'Region', 'Arctic']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)
```

```
['Inuit', 'Signed', 'in', 'Canada']:  
(['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)
```

```

['Inuit', 'Total', 'signers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Mehek', 'Region', 'Sandaun', 'Province']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mehek', 'Language', 'codes', 'ISO', '639-3', 'nux']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ugandan', 'Signed', 'in', 'Uganda']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Ugandan', 'Total', 'signers', '(', 'unknown', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\\
(', 'unknown', '\\)']]], 6, True)

['Canara', 'Konkani', 'Total', 'speakers', '2', 'million']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total',
'speakers', 'NUM', 'million']], ['NUM', '2']], 6, True)

['Canara', 'Region', 'Karnataka', 'and', 'Kerala', '(', 'See', 'text',
')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Karnataka',
'and', 'Kerala', '\\((', 'See', 'text', '\\)')]], 9, True)

['Canara', 'Spoken', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Canara', 'Writing', 'system', 'Devanagari', '(', 'official', ')']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', '\\((', 'Type: ANY', '\\)')]], 7, True)

['Canara', 'Writing', 'system', 'Roman', 'Kannada']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY']]], 5, True)

['Canara', 'Writing', 'system', 'Malayalam', 'and', 'Persian']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',
'Type: ANY', 'Type: ANY', 'Type: ANY']]], 6, True)

```

```

['Canara', 'Pronunciation', 'ko', '\xc5', '\x8b', 'k', '\xc9', '\xb3',
'i']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'ko',
'\xc5', '\x8b', 'k', '\xc9', '\xb3', 'i']]], 9, True)

['Canara', 'Official', 'language', 'in', 'India']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',
'in', 'Type: ANY']]], 5, True)

['Canara', 'Recognised', 'minority', 'language', 'in', 'Karnataka']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'Type: ANY']]], 6, True)

['Canara', 'Recognised', 'minority', 'language', 'in', 'Kerala']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in', 'Type: ANY']]], 6, True)

['Canara', 'Recognised', 'minority', 'language', 'in']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Recognised', 'minority',
'language', 'in']]], 5, True)

['Canara', 'Language', 'codes', 'ISO', '639-2', 'kok']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Canara', 'Language', 'codes', 'ISO', '639-3', 'variously']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Canara', 'Language', 'codes', 'ISO', '639-3', ':', 'kex', '\xe2', '\x80',
'\x93', 'Kanara', 'Konkani']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', ':', 'kex', '\xe2', '\x80', '\x93', 'Kanara', 'Type: ANY']],
['NUM', '639-3']], 12, True)

['Canara', 'Language', 'codes', 'ISO', '639-3', 'gom', '\xe2', '\x80',
'\x93', 'Goan', 'Konkani']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'gom', '\xe2', '\x80', '\x93', 'Goan', 'Type: ANY']], ['NUM',
'639-3']], 11, True)

['Mbara', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Mbara', 'Total', 'speakers', '0', '(', 'extinct', ')']:

```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', '\\(', 'extinct', '\\)']], ['NUM', '0']], 7, True)

['Mbara', 'Language', 'codes', 'ISO', '639-3', 'mvl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Laua', 'Language', 'codes', 'ISO', '639-3', 'luf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nunivak', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Nunivak', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Nunivak', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Nunivak', 'Language', 'codes', 'ISO', '639-3', 'esu-nun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chevak', 'Cup', '\\xe2', '\\x80', '\\x99', 'ik', 'language', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Cup', '\\xe2', '\\x80', '\\x99', 'ik', 'language', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']], 12, True)

['Chevak', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Chevak', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Chevak', 'Language', 'codes', 'ISO', '639-3', 'esu-hoo', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

['ISO', 'NUM', 'esu-hoo', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Utkuhiksalik', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\xe2', '\x80', '\x93']]], 6, True)

['Utkuhiksalik', 'Spoken', 'in', 'Canada']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Utkuhiksalik', 'Region', 'North', 'America']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Utkuhiksalik', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Isanzu', 'Spoken', 'in', 'Tanzania']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Isanzu', 'Total', 'speakers', '32', ',', '000', 'in', '1987']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', ',', 'NUM', 'in', 'Type: ANY']]], ['NUM', '32'], ['NUM', '000']], 8, True)

['Isanzu', 'Language', 'codes', 'ISO', '639-3', 'isn']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Old', 'Burmese', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Total', 'speakers', '\xe2', '\x80', '\x93']]], 7, True)

['Old', 'Spoken', 'in', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', '\xe2', '\x80', '\x93']]], 6, True)

['Old', 'Region', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Writing', 'system', 'Burmese', 'script']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'obr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Kalusa', 'Regulated', 'by', 'Community', 'moderation']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'Type: ANY']]], 5, True)

['Kalusa', 'Total', 'speakers', 'Several', 'with', 'a', 'recorded', 'level', 'of', 'proficiency']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'Several', 'with', 'a', 'recorded', 'level', 'of', 'proficiency']]], 10, True)

['Kalusa', 'Date', 'founded', '2006']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Kalusa', 'Category', '(', 'purpose', ')', 'Constructed', 'language', 'with', 'elements', 'of', 'the', 'subgenres', 'artistic', 'language', 'and', 'personal', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'purpose', '\\)', 'Constructed', 'language', 'with', 'elements', 'of', 'the', 'subgenres', 'artistic', 'language', 'and', 'personal', 'language']]], 17, True)

['Kalusa', 'Pronunciation', '/', 'k', '\\xc3', '\\xa6', '\\xcb', '\\x88', 'lu', '\\xcb', '\\x90', 's', '\\xc9', '\\x99', '/', 'or', '/', '\\xcb', '\\x88', 'k', '\\xc3', '\\xa6', 'lu', '\\xcb', '\\x90', 's', '\\xc9', '\\x99', '/']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '/', 'k', '\\xc3', '\\xa6', '\\xcb', '\\x88', 'lu', '\\xcb', '\\x90', 's', '\\xc9', '\\x99', '/', 'or', '/', '\\xcb', '\\x88', 'k', '\\xc3', '\\xa6', 'lu', '\\xcb', '\\x90', 's', '\\xc9', '\\x99', '/']]], 29, True)

['Kalusa', 'Created', 'by', 'Gary', 'Shannon', ',', 'et', 'al']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Created', 'by', 'Gary', 'Shannon', ',', 'et', 'al']]], 8, True)

['Kalusa', 'Category', '(', 'sources', ')', 'An', 'a', 'posteriori', 'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'An', 'a', 'posteriori', 'language']]], 9, True)

['Kalusa', 'Category', '(', 'sources', ')', 'with', 'elements', 'of', 'Tazhu']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'with', 'elements', 'of', 'Type: ANY']]], 9, True)

['Kalusa', 'Category', '(', 'sources', ')', 'Madjal']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'Type: ANY']]], 6, True)

['Kalusa', 'Category', '(', 'sources', ')', 'Swahili']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'Type: ANY']]], 6, True)

['Kalusa', 'Category', '(', 'sources', ')', 'Tok', 'Pisin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'Type: ANY', 'Type: ANY']]], 7, True)

['Kalusa', 'Category', '(', 'sources', ')', 'and', 'Indo-European', 'languages']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Category', '\\(', 'sources', '\\)', 'and', 'Indo-European', 'languages']]], 8, True)

['Kalusa', 'Setting', 'and', 'usage', 'Testing', 'principles', 'of', 'collaborative', 'corpus', 'driven', 'language']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Setting', 'and', 'usage', 'Testing', 'principles', 'of', 'collaborative', 'corpus', 'driven', 'language']]], 11, True)

['Kalusa', 'Language', 'codes', 'ISO', '639-2', 'art']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-2']], 6, True)

['Kalusa', 'Language', 'codes', 'ISO', '639-3', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 8, True)

['Natsilik', 'Total', 'speakers', '\\xe2', '\\x80', '\\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers', '\\xe2', '\\x80', '\\x93']]], 6, True)

['Natsilik', 'Spoken', 'in', 'Canada']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Natsilik', 'Region', 'North', 'America']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Natsilik', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Natsilik', 'Language', 'codes', 'ISO', '639-3', 'ikt-net']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Patax\xc3\xb3', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Patax\xc3\xb3', 'Total', 'speakers', '(', 'Extinct', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', '\\(', 'Extinct', '\\)']], 6, True)

['Patax\xc3\xb3', 'Language', 'codes', 'ISO', '639-3', 'pth']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Aghem', 'Spoken', 'in', 'Cameroon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Aghem', 'Region', 'Menchum']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Aghem', 'Language', 'codes', 'ISO', '639-3', 'agq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vaccarizzo', 'Spoken', 'in', 'Italy']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Vaccarizzo', 'Writing', 'system', 'Latin', 'alphabet']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']], 5, True)

['Vaccarizzo', 'Language', 'codes', 'ISO', '639-3', 'aae']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Vaccarizzo', 'Language', 'codes', 'Linguasphere', '55-AAA-ahd']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'Linguasphere', 'Type: ANY']]]], 5, True)

['Kalix', 'Pronunciation', 'kj', '\xc5', '\x93', '\xc9', '\xad', '\xc9',  
'\xaa', 's', '\xcb', '\x88', 'm', '\xc9', '\x94', '\xcb', '\x90', '\xc9',  
'\xad', '\xc9', '\x9b']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', 'kj',  
'\xc5', '\x93', '\xc9', '\xad', '\xc9', '\xaa', 's', '\xcb', '\x88', 'm',  
'\xc9', '\x94', '\xcb', '\x90', '\xc9', '\xad', '\xc9', '\x9b']]]], 21,  
True)

['Kalix', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['If\xc3\xa8', 'Spoken', 'in', 'Togo', ',', 'Benin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]]], 6, True)

['If\xc3\xa8', 'Writing', 'system', 'Latin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

['If\xc3\xa8', 'Dialects', 'Tschetti']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['If\xc3\xa8', 'Dialects', 'Djama']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['If\xc3\xa8', 'Dialects', 'Dadja']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]]], 3, True)

['If\xc3\xa8', 'Language', 'codes', 'ISO', '639-3', 'ife']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tokano', 'Dialects', 'Lower', 'Asaro']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]]], 4, True)

```

['Tokano', 'Dialects', 'Zuhuzuho']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Tokano', 'Writing', 'system', 'Latin']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tokano', 'Language', 'codes', 'ISO', '639-3', 'zuh']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Zambian', 'Signed', 'in', 'Zambia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zambian', 'Total', 'signers', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'signers', '\xe2', '\x80', '\x93']], 6, True)

['Zangwal', 'Spoken', 'in', 'Nigeria']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zangwal', 'Region', 'Bauchi', 'State']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Zangwal', 'Language', 'codes', 'ISO', '639-3', 'zah']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Old', 'Spoken', 'in', 'Sultanate', 'of', 'R', '\xc3', '\xbb', 'm', ',', 'Anatolian', 'Beyliks']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Sultanate', 'of', 'R', '\xc3', '\xbb', 'm', ',', 'Anatolian', 'Beyliks']], 12, True)

['Old', 'Official', 'language', 'in', 'Karamano', '\xc4', '\x9f', 'lu']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language', 'in', 'Karamano', '\xc4', '\x9f', 'lu']], 8, True)

['Old', 'Regulated', 'by', 'No', 'official', 'regulation']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type: ANY', 'of', 'Type: ANY']], 6, True)

```

```
['Old', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Zaramo', 'Total', 'speakers', 'few']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zaramo', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zaramo', 'Region', 'Pwani']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Zaramo', 'Writing', 'system', 'Latin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Zaramo', 'Language', 'codes', 'ISO', '639-3', 'zaj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Burduna', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Burduna', 'Language', 'codes', 'ISO', '639-3', 'bxn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Djeebbana', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Djeebbana', 'Total', 'speakers', 'none', 'fluent']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'none', 'fluent']]], 5, True)
```

```
['Djeebbana', 'Language', 'codes', 'ISO', '639-3', 'djj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Burarra', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'Type: ANY']]], 5, True)

['Burarra', 'Total', 'speakers', '700']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Burarra', 'Language', 'codes', 'ISO', '639-3', 'bvr']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nakkara', 'Spoken', 'in', 'Northern', 'Territory']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Nakkara', 'Language', 'codes', 'ISO', '639-3', 'nck']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Andegerebinha', 'Spoken', 'in', 'Australia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Andegerebinha', 'Language', 'codes', 'ISO', '639-3', 'adg']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Alumu-Tesu', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Alumu-Tesu', 'Region', 'Nassarawa', 'State']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Alumu-Tesu', 'Dialects', 'Alumu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Alumu-Tesu', 'Dialects', 'Tesu']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Alumu-Tesu', 'Language', 'codes', 'ISO', '639-3', 'aab']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

```
['Aranadan', 'Spoken', 'in', 'India']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aranadan', 'Region', 'Malappuram', 'District']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aranadan', 'Language', 'codes', 'ISO', '639-3', 'aaf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Southern', 'Spoken', 'in', 'Albania', '(', 'Durr', '\xc3', '\xab', 's',  
,',', 'Elbasan', ',,', 'Tiran', '\xc3', '\xab', ')', ',,', 'western',  
'Macedonia', 'between', 'Debar', 'and', 'Struga']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Albania', '\\(', 'Durr', '\xc3', '\xab', 's', ',,', 'Elbasan', ',,',  
'Tiran', '\xc3', '\xab', '\\)'], ',,', 'western', 'Macedonia', 'between',  
'Debar', 'and', 'Struga']]], 23, True)
```

```
['Southern', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]],  
3, True)
```

```
['Southern', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Southern', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Ambrak', 'Region', 'Nuku', 'District', ',,', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Nuku',  
'District', ',,', 'Sandaun', 'Province']]], 7, True)
```

```
['Ambrak', 'Language', 'codes', 'ISO', '639-3', 'aag']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aramanik', 'Spoken', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Aramanik', 'Region', 'Arusha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
```



ANY']]], 3, True)

['Aramanik', 'Language', 'codes', 'ISO', '639-3', 'aam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pamona', 'Spoken', 'in', 'Sulawesi', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Pamona', 'Language', 'codes', 'ISO', '639-3', 'pmf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Papantla', 'Spoken', 'in', 'Mexico', '(', 'Veracruz', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Papantla', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Papantla', 'Language', 'codes', 'ISO', '639-3', 'top']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mambila', 'Spoken', 'in', 'Nigeria', 'and', 'Cameroon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'and', 'Type: ANY']], 6, True)

['Mambila', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mambila', 'Language', 'codes', 'ISO', '639-3', 'mcu', '\\xe2', '\\x80',  
'\\x93', 'Camerounian', 'Mambila']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'mcu', '\\xe2', '\\x80', '\\x93', 'Camerounian', 'Type: ANY']],  
['NUM', '639-3']], 11, True)

['Cofi', 'dialect', 'Total', 'speakers', 'unknown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'dialect', 'Total',  
'speakers', 'unknown']], 5, True)

['Cofi', 'Spoken', 'in', 'Wales']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Cofi', 'Region', 'Gwynedd']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Cofi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Western', 'Spoken', 'in', 'Highlands', 'of', 'Irian', 'Jaya']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Highlands', 'of', 'Irian', 'Jaya']]], 7, True)

['Western', 'Language', 'codes', 'ISO', '639-3', 'dnw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Terei', 'Spoken', 'in', 'Bougainville']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Terei', 'Language', 'codes', 'ISO', '639-3', 'buo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Galela', 'Language', 'codes', 'ISO', '639-3', 'gbi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tehit', 'Spoken', 'in', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tehit', 'Language', 'codes', 'ISO', '639-3', 'kps']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amerbaken', 'Total', 'speakers', '7000']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Amerbaken', 'Dialects', 'Sirir']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Amberbaken', 'Dialects', 'Ajiw']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Amberbaken', 'Language', 'codes', 'ISO', '639-3', 'akc']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hattam', 'Dialects', 'Moi', '(', 'Moire', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\\\(', 'Type: ANY', '\\\\)']]], 6, True)

['Hattam', 'Dialects', 'Tinam']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hattam', 'Dialects', 'Miriei']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hattam', 'Dialects', 'Adihup']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hattam', 'Dialects', 'Uran']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Hattam', 'Language', 'codes', 'ISO', '639-3', 'had']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Abun', 'Dialects', 'Abun', 'Je']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Abun', 'Language', 'codes', 'ISO', '639-3', 'kgr']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['West', 'Spoken', 'in', 'Indonesia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['West', 'Language', 'codes', 'ISO', '639-3', 'mqs']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Morori', 'Spoken', 'in', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Morori', 'Language', 'codes', 'ISO', '639-3', 'mok']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Amung', 'Spoken', 'in', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Amung', 'Language', 'codes', 'ISO', '639-3', 'uhn']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dem', 'Spoken', 'in', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dem', 'Language', 'codes', 'ISO', '639-3', 'dem']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mor', 'Spoken', 'in', 'West', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Mor', 'Total', 'speakers', '25']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Mor', 'Language', 'codes', 'ISO', '639-3', 'moq']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kapori', 'Spoken', 'in', 'Papua']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kapori', 'Total', 'speakers', '200']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Kapori', 'Language', 'codes', 'ISO', '639-3', 'khp']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tanah', 'Spoken', 'in', 'West', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Tanah', 'Language', 'codes', 'ISO', '639-3', 'tcm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karas', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Karas', 'Language', 'codes', 'ISO', '639-3', 'kgv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Pawaia', 'Dialects', 'Aurama', '(', 'Turoha', ',', 'Uri', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Aurama', '\\\\(', 'Turoha', ',', 'Uri', '\\\\)']], 8, True)

['Pawaia', 'Dialects', 'Hauruha']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Pawaia', 'Language', 'codes', 'ISO', '639-3', 'pwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kamula', 'Language', 'codes', 'ISO', '639-3', 'xla']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guhu-Samane', 'Dialects', 'Sekare']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Guhu-Samane', 'Language', 'codes', 'ISO', '639-3', 'ghs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wersing', 'Dialects', 'Kolana']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

ANY']]], 3, True)

['Wersing', 'Dialects', 'Maneta']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wersing', 'Dialects', 'Langkuru', '(', 'Pureman', ',', 'Mademang', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Langkuru', '\\(', 'Pureman', ',', 'Mademang', '\\)']]], 8, True)

['Wersing', 'Region', 'Alor']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Wersing', 'Spoken', 'in', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lele', 'Total', 'speakers', '1300']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lele', 'Region', 'northeastern', 'Manus', 'Island', ',', 'Manus', 'Province']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'northeastern', 'Manus', 'Island', ',', 'Manus', 'Province']]], 8, True)

['Lele', 'Language', 'codes', 'ISO', '639-3', 'lle']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lele', 'Spoken', 'in', 'Guinea']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Lele', 'Region', 'Kissidougou', 'and', 'Gu', '\\xc3', '\\xa9', 'ck', '\\xc3', '\\xa9', 'dou', 'Prefectures']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Kissidougou', 'and', 'Gu', '\\xc3', '\\xa9', 'ck', '\\xc3', '\\xa9', 'dou', 'Prefectures']]], 12, True)

['Lele', 'Language', 'codes', 'ISO', '639-3', 'llc']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Faita', 'Spoken', 'in', 'Madang', 'Province']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Faita', 'Total', 'speakers', '50']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Faita', 'Language', 'codes', 'ISO', '639-3', 'faj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Paynamar', 'Spoken', 'in', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Paynamar', 'Language', 'codes', 'ISO', '639-3', 'pmr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Wadaginam', 'Spoken', 'in', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Wadaginam', 'Total', 'speakers', '950']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Wadaginam', 'Language', 'codes', 'ISO', '639-3', 'wdg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Osum', 'Spoken', 'in', 'Madang', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']]], 5, True)

['Osum', 'Total', 'speakers', '1200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Osum', 'Language', 'codes', 'ISO', '639-3', 'omo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Massep', 'Language', 'codes', 'ISO', '639-3', 'mvs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Molof', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Molof', 'Language', 'codes', 'ISO', '639-3', 'msl']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Samarokena', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Samarokena', 'Language', 'codes', 'ISO', '639-3', 'tmj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tofanma', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tofanma', 'Total', 'speakers', '250']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Tofanma', 'Language', 'codes', 'ISO', '639-3', 'tlg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Komyandaret', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Komyandaret', 'Total', 'speakers', '300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Komyandaret', 'Language', 'codes', 'ISO', '639-3', 'kzv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Manggarainese', 'Spoken', 'in', 'Manggarai', '(', 'Indonesia', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 7, True)

['Manggarainese', 'Writing', 'system', 'Latin', 'alphabet']:



([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system', 'Type: ANY', 'Type: ANY']]], 5, True)

['Manggarainese', 'Language', 'codes', 'ISO', '639-3', 'mgy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yoke', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', ',', 'Type: ANY']]], 6, True)

['Yoke', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yoke', 'Language', 'codes', 'ISO', '639-3', 'yki']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Hagahai', 'Language', 'codes', 'ISO', '639-3', 'pnn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ayapaneco', 'Total', 'speakers', '2']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ayapaneco', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Ayapaneco', 'Region', 'Jalpa', 'de', 'M', '\xc3', '\xa9', 'ndez', ',', 'Tabasco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Jalpa', 'de', 'M', '\xc3', '\xa9', 'ndez', ',', 'Tabasco']]], 10, True)

['Ayapaneco', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Yukulta', 'Language', 'extinction', 'extinct']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yukulta', 'Spoken', 'in', 'Queensland']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

'Type: ANY']]], 4, True)

['Yukulta', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yukulta', 'Language', 'codes', 'ISO', '639-3', ':', 'gcd', '\xe2',  
'\x80', '\x93', 'Ganggalida']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'gcd', '\xe2', '\x80', '\x93', 'Ganggalida']], ['NUM',  
'639-3']], 11, True)

['Yukulta', 'Language', 'codes', 'ISO', '639-3', 'nny', '\xe2', '\x80',  
'\x93', 'Nyannga']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'nny', '\xe2', '\x80', '\x93', 'Nyannga']], ['NUM',  
'639-3']], 10, True)

['Tor', 'Spoken', 'in', 'Mali']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Tor', 'Official', 'language', 'in', 'Mali']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Official', 'language',  
'in', 'Type: ANY']]], 5, True)

['Tor', 'Regulated', 'by', 'No', 'official', 'regulation']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Regulated', 'by', 'Type:  
ANY', 'of', 'Type: ANY']]], 6, True)

['Tor', 'Language', 'codes', 'ISO', '639-3', 'dts']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lomwe', 'Spoken', 'in', 'Mozambique']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lomwe', 'Language', 'codes', 'ISO', '639-3', 'ngl']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Demta', 'Spoken', 'in', 'Papua', ',', 'Indonesia']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Demta', 'Total', 'speakers', '1300']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Demta', 'Language', 'codes', 'ISO', '639-3', 'dmy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Meax', 'Language', 'codes', 'ISO', '639-3', 'mej']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Rasawa', 'Spoken', 'in', 'New', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Rasawa', 'Language', 'codes', 'ISO', '639-3', 'rac']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Saponi', 'Spoken', 'in', 'New', 'Guinea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Saponi', 'Language', 'codes', 'ISO', '639-3', 'spi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nakwi', 'Region', 'East', 'Sepik', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Nakwi', 'Language', 'codes', 'ISO', '639-3', 'nax']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Bo', 'Dialects', 'Kaboru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bo', 'Dialects', 'Nikiyama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Bo', 'Dialects', 'Umuruta']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:

ANY'[][]], 3, True)

['Bo', 'Region', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Bo', 'Language', 'codes', 'ISO', '639-3', 'bpw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Ama', 'Region', 'East', 'Sepik', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Ama', 'Language', 'codes', 'ISO', '639-3', 'amm']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Nimo', 'Region', 'Sandaun', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nimo', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']]], ['NUM', '639-2']], 8, True)

['Nimo', 'Language', 'codes', 'ISO', '639-3', 'niw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Owiniga', 'Region', 'East', 'Sepik', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 5, True)

['Owiniga', 'Language', 'codes', 'ISO', '639-3', 'owi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Fas', 'Language', 'codes', 'ISO', '639-3', 'fqs']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Baibai', 'Total', 'speakers', '340']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Baibai', 'Language', 'codes', 'ISO', '639-3', 'bbf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kwomtari', 'Language', 'codes', 'ISO', '639-3', 'kwo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Guriaso', 'Language', 'codes', 'ISO', '639-3', 'grx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kiwai', 'Dialects', 'Doumori']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kiwai', 'Dialects', 'Coast', 'Kiwai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kiwai', 'Dialects', 'Daru', 'Kiwai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kiwai', 'Dialects', 'Eastern', 'Kiwai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kiwai', 'Dialects', 'Island', 'Kiwai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Kiwai', 'Dialects', 'Gibaio']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kiwai', 'Dialects', 'Kope', '(', 'Gope', ',', 'Era', 'River', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Kope', '\\  
(', 'Gope', ',', 'Era', 'River', '\\)']], 9, True)
```

```
['Kiwai', 'Dialects', 'Urama']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Kiwai', 'Dialects', 'Arigibi', '(', 'Anigibi', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',
```

'\\(', 'Type: ANY', '\\)']]]], 6, True)

['Kiwai', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Kiwai', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kiwai', 'Language', 'codes', 'ISO', '639-3', ':', 'kiw', '\\xe2', '\\x80',  
'\\x93', 'Northeast', 'Kiwai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'kiw', '\\xe2', '\\x80', '\\x93', 'Northeast', 'Type:  
ANY']], ['NUM', '639-3']], 12, True)

['Abelam', 'Language', 'codes', 'ISO', '639-3', 'abt']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iatm\\xc3\\xbcl', 'Dialects', 'Nyaura']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Iatm\\xc3\\xbcl', 'Dialects', 'Palimbei']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Iatm\\xc3\\xbcl', 'Language', 'codes', 'ISO', '639-3', 'ian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Boiken', 'Language', 'codes', 'ISO', '639-3', 'bzf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Hawu', 'Dialects', 'Seba', '(', 'H', '\\xc9', '\\x99', '\\xc9', '\\x93', 'a',  
)']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Seba', '\\  
(', 'H', '\\xc9', '\\x99', '\\xc9', '\\x93', 'a', '\\)']], 11, True)

['Hawu', 'Dialects', 'Timu', '(', 'Dimu', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\(', 'Type: ANY', '\\)']], 6, True)

['Hawu', 'Dialects', 'Liae']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Hawu', 'Dialects', 'Mesara', '(', 'Mehara', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Hawu', 'Dialects', 'Raijua', '(', 'Raidjua', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\\(', 'Type: ANY', '\\)']], 6, True)

['Hawu', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']], 8, True)

['Hawu', 'Pronunciation', '\\xcb', '\\x88', 'havu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb', '\\x88', 'havu']], 5, True)

['Hawu', 'Language', 'codes', 'ISO', '639-2', '\\xe2', '\\x80', '\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\\xe2', '\\x80', '\\x93']], ['NUM', '639-2']], 8, True)

['Hawu', 'Language', 'codes', 'ISO', '639-3', 'hvn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Dhao', 'Pronunciation', '\\xcb', '\\x88', '\\xc9', '\\x96', '\\xcd', '\\xa1', '\\xca', '\\x90', 'ao']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Pronunciation', '\\xcb', '\\x88', '\\xc9', '\\x96', '\\xcd', '\\xa1', '\\xca', '\\x90', 'ao']], 11, True)

['Dhao', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Lesser', 'Sunda', 'Islands', ',', 'Indonesia']], 8, True)

['Dhao', 'Language', 'codes', 'ISO', '639-3', 'nfa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Havu', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'Congo']], 7, True)

['Havu', 'Region', 'Sud-Kivu', 'Province']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Havu', 'Language', 'codes', 'ISO', '639-3', 'hav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jino', 'Spoken', 'in', 'Jino', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Jino', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Jino', 'Language', 'codes', 'ISO', '639-3', ':', 'jiu', '\xe2', '\x80',  
\x93', 'Youle', 'Jinuo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'jiu', '\xe2', '\x80', '\x93', 'Youle', 'Jinuo']],  
['NUM', '639-3']], 12, True)

['Jino', 'Language', 'codes', 'ISO', '639-3', 'jiy', '\xe2', '\x80',  
\x93', 'Buyuan', 'Jinuo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'jiy', '\xe2', '\x80', '\x93', 'Buyuan', 'Jinuo']], ['NUM',  
'639-3']], 11, True)

['Tegem', 'Spoken', 'in', 'Nuba', 'Hills', ',', 'Sudan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Nuba',  
'Hills', ',', 'Sudan']]], 7, True)

['Tegem', 'Language', 'codes', 'ISO', '639-3', 'laf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Baa', 'Dialects', 'Gyakan']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Baa', 'Dialects', 'Kwa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Baa', 'Region', 'Adamawa', 'State']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)



```
['Baa', 'Spoken', 'in', 'Nigeria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Baa', 'Language', 'codes', 'ISO', '639-3', 'kwb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tyefo', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Tyefo', 'Language', 'codes', 'ISO', '639-3', 'tiq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Vyemo', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Vyemo', 'Language', 'codes', 'ISO', '639-3', 'vig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tusya', 'Spoken', 'in', 'Burkina', 'Faso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Tusya', 'Total', 'speakers', '20', ',', '000', '(', '?', ')', ',', '40',  
, ',', '000', '(', '?', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', ',', 'NUM', '\\(', '?', '\\)', ',', 'NUM', ',', 'NUM', '\\(', '?',  
'\\)'']], ['NUM', '20'], ['NUM', '000'], ['NUM', '40'], ['NUM', '000']], 16,  
True)
```

```
['Tusya', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Tusya', 'Language', 'codes', 'ISO', '639-3', ':', 'wib', '\\xe2', '\\x80',  
'\\x93', 'Southern']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', ':', 'wib', '\\xe2', '\\x80', '\\x93', 'Type: ANY']]], ['NUM',  
'639-3']], 11, True)
```

```
['Tusya', 'Language', 'codes', 'ISO', '639-3', 'tsp', '\\xe2', '\\x80',
```

```

'\x93', 'Northern']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'tsp', '\xe2', '\x80', '\x93', 'Type: ANY']], ['NUM',
'639-3']], 10, True)

['Betoi', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Betoi', 'Dialects', 'Situfa']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Betoi', 'Dialects', 'Airico']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Betoi', 'Dialects', 'Lolaca']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Betoi', 'Dialects', 'Jirara']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Betoi', 'Dialects', 'Betoi']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Betoi', 'Dialects', 'Ele']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']]], 3, True)

['Betoi', 'Region', 'Orinoco', 'llanos']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Betoi', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',
'Linguist', 'List']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',
'639-3']], 10, True)

['Betoi', 'Language', 'codes', 'ISO', '639-3', 'qtd']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

['Huamo\xc3\xa9', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Huamo\xc3\xa9', 'Region', 'Floresta', ',', 'Pernambuco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Huamo\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'uam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Katembr\xc3\xad', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Katembr\xc3\xad', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Katembr\xc3\xad', 'Region', 'to', 'Guyanese', 'border']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Katembr\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'qef', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qef', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Namla', 'Spoken', 'in', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Namla', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)

['Namla', 'Language', 'codes', 'ISO', '639-3', 'naa']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Esmeralda', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Esmeralda', 'Total', 'speakers', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'speakers',  
'\xe2', '\x80', '\x93']]], 6, True)
```

```
['Esmeralda', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Esmeralda', 'Language', 'codes', 'ISO', '639-3', '0ip']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Maratino', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Maratino', 'Spoken', 'in', 'NE', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Maratino', 'Language', 'codes', 'ISO', '639-3', 'q06', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'q06', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)
```

```
['Naolan', 'Spoken', 'in', 'NE', 'Mexico']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)
```

```
['Naolan', 'Language', 'codes', 'ISO', '639-3', 'ql8', '\xe2', '\x80',  
'\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'ql8', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)
```

```
['Tux\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Tux\xc3\xa1', 'Region', 'Bahia', ',', 'Pernambuco']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Tux\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'tud']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Taruma', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Taruma', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Taruma', 'Language', 'codes', 'ISO', '639-3', 'qoi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Matanaw\xc3\xad', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Matanaw\xc3\xad', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Matanaw\xc3\xad', 'Region', 'Amazonas']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Matanaw\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'qfu', '\xe2',  
'\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qfu', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Huarpe', 'Spoken', 'in', 'Argentina', ',', 'dispossessed', 'to',  
'Chile']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Argentina', ',', 'dispossessed', 'to', 'Chile']]], 8, True)

['Huarpe', 'Dialects', 'Millcayac']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Huarpe', 'Dialects', 'Allentiac']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Huarpe', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Huarpe', 'Language', 'codes', 'ISO', '639-3', 'qbt', '(', 'Alyentiyak', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qbt', '\\(', 'Alyentiyak', '\\)']], ['NUM', '639-3']], 9, True)

['Huarpe', 'Language', 'codes', 'ISO', '639-3', 'qgb', '(', 'Milykayak', ')']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qgb', '\\(', 'Milykayak', '\\)']], ['NUM', '639-3']], 9, True)

['Chimane', 'Dialects', 'Covendo', 'Moset', '\xc3', '\xa9', 'n']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', 'Moset', '\xc3', '\xa9', 'n']], 7, True)

['Chimane', 'Region', 'western', 'Amazon']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chimane', 'Spoken', 'in', 'Bolivia']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chimane', 'Language', 'codes', 'ISO', '639-3', 'cas']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Eborna', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Democratic', 'Republic', 'of', 'the', 'Congo']], 8, True)

['Eborna', 'Total', 'speakers']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']], 3, True)

['Eborna', 'Language', 'codes', 'ISO', '639-3', 'bxx']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```
['Carabayo', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Carabayo', 'Total', 'speakers', '150', '(', 'aerial', 'survey', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', '\\(', 'aerial', 'survey', '\\)']]], ['NUM', '150']], 8, True)
```

```
['Carabayo', 'Language', 'codes', 'ISO', '639-3', 'cby']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Guaycura', 'Spoken', 'in', 'Baja', 'California']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)
```

```
['Guaycura', 'Language', 'codes', 'ISO', '639-3', 'qjg', '\\xe2', '\\x80',  
'\\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qjg', '\\xe2', '\\x80', '\\x93']]], ['NUM', '639-3']], 9, True)
```

```
['Guaycura', 'Language', 'codes', 'ISO', '639-3', 'which', 'is', 'still',  
'spoken']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'which', 'is', 'still', 'spoken']]], ['NUM', '639-3']], 9,  
True)
```

```
['Guaycura', 'Language', 'codes', 'ISO', '639-3', 'is', 'a', 'Yuman',  
'language']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'is', 'a', 'Yuman', 'language']]], ['NUM', '639-3']], 9, True)
```

```
['Pankarar\\xc3\\xba', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Pankarar\\xc3\\xba', 'Region', 'Pernambuco', ',', 'Alagoas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Pankarar\\xc3\\xba', 'Language', 'codes', 'ISO', '639-3', 'either']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Pankarar\\xc3\\xba', 'Language', 'codes', 'ISO', '639-3', ':', 'paz',  
'\\xe2', '\\x80', '\\x93', 'Pankarar', '\\xc3', '\\xba']:
```

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', ':', 'paz', '\xe2', '\x80', '\x93', 'Pankarar', '\xc3', '\xba']], ['NUM', '639-3']], 13, True)

['Pankarar\xc3\xba', 'Language', 'codes', 'ISO', '639-3', 'pax', '\xe2', '\x80', '\x93', 'Pankarar', '\xc3', '\xa9']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'pax', '\xe2', '\x80', '\x93', 'Pankarar', '\xc3', '\xa9']], ['NUM', '639-3']], 12, True)

['Dusner', 'Total', 'speakers', '3', 'speakers']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers', 'NUM', 'speakers']], ['NUM', '3']], 5, True)

['Dusner', 'Spoken', 'in', 'Indonesia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dusner', 'Region', 'Southeast', 'Asia']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Dusner', 'Language', 'codes', 'ISO', '639-3', 'dsn']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Alag\xc3\xbcilac', 'Language', 'extinction', '?']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Alag\xc3\xbcilac', 'Spoken', 'in', 'Guatemala']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Alag\xc3\xbcilac', 'Region', 'Motagua', 'River', 'Ethnicity']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Alag\xc3\xbcilac', 'Region', 'Alaguilac', 'people']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Favorlang', 'Spoken', 'in', 'Taiwan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)



```
['Favorlang', 'Total', 'speakers', '0']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Favorlang', 'Language', 'codes', 'ISO', '639-3', 'bzg']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Chol\xc3\xb3n', 'Spoken', 'in', 'Huallaga', 'River', 'valley']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Huallaga', 'River', 'valley']]], 6, True)
```

```
['Chol\xc3\xb3n', 'Language', 'codes', 'ISO', '639-3', 'cht']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Hibito', 'Spoken', 'in', 'Bobonaje', 'River', 'valley']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',
'Bobonaje', 'River', 'valley']]], 6, True)
```

```
['Hibito', 'Language', 'codes', 'ISO', '639-3', 'hib']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Himarim\xc3\xa3', 'Total', 'speakers', '40']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Himarim\xc3\xa3', 'Spoken', 'in', 'Brazil']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)
```

```
['Himarim\xc3\xa3', 'Region', 'Tapau', '\xc3', '\xa1', 'River', 'valley',
'Ethnicity']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Tapau',
'\xc3', '\xa1', 'River', 'valley', 'Ethnicity']]], 8, True)
```

```
['Himarim\xc3\xa3', 'Region', 'Hi-Merim', '\xc3', '\xa3', 'people']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Hi-Merim',
'\xc3', '\xa3', 'people']]], 6, True)
```

```
['Himarim\xc3\xa3', 'Language', 'codes', 'ISO', '639-3', 'hir']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Yar\xc3\xad', 'Total', 'speakers', '760']:
```

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Yar\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'yri']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nat\xc3\xba', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Nat\xc3\xba', 'Region', 'Pernambuco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tarairi\xc3\xba', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tarairi\xc3\xba', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Tarairi\xc3\xba', 'Region', 'Pernambuco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Tarairi\xc3\xba', 'Region', 'Para', '\xc3', '\xad', 'ba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Para', '\xc3', '\xad', 'ba']]], 6, True)

['Tarairi\xc3\xba', 'Region', 'Rio', 'Grande', 'do', 'Norte']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Rio', 'Grande', 'do', 'Norte']]], 6, True)

['Jingulu', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

['Jingulu', 'Language', 'codes', 'ISO', '639-3', 'jig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Nung', 'Spoken', 'in', 'Vietnam']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```

['Nung', 'Language', 'codes', 'ISO', '639-3', 'nut']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guamo', 'Language', 'extinction', '?']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Guamo', 'Spoken', 'in', 'Venezuela']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Westrobothnian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',
'\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Panzaleo', 'Spoken', 'in', 'Ecuador']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Panzaleo', 'Region', 'Quito', 'Ethnicity']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Panzaleo', 'Region', 'Panzaleo', 'people']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Panzaleo', 'Language', 'codes', 'ISO', '639-3', 'None']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ngarnka', 'Spoken', 'in', 'Australia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

['Ngarnka', 'Language', 'codes', 'ISO', '639-3', 'nji']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kambiw\xc3\xa1', 'Spoken', 'in', 'Pernambuco']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']]], 4, True)

```

['Kambiw\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'xbw']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Biatah', 'Spoken', 'in', 'Malaysia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Biatah', 'Region', 'Borneo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Biatah', 'Language', 'codes', 'ISO', '639-2', 'bth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-2']], 6, True)

['Biatah', 'Language', 'codes', 'ISO', '639-3', 'bth']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaimb\xc3\xa9', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kaimb\xc3\xa9', 'Region', 'Bah', '\xc3', '\xad', 'a']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Bah', '\xc3',  
'\xad', 'a']], 6, True)

['Kaimb\xc3\xa9', 'Language', 'codes', 'ISO', '639-3', 'xai']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wakon\xc3\xa1', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Wakon\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Wakon\xc3\xa1', 'Region', 'Alagoas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

['Wakon\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'waf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kapinaw\xc3\xa2', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kapinaw\xc3\xa2', 'Spoken', 'in', 'Pernambuco']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kapinaw\xc3\xa2', 'Language', 'codes', 'ISO', '639-3', 'xpn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Siona', 'Total', 'speakers', '600']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Siona', 'Spoken', 'in', 'Colombia', ',', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Siona', 'Region', 'Putumayo', 'River', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)

['Siona', 'Region', 'Siona', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Siona', 'Language', 'codes', 'ISO', '639-3', 'snn']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ar\xc3\xa1ra', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ar\xc3\xa1ra', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Ar\xc3\xa1ra', 'Region', 'Mato', 'Grosso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

```
['Ar\xc3\xa1ra', 'Language', 'codes', 'ISO', '639-3', 'axg']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wasu', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Wasu', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Wasu', 'Region', 'Alagoas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Wasu', 'Language', 'codes', 'ISO', '639-3', 'wsu']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Truk\xc3\xa1', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Truk\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Truk\xc3\xa1', 'Region', 'Pernambuco', ',', 'Bahia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Truk\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'tka']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Cagua', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Cagua', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Cagua', 'Language', 'codes', 'ISO', '639-3', 'cbh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Chipiajes', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chipiajes', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Chipiajes', 'Language', 'codes', 'ISO', '639-3', 'cbe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Coxima', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Coxima', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Coxima', 'Language', 'codes', 'ISO', '639-3', 'kox']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Natagaimas', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Natagaimas', 'Region', 'Tolima']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Natagaimas', 'Language', 'codes', 'ISO', '639-3', 'nts']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Iapama', 'Total', 'speakers', '200']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Iapama', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

```
['Iapama', 'Region', 'Par', '\xc3', '\xa1', '\xe2', '\x80', '\x93', 'Amap',  
'\xc3', '\xa1', 'border']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Par', '\xc3',  
'\xa1', '\xe2', '\x80', '\x93', 'Amap', '\xc3', '\xa1', 'border']]], 12,  
True)
```

```
['Iapama', 'Language', 'codes', 'ISO', '639-3', 'iap']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Kamba', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kamba', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Kamba', 'Region', 'Mato', 'Grosso', 'do', 'Sul']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mato',  
'Grosso', 'do', 'Sul']]], 6, True)
```

```
['Kamba', 'Language', 'codes', 'ISO', '639-2', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-2']], 8, True)
```

```
['Kamba', 'Language', 'codes', 'ISO', '639-3', 'xba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Shenenawa', 'Spoken', 'in', 'Brazil']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Shenenawa', 'Region', 'Acre']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Shenenawa', 'Language', 'codes', 'ISO', '639-3', 'swo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tapeba', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```



```
['Tapeba', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tapeba', 'Region', 'Cear', '\xc3', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Tapeba', 'Language', 'codes', 'ISO', '639-3', 'tbb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Tingui-Boto', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tingui-Boto', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Tingui-Boto', 'Region', 'Alagoas']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Tingui-Boto', 'Language', 'codes', 'ISO', '639-3', 'tgv']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Trememb\xc3\xa1', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Trememb\xc3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Trememb\xc3\xa1', 'Region', 'Cear', '\xc3', '\xa1']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Trememb\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'tme']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aranama', 'Spoken', 'in', 'Texas', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:
```

ANY', 'Type: ANY']]], 5, True)

['Aranama', 'Spoken', 'in', 'Aranama', ',', 'Tamique']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Aranama', 'Language', 'codes', 'ISO', '639-3', 'xrt']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Mawa', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mawa', 'Spoken', 'in', 'Nigeria']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mawa', 'Region', 'Bauchi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Mawa', 'Language', 'codes', 'ISO', '639-3', 'wma']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Lepki', 'Region', 'Jayawijaya', 'Kabupaten']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lepki', 'Region', 'Okbibab', 'Kecamatan']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Lepki', 'Region', 'on', 'the', 'Sogber', 'River']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'on', 'the',  
'Sogber', 'River']]], 6, True)

['Lepki', 'Language', 'codes', 'ISO', '639-3', 'lpe']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Murkim', 'Region', 'Pegunungan', 'Bintang', 'Regency', ',', 'Batom',  
'subdistrict']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Pegunungan',  
'Bintang', 'Regency', ',', 'Batom', 'subdistrict']]], 8, True)

['Murkim', 'Language', 'codes', 'ISO', '639-3', 'rmh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Doso', 'Language', 'codes', 'ISO', '639-3', 'dol']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kehu', 'Spoken', 'in', 'West', 'Papua']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Kehu', 'Region', 'Wapoga', 'River', ',', 'in', 'the', 'foothills',  
'inland', 'from', 'Cenderawasih', 'Bay']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Wapoga',  
'River', ',', 'in', 'the', 'foothills', 'inland', 'from', 'Cenderawasih',  
'Bay']], 12, True)

['Kehu', 'Language', 'codes', 'ISO', '639-3', 'khh']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Sorothaptic', 'Spoken', 'in', 'Urnfield', 'culture']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Sorothaptic', 'Language', 'codes', 'ISO', '639-3', 'sxo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kaskian', 'Spoken', 'in', 'northeastern', 'Anatolia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'northeastern', 'Anatolia', 'Ethnicity']], 6, True)

['Kaskian', 'Spoken', 'in', 'Kaskians']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Kaskian', 'Language', 'codes', 'ISO', '639-3', 'zsk']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Buol', 'Spoken', 'in', 'Indonesia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Buol', 'Region', 'Central', 'Sulawesi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Buol', 'Language', 'codes', 'ISO', '639-3', 'blf']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paleo-Sardinian', 'Spoken', 'in', 'Corsica', 'et', 'Sardinia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Corsica', 'et', 'Sardinia', 'Ethnicity']], 7, True)

['Paleo-Sardinian', 'Spoken', 'in', 'Nuragic', 'civilization']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Paleo-Sardinian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Paleo-Sardinian', 'Language', 'codes', 'ISO', '639-3', 'Nuragic', 'tribes']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Nuragic', 'tribes']], ['NUM', '639-3']], 7, True)

['Ta\x3\xadno', 'Spoken', 'in', 'Caribbean', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Ta\x3\xadno', 'Spoken', 'in', 'Ta', '\xc3', '\xad', 'no', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Ta', '\xc3', '\xad', 'no', 'people']], 8, True)

['Ta\x3\xadno', 'Dialects', 'Baicawa', '(', 'Hispaniola', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY', '\(', 'Type: ANY', '\)']], 6, True)

['Ta\x3\xadno', 'Dialects', 'Cayaba', '(', 'Haiti', 'and', 'Florida', 'Keys', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Cayaba', '\(', 'Haiti', 'and', 'Florida', 'Keys', '\)']], 9, True)

['Ta\x3\xadno', 'Dialects', 'Cubaba', '(', 'Cuba', 'and', 'Hispaniola',

[')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Cubaba', '\\  
(', 'Cuba', 'and', 'Hispaniola', '\\)']]], 8, True)

['Ta\\xc3\\xadno', 'Dialects', 'Eyeri', '(', 'Puerto', 'Rico', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Eyeri', '\\  
(', 'Puerto', 'Rico', '\\)']]], 7, True)

['Ta\\xc3\\xadno', 'Dialects', 'Lucayo', '(', 'Bahamas', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\((', 'Type: ANY', '\\)')]]], 6, True)

['Ta\\xc3\\xadno', 'Language', 'codes', 'ISO', '639-3', 'tnq']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Warndarang', 'Language', 'extinction', '1974']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Warndarang', 'Spoken', 'in', 'Arnhem', 'Land']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Warndarang', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Warndarang', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Warndarang', 'Language', 'codes', 'ISO', '639-3', 'wnd']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mara', 'Spoken', 'in', 'Arnhem', 'Land']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Mara', 'Spoken', 'in', 'Northern', 'Territory']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Mara', 'Spoken', 'in', 'Australia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',

'Type: ANY']]], 4, True)

['Mara', 'Language', 'codes', 'ISO', '639-3', 'mec']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Palta', 'Spoken', 'in', 'Ecuador']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palta', 'Region', 'Amazonia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palta', 'Region', 'Palta', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Palta', 'Language', 'codes', 'ISO', '639-3', 'None']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Kukur\x3\xa1', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kukur\x3\xa1', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Kukur\x3\xa1', 'Region', 'Mato', 'Grosso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chono', 'Language', 'extinction', '1875']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chono', 'Spoken', 'in', 'Chile']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Chono', 'Region', 'Chilo', '\xc3', '\xa9', 'Archipelago', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Chilo',  
'\xc3', '\xa9', 'Archipelago', 'Ethnicity']]], 7, True)

```
['Chono', 'Region', 'Chono', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Chono', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Chono', 'Language', 'codes', 'ISO', '639-3', '064']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Jeik\xc3\xb3', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Umot\xc3\xadna', 'Language', 'extinction', '1988']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Umot\xc3\xadna', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Umot\xc3\xadna', 'Region', 'Mato', 'Grosso']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Umot\xc3\xadna', 'Language', 'codes', 'ISO', '639-3', 'umo']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Gamela', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gamela', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Gamela', 'Region', 'Maranh', '\xc3', '\xa3', 'o']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Maranh',  
'\xc3', '\xa3', 'o']]], 6, True)
```

['Gamela', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Gamela', 'Language', 'codes', 'ISO', '639-3', '1hz']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Miriti', 'Spoken', 'in', 'Brazil']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Miriti', 'Total', 'speakers', '0']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Miriti', 'Language', 'codes', 'ISO', '639-3', 'mmv']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Tawasa', 'Spoken', 'in', 'United', 'States']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']], 5, True)

['Tawasa', 'Region', 'eastern', 'Alabama', 'Ethnicity']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Tawasa', 'Region', 'Tawasa', 'people']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Berik', 'Total', 'speakers', '1200']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Berik', 'Writing', 'system', 'Latin']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Berik', 'Language', 'codes', 'ISO', '639-3', 'bkl']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)



['Cabiyari', 'Spoken', 'in', 'Colombia', ',', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Cabiyari', 'Language', 'codes', 'ISO', '639-3', 'cbb']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)

['Peba', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Peba', 'Spoken', 'in', 'western', 'Amazon']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Peba', 'Language', 'codes', 'ISO', '639-3', 'q14', '\xe2', '\x80',  
'\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'q14', '\xe2', '\x80', '\x93']]], ['NUM', '639-3']], 9, True)

['Timote', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Timote', 'Total', 'speakers', 'perhaps', '200', 'in', 'the', 'village',  
'of', 'Mut', '\xc3', '\xba', 's']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'perhaps', 'NUM', 'in', 'the', 'village', 'of', 'Mut', '\xc3', '\xba',  
's']]], ['NUM', '200']], 13, True)

['Timote', 'Dialects', 'Timote']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Timote', 'Dialects', 'Cuica']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Timote', 'Dialects', '?', 'Mut', '\xc3', '\xba', '(', 'Loco', ')']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', '?', 'Mut',  
'\xc3', '\xba', '\\(', 'Loco', '\\)']]], 9, True)

['Timote', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',

```
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Timote', 'Language', 'codes', 'ISO', '639-3', 'qfn', '(', 'Cuica', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qfn', '\\(', 'Cuica', '\\)']], ['NUM', '639-3']], 9, True)
```

```
['Jebero', 'Spoken', 'in', 'Per', '\xc3', '\xba']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Per', '\xc3', '\xba']], 6, True)
```

```
['Jebero', 'Language', 'codes', 'ISO', '639-3', 'jeb']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taparita', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Taparita', 'Spoken', 'in', 'Venezuela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Taparita', 'Region', 'Amazon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Taparita', 'Language', 'codes', 'ISO', '639-3', 'qlf', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'qlf', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)
```

```
['Otomako', 'Language', 'extinction', '?']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Otomako', 'Spoken', 'in', 'Venezuela']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Otomako', 'Region', 'Amazon']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)
```

```
['Otomako', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:
```

```
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)
```

```
['Otomako', 'Language', 'codes', 'ISO', '639-3', 'apg']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Sanapan\xc3\xa1', 'Spoken', 'in', 'Paraguay', 'Ethnicity']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Sanapan\xc3\xa1', 'Spoken', 'in', 'Sanapan', '\xc3', '\xa1', 'people']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Sanapan', '\xc3', '\xa1', 'people']], 7, True)
```

```
['Sanapan\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', 'sap']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Enxet', 'Spoken', 'in', 'Paraguay', 'Ethnicity']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Enxet', 'Spoken', 'in', 'Enxet', 'people']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Enxet', 'Language', 'codes', 'ISO', '639-3', 'leg']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Aweti', 'Spoken', 'in', 'Brazil', 'Ethnicity']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Aweti', 'Spoken', 'in', 'Aweti', 'people']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)
```

```
['Aweti', 'Total', 'speakers', '140']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)
```

```
['Aweti', 'Language', 'codes', 'ISO', '639-3', 'awe']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
```

'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mawe', 'Spoken', 'in', 'Brazil', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Mawe', 'Spoken', 'in', 'Mawe', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Mawe', 'Language', 'codes', 'ISO', '639-3', 'mav']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Karo', 'Spoken', 'in', 'Brazil']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karo', 'Total', 'speakers', '170']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Karo', 'Dialects', 'Arara']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Karo', 'Dialects', 'Uruku']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)

['Karo', 'Language', 'codes', 'ISO', '639-3', 'arr']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Molala', 'Total', 'speakers', '0']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Molala', 'Spoken', 'in', 'United', 'States']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', 'Type: ANY']]], 5, True)

['Molala', 'Region', 'Molala', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Molala', 'Language', 'codes', 'ISO', '639-3', 'mbe']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Cochim\xc3\xad', 'Spoken', 'in', 'Mexico']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Cochim\xc3\xad', 'Region', 'Baja', 'California', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Cochim\xc3\xad', 'Region', 'Cochim', '\xc3', '\xad', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Cochim', '\xc3', '\xad', 'people']], 6, True)

['Cochim\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'coj']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Gaya', 'Spoken', 'in', 'Gaya', 'confederacy']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type: ANY', 'Type: ANY']], 5, True)

['Gaya', 'Region', 'Manchuria']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Gaya', 'Language', 'codes', 'ISO', '639-3', 'zra']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Guambiano', 'Spoken', 'in', 'Colombia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Guambiano', 'Region', 'Cauca', 'Department', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 5, True)

['Guambiano', 'Language', 'codes', 'ISO', '639-3', 'gum']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mucuch\xc3\xad', 'Language', 'extinction', '?']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

'Type: ANY']]], 4, True)

['Mucuch\xc3\xad', 'Spoken', 'in', 'Venezuela']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)

['Mucuch\xc3\xad', 'Dialects', 'Mucuch', '\xc3', '\xad']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Mucuch',  
'\xc3', '\xad']]], 5, True)

['Mucuch\xc3\xad', 'Dialects', 'Mirrip', '\xc3', '\xba']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Mirrip',  
'\xc3', '\xba']]], 5, True)

['Mucuch\xc3\xad', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80',  
'\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)

['Mucuch\xc3\xad', 'Language', 'codes', 'ISO', '639-3', 'qcx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Mijikenda', 'Spoken', 'in', 'Kenya', ',', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']]], 6, True)

['Mijikenda', 'Region', 'Mombasa', 'and', 'Kwale', 'districts', 'in',  
'Kenya', ';', 'Muheza', 'and', 'Tanga', 'districts', 'in', 'Tanzania']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Mombasa',  
'and', 'Kwale', 'districts', 'in', 'Kenya', ';', 'Muheza', 'and', 'Tanga',  
'districts', 'in', 'Tanzania']]], 15, True)

['Mijikenda', 'Language', 'codes', 'ISO', '639-3', 'dig']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Lembena', 'Language', 'codes', 'ISO', '639-3', 'net']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Paunaka', 'Spoken', 'in', 'Chiquitan', '\xc3', '\xad', 'a', ',',  
'Bolivia']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in',  
'Chiquitan', '\xc3', '\xad', 'a', ',', 'Bolivia']]], 9, True)

['Paunaka', 'Total', 'speakers', '5', 'official', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers',  
'NUM', 'of', 'Type: ANY']], ['NUM', '5']], 6, True)

['Paunaka', 'Language', 'codes', 'ISO', '639-3', 'qbz', '\xe2', '\x80',  
'\x93', 'Paunaka-Paikone']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'qbz', '\xe2', '\x80', '\x93', 'Paunaka-Paikone']], ['NUM',  
'639-3']], 10, True)

['Misima-Paneati', 'Writing', 'system', 'Latin', 'alphabet']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Writing', 'system',  
'Type: ANY', 'Type: ANY']], 5, True)

['Misima-Paneati', 'Language', 'codes', 'ISO', '639-3', 'mpx']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Danezaa', 'Spoken', 'in', 'Canada']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Danezaa', 'Region', 'British', 'Columbia', 'Ethnicity']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']], 5, True)

['Danezaa', 'Region', 'Danezaa', 'people']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Danezaa', 'Language', 'codes', 'ISO', '639-3', 'bea']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ese\xe2\x80\x99ejja', 'Total', 'speakers', '1800']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)

['Ese\xe2\x80\x99ejja', 'Spoken', 'in', 'Bolivia', ',', 'Peru']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Spoken', 'in', 'Type:  
ANY', ',', 'Type: ANY']], 6, True)

['Ese\xe2\x80\x99ejja', 'Region', 'Beni']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)

```
['Ese\xe2\x80\x99ejja', 'Language', 'codes', 'ISO', '639-3', 'ese']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Taizhou', 'Wu', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Taizhou', 'Region', 'Zhejiang']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Taizhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Taizhou', 'Language', 'codes', 'ISO', '639-3', 'wu-taz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Wuzhou', 'Wu', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Wuzhou', 'Region', 'Zhejiang']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']], 3, True)
```

```
['Wuzhou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93',  
'Linguist', 'List']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM',  
'639-3']], 10, True)
```

```
['Wuzhou', 'Language', 'codes', 'ISO', '639-3', 'wu-wuz']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Xuanzhou', 'Wu', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']], 4, True)
```

```
['Xuanzhou', 'Region', 'Anhui']:
```



([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Xuanzhou', 'Language', 'codes', 'ISO', '639-3', 'wu-xua', '\xe2', '\x80', '\x93']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'wu-xua', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Old', 'Xiang', 'Total', 'speakers']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Old', 'Region', 'Hunan']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Old', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Old', 'Language', 'codes', 'ISO', '639-3', 'hsn-luo']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Ji-Xu', 'Xiang', 'Total', 'speakers']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Ji-Xu', 'Region', 'Hunan']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Ji-Xu', 'Language', 'codes', 'ISO', '639-3', 'hsn-jis', '\xe2', '\x80', '\x93']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'hsn-jis', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Yong-Xun', 'Yue', 'Total', 'speakers']:  
([['Start', 'Phrase']], ['Phrase', 'Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Yong-Xun', 'Region', 'Guangxi']:

([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Yong-Xun', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Yong-Xun', 'Language', 'codes', 'ISO', '639-3', 'Yong-Xun']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Yong-Xun', 'Language', 'codes', 'ISO', '639-3', 'to', 'the', 'left', 'in', 'brown']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'to', 'the', 'left', 'in', 'brown']], ['NUM', '639-3']], 10, True)

['Yong-Xun', 'Language', 'codes', 'ISO', '639-3', 'is', 'labeled', 'Nanning']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'is', 'labeled', 'Type: ANY']], ['NUM', '639-3']], 8, True)

['Qin-Lian', 'Yue', 'Total', 'speakers']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']], 4, True)

['Qin-Lian', 'Region', 'Guangxi']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']], 3, True)

['Qin-Lian', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93', 'Linguist', 'List']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', '\xe2', '\x80', '\x93', 'Linguist', 'List']], ['NUM', '639-3']], 10, True)

['Qin-Lian', 'Language', 'codes', 'ISO', '639-3', 'yue-qin']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Qin-Lian', 'Language', 'codes', 'ISO', '639-3', 'Qin-Lian']:  
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Qin-Lian', 'Language', 'codes', 'ISO', '639-3', 'to', 'the', 'left', 'in', 'green']:

([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'to', 'the', 'left', 'in', 'green']], ['NUM', '639-3']], 10, True)

['Qin-Lian', 'Language', 'codes', 'ISO', '639-3', 'is', 'labeled', 'Beihai']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'is', 'labeled', 'Type: ANY']], ['NUM', '639-3']], 8, True)

['Yuehai', 'Total', 'speakers']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Total', 'speakers']]], 3, True)

['Yuehai', 'Dialects', 'Guangfu']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yuehai', 'Dialects', 'Sanyi']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yuehai', 'Dialects', 'Zhongshan']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yuehai', 'Dialects', 'Guan-Bao']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yuehai', 'Region', 'Guangdong']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY']]], 3, True)

['Yuehai', 'Language', 'codes', 'ISO', '639-3', 'yue-yue', '\xe2', '\x80', '\x93']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'yue-yue', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 9, True)

['Yuehai', 'Language', 'codes', 'ISO', '639-3', 'at', 'right']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes', 'ISO', 'NUM', 'at', 'right']], ['NUM', '639-3']], 7, True)

['Goulou', 'Yue', 'Total', 'speakers']:  
([['Start', 'Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY', 'Type: ANY']]], 4, True)

```
['Goulou', 'Region', 'Guangxi', ',', 'Guangdong']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Region', 'Type: ANY',  
'Type: ANY', 'Type: ANY']]], 5, True)
```

```
['Goulou', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Goulou', 'Language', 'codes', 'ISO', '639-3', 'Goulou']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)
```

```
['Goulou', 'Language', 'codes', 'ISO', '639-3', 'to', 'the', 'left', 'in',  
'chartreuse']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'to', 'the', 'left', 'in', 'chartreuse']], ['NUM', '639-3']],  
10, True)
```

```
['Goulou', 'Language', 'codes', 'ISO', '639-3', 'is', 'labeled', 'Yulin']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'is', 'labeled', 'Type: ANY']], ['NUM', '639-3']], 8, True)
```

```
['Luo-Guang', 'Yue', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Luo-Guang', 'Region', 'Guangdong']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Luo-Guang', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```
['Wu-Hua', 'Yue', 'Total', 'speakers']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Wu-Hua', 'Region', 'Guangxi']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Wu-Hua', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)
```

```

['Wu-Hua', 'Language', 'codes', 'ISO', '639-3', 'Wu-Hua']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

['Wu-Hua', 'Language', 'codes', 'ISO', '639-3', 'to', 'the', 'left', 'in',
'pink']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'to', 'the', 'left', 'in', 'pink']], ['NUM', '639-3']], 10,
True)

['Wu-Hua', 'Language', 'codes', 'ISO', '639-3', 'is', 'labeled', 'Wuzhou']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'is', 'labeled', 'Type: ANY']], ['NUM', '639-3']], 8, True)

['Hainan', 'Yue', 'Total', 'speakers']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Hainan', 'Region', 'Hainan']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:
ANY']], 3, True)

['Hainan', 'Language', 'codes', 'ISO', '639-3', '\xe2', '\x80', '\x93']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', '\xe2', '\x80', '\x93']], ['NUM', '639-3']], 8, True)

['Tambora', 'Language', 'extinction', '1815']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tambora', 'Spoken', 'in', 'Indonesia']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tambora', 'Region', 'Sumbawa', 'Ethnicity']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tambora', 'Region', 'Tambora', 'culture']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',
'Type: ANY']], 4, True)

['Tambora', 'Language', 'codes', 'ISO', '639-3', 'xxt']:
([['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',
'ISO', 'NUM', 'Type: ANY']], ['NUM', '639-3']], 6, True)

```

```
['Bima', 'Spoken', 'in', 'Sumbawa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Bima', 'Dialects', 'Kolo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Bima', 'Dialects', 'Sangar', '(', 'Sanggar', ')']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Dialects', 'Type: ANY',  
'\\(', 'Type: ANY', '\\)']]], 6, True)
```

```
['Bima', 'Dialects', 'Toloweri']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Bima', 'Dialects', 'Bima']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Bima', 'Dialects', 'Mbojo']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type:  
ANY']]], 3, True)
```

```
['Bima', 'Language', 'codes', 'ISO', '639-3', 'bhp']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

```
['Sumbawa', 'Spoken', 'in', 'Sumbawa']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Type: ANY', 'Type: ANY',  
'Type: ANY']]], 4, True)
```

```
['Sumbawa', 'Language', 'codes', 'ISO', '639-3', 'smw']:  
([[['Start', ['Phrase']], ['Phrase', ['Type: ANY', 'Language', 'codes',  
'ISO', 'NUM', 'Type: ANY']]], ['NUM', '639-3']], 6, True)
```

## FAILED RULES

```
['Arabic', 'Spoken', 'in', 'Primarily', 'in', 'the', 'Arab', 'states',  
'of', 'the', 'Middle', 'East', 'and', 'North', 'Africa', ';']  
['Arabic', 'Standard', 'forms', 'Modern', 'Standard', 'Arabic']  
['Arabic', 'Official', 'language', 'in', 'Official', 'language', 'of',
```

'26', 'states', ',', 'the', 'third', 'most', 'after', 'English', 'and',  
'French']  
['Arabic', 'Regulated', 'by', 'Algeria', ':', 'Supreme', 'Council', 'of',  
'the', 'Arabic', 'language', 'in', 'Algeria']  
['Arabic', 'Regulated', 'by', 'Egypt', ':', 'Academy', 'of', 'the',  
'Arabic', 'Language', 'in', 'Cairo']  
['Arabic', 'Regulated', 'by', 'Iraq', ':', 'Iraqi', 'Academy', 'of',  
'Sciences']  
['Arabic', 'Regulated', 'by', 'Libya', ':', 'Academy', 'of', 'the',  
'Arabic', 'Language', 'in', 'Jamahiriya']  
['Arabic', 'Regulated', 'by', 'Morocco', ':', 'Academy', 'of', 'the',  
'Arabic', 'Language', 'in', 'Rabat']  
['Arabic', 'Regulated', 'by', 'Somalia', ':', 'Academy', 'of', 'the',  
'Arabic', 'Language', 'in', 'Mogadishu']  
['Arabic', 'Regulated', 'by', 'Southern', 'Sudan', '&', 'amp', ';']  
['Arabic', 'Regulated', 'by', 'Sudan', ':', 'Academy', 'of', 'the',  
'Arabic', 'Language', 'in', 'Khartoum']  
['Arabic', 'Regulated', 'by', 'Israel', ':', 'Academy', 'of', 'the',  
'Arabic', 'Language', 'in', 'Israel']  
['Arabic', 'Language', 'codes', 'ISO', '639-3', 'ara', '\xe2', '\x80',  
'\x93', 'Arabic', '(', 'generic', ')']  
['Arabic', 'Language', 'codes', 'ISO', '639-3', '(', 'see', 'varieties',  
'of', 'Arabic', 'for', 'the', 'individual', 'codes', ')']  
['Arabic', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
'Arabic', 'as', 'an', 'official', 'language', 'in', 'the', 'Arab', 'World']  
['Arabic', 'Language', 'codes', 'ISO', '639-3', 'Majority', 'Arabic',  
'speakers', '(', 'blue', ')', 'and', 'minority', 'Arabic', 'speakers', '(',  
'green', ')']  
['Austrian', 'Spoken', 'in', 'Italy', '(', 'South', 'Tyrol', ')']  
['Austrian', 'Total', 'speakers', '8', '.', '5', 'million']  
['American', 'Language', 'codes', 'ISO', '639-3', 'ase', 'sign',  
'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
'\xe2', '\x80', '\x94', 'legal', 'recognition']  
['Armenian', 'Total', 'speakers', '6', '.', '7', 'million']  
['Armenian', 'Language', 'codes', 'ISO', '639-3', ':', 'hye', '\xe2',  
'\x80', '\x93', 'Modern', 'Armenian']  
['Armenian', 'Language', 'codes', 'ISO', '639-3', 'xcl', '\xe2', '\x80',  
'\x93', 'Classical', 'Armenian']  
['Armenian', 'Language', 'codes', 'ISO', '639-3', 'axm', '\xe2', '\x80',  
'\x93', 'Middle', 'Armenian']  
['Armenian', 'Language', 'codes', 'Linguasphere', '57-AAA-a', '(', '31',  
'varieties', ')']  
['Aramaic', 'Total', 'speakers', '~', '500', ',', '000']  
['Aramaic', 'Region', 'Throughout', 'the', 'Middle', 'East', ',', 'Europe',  
'and', 'America']  
['Aramaic', 'Writing', 'system', 'Arabic', 'Abjad', '(', 'vernacular', '}',

'with', 'a', 'handful', 'of', 'inscriptions', 'found', 'in', 'Demotic',  
'and', 'Chinese']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', ':', 'arc', '\xe2',  
\x80', \x93', 'Imperial', 'and', 'Official', 'Aramaic', '(', '700-300',  
'BCE', ')']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'oar', '\xe2', '\x80',  
\x93', 'Old', 'Aramaic', '(', 'before', '700', 'BCE', ')']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'aii', '\xe2', '\x80',  
\x93', 'Assyrian', 'Neo-Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'aij', '\xe2', '\x80',  
\x93', 'Lishanid', 'Noshan']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'amw', '\xe2', '\x80',  
\x93', 'Western', 'Neo-Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'bhn', '\xe2', '\x80',  
\x93', 'Bohtan', 'Neo-Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'bjf', '\xe2', '\x80',  
\x93', 'Barzani', 'Jewish', 'Neo-Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'cld', '\xe2', '\x80',  
\x93', 'Chaldean', 'Neo-Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'jpa', '\xe2', '\x80',  
\x93', 'Jewish', 'Palestinian', 'Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'kqd', '\xe2', '\x80',  
\x93', 'Koy', 'Sanjaq', 'Surat']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'lsd', '\xe2', '\x80',  
\x93', 'Lishana', 'Deni']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'mid', '\xe2', '\x80',  
\x93', 'Modern', 'Mandaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'myz', '\xe2', '\x80',  
\x93', 'Classical', 'Mandaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'sam', '\xe2', '\x80',  
\x93', 'Samaritan', 'Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'syc', '\xe2', '\x80',  
\x93', 'Syriac', '(', 'classical', ')']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'tmr', '\xe2', '\x80',  
\x93', 'Jewish', 'Babylonian', 'Aramaic']

['Aragonese', 'Total', 'speakers', '10', ',', '000', '(', '30', ',', '000',  
'total', ')']

['Aragonese', 'Language', 'codes', 'Linguasphere', 'Language',  
'distribution', 'in', 'Aragon', '(', 'Aragonese', 'in', 'red', ')', '.',  
'Spanish', 'is', 'spoken', 'across', 'the', 'whole', 'area', ',', 'but',  
'the', 'yellow-green', 'part', 'of', 'Aragon', 'is', 'monolingually',  
'Spanish-speaking']

['Bulgarian', 'Regulated', 'by', 'Institute', 'for', 'the', 'Bulgarian',  
'language', 'at', 'the', 'Bulgarian', 'Academy', 'of', 'Sciences', '(',  
\xd0', \x98', \xd0', \xbd', \xd1', \x81', \xd1', \x82', \xd0',  
\xb8', \xd1', \x82', \xd1', \x83', \xd1', \x82', \xd0', \xb7',



'\xd0', '\xb0', '\xd0', '\xb1', '\xd1', '\x8a', '\xd0', '\xbb', '\xd0', '\xb3', '\xd0', '\xb0', '\xd1', '\x80', '\xd1', '\x81', '\xd0', '\xba', '\xd0', '\xb8', '\xd0', '\xb5', '\xd0', '\xb7', '\xd0', '\xb8', '\xd0', '\xba', '\xd0', '\xba', '\xd1', '\x8a', '\xd0', '\xbc', '\xd0', '\x91', '\xd1', '\x8a', '\xd0', '\xbb', '\xd0', '\xb3', '\xd0', '\xb0', '\xd1', '\x80', '\xd1', '\x81', '\xd0', '\xba', '\xd0', '\xb0', '\xd1', '\x82', '\xd0', '\xb0', '\xd0', '\xb0', '\xd0', '\xba', '\xd0', '\xb0', '\xd0', '\xb4', '\xd0', '\xb5', '\xd0', '\xbc', '\xd0', '\xb8', '\xd1', '\x8f', '\xd0', '\xbd', '\xd0', '\xb0', '\xd0', '\xbd', '\xd0', '\xb0', '\xd1', '\x83', '\xd0', '\xba', '\xd0', '\xb8', '\xd1', '\x82', '\xd0', '\xb5', '(', '\xd0', '\x91', '\xd0', '\x90', '\xd0', '\x9d', ')', ')']

['Bulgarian', 'Language', 'codes', 'Linguasphere', '53-AAA-hb', '&', 'lt', ';', '53-AAA-h']

['Bambara', 'Total', 'speakers', '4', ',', '000', ',', '000', ',', 'second', 'language', 'speakers', '8-10', 'millions', '(', 'estimation', ',', '2010', ';', 'the', 'figures', 'by', 'the', 'Ethnologue', 'of', '2', ',', '700', ',', '000', 'are', 'out', 'of', 'date', ')']

['Bambara', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]

['Breton', 'Total', 'speakers', 'ca', '.', '200', ',', '000']

['Breton', 'Language', 'codes', 'Linguasphere', '50-ABB-b', '(', 'varieties']

['Breton', 'Language', 'codes', 'Linguasphere', ':', '50-ABB-ba', 'to', '-be', ')']

['Blissymbols', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']

['Brahui', 'Total', 'speakers', '2', '.', '2', 'million', '(', '2005', ')']

['Brahui', 'Regulated', 'by', 'Brahui', 'Language', 'Board', '(', 'Pakistan', ')']

['Brahui', 'Language', 'codes', 'ISO', '639-3', 'Brahui', '(', 'at', 'far', 'upper-left', ')', 'is', 'quite', 'geographically', 'isolated', 'from', 'the', 'other', 'Dravidian', 'languages']

['Khmer', 'Total', 'speakers', '15', '.', '7', 'to', '21', '.', '6', 'million', '(', '2004', ')']

['Khmer', 'Total', 'speakers', 'Native', 'speakers', ':', '14', '.', '7', 'to', '20', '.', '6', 'million']

['Khmer', 'Total', 'speakers', 'Cambodia', ':', '12', '.', '1', 'million']

['Khmer', 'Total', 'speakers', 'Vietnam', ':', '1', ',', '055', ',', '174']

['Khmer', 'Total', 'speakers', 'Thailand', ':', '1', '.', '2', 'million']

['Khmer', 'Total', 'speakers', 'USA', ':', '190', ',', '000']

['Khmer', 'Total', 'speakers', 'France', ':', 'ca', '.', '50', ',', '000']

['Khmer', 'Total', 'speakers', 'Australia', ':', '22', ',', '000']

['Khmer', 'Total', 'speakers', 'Canada', ':', '16', ',', '500', '2nd', 'language', 'speakers', ':', '1', 'million', 'in', 'Cambodia']

['Khmer', 'Language', 'codes', 'ISO', '639-3', ':', 'khm', '\xe2', '\x80', '\x93', 'Central', 'Khmer']

['Khmer', 'Language', 'codes', 'ISO', '639-3', 'kxm', '\xe2', '\x80',

'\x93', 'Northern', 'Khmer']  
 ['Khmer', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Catalan', 'Valencian', 'Total', 'speakers', '11', '.', '5', 'million',  
 '(', '2006', ')']  
 ['Catalan', 'Spoken', 'in', 'Northern', 'Catalonia', '(', 'Pyr', '\xc3',  
 '\xa9', 'n', '\xc3', '\xa9', 'es-Orientales', ')']  
 ['Catalan', 'Regulated', 'by', 'Institut', 'd'Estudis', 'Catalans']  
 ['Chinese', 'Total', 'speakers', 'approx', '1', '.', '3', 'billion']  
 ['Chinese', 'Dialects', 'Eastern', 'Min', '(', 'including', 'Fuchow', ')']  
 ['Chinese', 'Dialects', 'Southern', 'Min', '(', 'including', 'Amoy', ',',  
 'Taiwanese', ')']  
 ['Chinese', 'Region', '(', 'minorities', ')', ':', 'Southeast', 'Asia',  
 ',', 'and', 'other', 'regions', 'with', 'Chinese', 'communities']  
 ['Chinese', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China', '(',  
 'PRC', ',', 'commonly', 'known', 'as', 'China', ')', ',', 'Republic', 'of',  
 'China', '(', 'ROC', ',', 'commonly', 'known', 'as', 'Taiwan', ')', ',',  
 'Singapore', ',', 'Malaysia', ',', 'the', 'United', 'States', ',', 'the',  
 'Philippines', ',', 'Australia', ',', 'Indonesia', ',', 'Mauritius', ',',  
 'Peru', ',', 'and', 'other', 'regions', 'with', 'Chinese', 'communities']  
 ['Chinese', 'Official', 'language', 'in', 'People's', 'Republic', 'of',  
 'China']  
 ['Chinese', 'Official', 'language', 'in', 'Singapore', '(', 'official',  
 ',', 'but', 'not', 'national', 'language', ')']  
 ['Chinese', 'Regulated', 'by', 'In', 'the', 'PRC', ':', 'National',  
 'Language', 'Regulating', 'Committee']  
 ['Chinese', 'Regulated', 'by', 'In', 'Singapore', ':', 'Promote',  
 'Mandarin', 'Council', '/', 'Speak', 'Mandarin', 'Campaign']  
 ['Chinese', 'Language', 'codes', 'ISO', '639-3', ':', 'cdo', '\xe2',  
 '\x80', '\x93', 'Min', 'Dong']  
 ['Chinese', 'Language', 'codes', 'ISO', '639-3', 'czo', '\xe2', '\x80',  
 '\x93', 'Min', 'Zhong']  
 ['Chinese', 'Language', 'codes', 'ISO', '639-3', 'mnp', '\xe2', '\x80',  
 '\x93', 'Min', 'Bei']  
 ['Chinese', 'Language', 'codes', 'ISO', '639-3', 'nan', '\xe2', '\x80',  
 '\x93', 'Min', 'Nan']  
 ['Chinese', 'Language', 'codes', 'ISO', '639-3', 'och', '\xe2', '\x80',  
 '\x93', 'Old', 'Chinese']  
 ['Chinese', 'Language', 'codes', 'ISO', '639-3', 'ltc', '\xe2', '\x80',  
 '\x93', 'Late', 'Middle', 'Chinese']  
 ['Cornish', 'Regulated', 'by', 'Cornish', 'Language', 'Partnership']  
 ['Czech', 'Regulated', 'by', 'Czech', 'Language', 'Institute']  
 ['Czech', 'Language', 'codes', 'Linguasphere', '53-AAA-da', '&', 'lt', ';',  
 '53-AAA-b', '.', '.', '.', '-d']

['Czech', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-AAA-daa', 'to', '53-AAA-dam', ')']

['Corsican', 'Writing', 'system', 'Latin', 'Alphabet', '(', 'Corsican', 'variant', ')']

['Danish', 'Regulated', 'by', 'Dansk', 'Sprogn', '\xc3', '\xa6', '\n', '(', '\xe2', '\x80', '\x9c', 'Danish', 'Language', 'Committee', '\xe2', '\x80', '\x9d', ')']

['Danish', 'Language', 'codes', 'Linguasphere', '52-AAA-bf', '&', 'amp', ';', '-ca', 'to', '-cj']

['Dioula', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]

['Esperanto', 'Total', 'speakers', 'Native', ':', '200', 'to', '1', ',', '000', '(', '1996', ',', 'est', '.', ')', ';', 'others', ':', '10', ',', '000', 'to', '2', ',', '000', ',', '000']

['Esperanto', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']

['Esperanto', 'Setting', 'and', 'usage', 'International', 'auxiliary', 'language']

['Etruscan', 'Language', 'extinction', '1st', 'century', 'AD']

['Estonian', 'Total', 'speakers', '1', ',', '048', ',', '660']

['Estonian', 'Spoken', 'in', 'Estonian', 'immigrant', 'communities', '(', 'see', 'Estonians', ')']

['Estonian', 'Regulated', 'by', 'Institute', 'of', 'the', 'Estonian', 'Language', '/', 'Eesti', 'Keele', 'Instituut', ',', 'Emakeele', 'Selts', '(', 'semi-official', ')']

['Estonian', 'Language', 'codes', 'ISO', '639-3', ':', 'ekk', '\xe2', '\x80', '\x93', 'Standard', 'Estonian']

['Frisian', 'Total', 'speakers', '500', ',', '000']

['Frisian', 'Regulated', 'by', 'unofficial', ':', 'the', 'Seelter', 'Buund', '(', 'for', 'Sater', 'Frisian', ')', ',', 'the', 'Nordfriisk', 'Instituut', '(', 'for', 'North', 'Frisian', ')']

['Frisian', 'Language', 'codes', 'ISO', '639-3', ':', 'fry', '\xe2', '\x80', '\x93', 'West', 'Frisian']

['Frisian', 'Language', 'codes', 'ISO', '639-3', 'stq', '\xe2', '\x80', '\x93', 'Saterland', 'Frisian']

['Frisian', 'Language', 'codes', 'ISO', '639-3', 'frr', '\xe2', '\x80', '\x93', 'North', 'Frisian']

['Frisian', 'Language', 'codes', 'Linguasphere', 'Original', 'and', '21st', 'century', 'distribution', 'of', 'Frisian', '(', 'without', 'the', 'island', 'of', 'R', '\xc3', '\xb8', 'm', '\xc3', '\xb8', '(', 'Danish', 'language', ')', ')']

['French', 'Writing', 'system', 'Latin', 'alphabet', '(', 'French', 'variant', ')']

['French', 'Official', 'language', 'in', 'Central', 'African', 'Republic']

['French', 'Official', 'language', 'in', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]

['French', 'Official', 'language', 'in', 'French', 'Southern', 'and',

'Antarctic', 'Lands']  
 ['French', 'Regulated', 'by', 'Acad', '\xc3', '\xa9', 'mie', 'fran',  
 '\xc3', '\xa7', 'aise', '(', 'French', 'Academy', ')']  
 ['West', 'Flemish', 'Total', 'speakers', '~', '1', '.', '16', 'million',  
 'speakers']  
 ['Faroese', 'Total', 'speakers', '60', ',', '000', '\xe2', '\x80', '\x93',  
 '80', ',', '000']  
 ['Faroese', 'Regulated', 'by', 'Faroese', 'Language', 'Board', 'F', '\xc3',  
 '\xb8', 'royska', 'm', '\xc3', '\xa1', 'lnevndin']  
 ['Faroese', 'Language', 'codes', 'Linguasphere', 'Faroese', 'keyboard',  
 'layout']  
 ['Persian', 'Total', 'speakers', 'ca', '.', '60-70', 'million', ',', 'as',  
 'first', 'language', '(', '2006', 'estimates', ')']  
 ['Persian', 'Region', 'Western', 'Asia', ',', 'Central', 'Asia']  
 ['Persian', 'Regulated', 'by', 'Academy', 'of', 'Persian', 'Language',  
 'and', 'Literature', '(', 'Iran', ')']  
 ['Persian', 'Language', 'codes', 'Linguasphere', '58-AAC', '(', 'Wider',  
 'Persian', ')', '&', 'gt', ';', '58-AAC-c', '(', 'Central', 'Persian', ')']  
 ['German', 'Spoken', 'in', 'Primarily', 'in', 'German-speaking', 'Europe',  
 ',', 'as', 'a', 'minority', 'language', 'and', 'amongst', 'the', 'German',  
 'diaspora', 'worldwide']  
 ['German', 'Writing', 'system', 'Latin', 'alphabet', '(', 'German',  
 'variant', ')']  
 ['German', 'Official', 'language', 'in', 'Italy', '(', 'South', 'Tyrol',  
 ')']  
 ['German', 'Regulated', 'by', '(', 'German', 'orthography', 'regulated',  
 'by', 'the', 'Council', 'for', 'German', 'Orthography', '(', 'Rat', 'f',  
 '\xc3', '\xbc', 'r', 'deutsche', 'Rechtschreibung', ')', ')']  
 ['German', 'Recognised', 'minority', 'language', 'in', 'Czech', 'Republic']  
 ['German', 'Recognised', 'minority', 'language', 'in', 'Namibia', '(',  
 'National', 'language', ';', 'official', 'language', '1984', '\xe2',  
 '\x80', '\x93', '90', ')']  
 ['German', 'Recognised', 'minority', 'language', 'in', 'Poland', '(',  
 'Auxiliary', 'language', 'in', '22', 'municipalities', 'in', 'Opole',  
 'Voivodeship', ')']  
 ['German', 'Recognised', 'minority', 'language', 'in', 'Slovakia', '(',  
 'Official', 'municipal', 'language', 'of', 'Krahule', '/', 'Blaufu',  
 '\xc3', '\x9f', ')']  
 ['German', 'Recognised', 'minority', 'language', 'in', 'Vatican', 'City',  
 '(', 'Administrative', 'and', 'commanding', 'language', 'of', 'the',  
 'Swiss', 'Guard', ')']  
 ['German', 'Language', 'codes', 'ISO', '639-3', ':', 'deu', '\xe2', '\x80',  
 '\x93', 'New', 'High', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'gmh', '\xe2', '\x80',  
 '\x93', 'Middle', 'High', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'goh', '\xe2', '\x80',

'\x93', 'Old', 'High', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'geh', '\xe2', '\x80',  
 '\x93', 'Hutterite', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'nds', '\xe2', '\x80',  
 '\x93', 'Low', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'sli', '\xe2', '\x80',  
 '\x93', 'Lower', 'Silesian']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'pdc', '\xe2', '\x80',  
 '\x93', 'Pennsylvania', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'swg', '\xe2', '\x80',  
 '\x93', 'Swabian', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'gsw', '\xe2', '\x80',  
 '\x93', 'Swiss', 'German']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'sxu', '\xe2', '\x80',  
 '\x93', 'Upper', 'Saxon']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'wae', '\xe2', '\x80',  
 '\x93', 'Walser', 'German']  
 ['German', 'Language', 'codes', 'Linguasphere', '52-AC', '(',  
 'Continental', 'West', 'Germanic', ')', '&', 'gt', ';', '52-ACB', '(',  
 'Deutsch', '&', 'amp', ';', 'Dutch', ')', '&', 'gt', ';']  
 ['German', 'Language', 'codes', 'Linguasphere', '52-ACB-d', '(', 'Central',  
 'German', 'incl', '.', '52-ACB', '\xe2', '\x80', '\x93', 'dl', '&', 'amp',  
 ';', '-dm', 'Standard', '/', 'Generalised', 'High', 'German', ')', '+',  
 '52-ACB-e', '&', 'amp', ';', '-f', '(', 'Upper', 'German', '&', 'amp', ';',  
 'Swiss', 'German', ')', '+', '52-ACB-g', '(', 'Yiddish', ')', '+', '52-ACB-  
 h', '(', '\xc3', '\xa9', 'migr', '\xc3', '\xa9', 'German', 'varieties',  
 'incl']  
 ['German', 'Language', 'codes', 'Linguasphere', '52-ACB-hc', 'Hutterite',  
 'German', '&', 'amp', ';', '52-ACB-he', 'Pennsylvania', 'German', 'etc',  
 '.', ')', '+', '52-ACB-i', '(', 'Yenish', ')', ';', 'totalling', '285',  
 'varieties', ':', '52-ACB-daa', 'to', '52-ACB-i']  
 ['German', 'Language', 'codes', 'Linguasphere', 'Nations', 'where',  
 'German', 'is', 'spoken']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', ':', 'grc', '\xe2', '\x80',  
 '\x93', 'Ancient', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'ell', '\xe2', '\x80',  
 '\x93', 'Modern', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'pnt', '\xe2', '\x80',  
 '\x93', 'Pontic', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'gmy', '\xe2', '\x80',  
 '\x93', 'Mycenaean', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'gkm', '\xe2', '\x80',  
 '\x93', 'Medieval', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'cpg', '\xe2', '\x80',  
 '\x93', 'Cappadocian', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'tsd', '\xe2', '\x80',

'\x93', 'Tsakonian', 'Greek']  
 ['Greek', 'Language', 'codes', 'Linguasphere', '56-AAA-a', '(',  
 'varieties']  
 ['Greek', 'Language', 'codes', 'Linguasphere', ':', '56-AAA-aa', 'to', '-  
 am', ')']  
 ['Gallurese', 'Total', 'speakers', '100', ',', '000', '(', '1993', ')']  
 ['Hebrew', 'Total', 'speakers', 'Total', 'Speakers', '&', 'lt', ';', '9',  
 ',', '000', ',', '000']  
 ['Hebrew', 'Total', 'speakers', 'First', 'Language', '5', ',', '300', ',',  
 '000', '(', '2009', ')', ';']  
 ['Hebrew', 'Total', 'speakers', 'Second', 'Language', '2', ',', '500', ',',  
 '000', '(', '2009', ')']  
 ['Hebrew', 'Total', 'speakers', 'Home', 'Language', '200', ',', '000', '(',  
 'approx', '.', ')', 'in', 'the', 'United', 'States', 'speak', 'Hebrew',  
 'at', 'home1']  
 ['Hebrew', 'Total', 'speakers', '1United', 'States', 'Census', '2000',  
 'PHC-T-37', '.', 'Ability', 'to', 'Speak', 'English', 'by', 'Language',  
 'Spoken', 'at', 'Home', ':', '2000', '.', 'Table', '1a', '.', 'PDF', '(',  
 '11', '.', '8', 'KB', ')']  
 ['Hebrew', 'Total', 'speakers', 'Palestinian', 'territories', 'Second',  
 'Language', '500', ',', '000', '-', '1', ',', '000', ',', '000']  
 ['Hebrew', 'Total', 'speakers', 'Extinct', 'as', 'a', 'regularly',  
 'spoken', 'language', 'by', 'the', '4th', 'century', 'CE', ',', 'but',  
 'survived', 'as', 'a', 'liturgical', 'and', 'literary', 'language', ';',  
 'revived', 'in', 'the', '1880s']  
 ['Hebrew', 'Spoken', 'in', 'Global', '(', 'as', 'a', 'liturgical',  
 'language', 'for', 'Judaism', ')', ',', 'in', 'West', 'Bank', ',', 'and',  
 'Gaza']  
 ['Hebrew', 'Pronunciation', 'standard', 'Israeli', ':', '(', '\xca',  
 '\x94', ')', 'iv', '\xcb', '\x88', '\xca', '\x81', 'it', '-', '(', '\xca',  
 '\x94', ')', 'iv', '\xcb', '\x88', '\xc9', '\xbe', 'it', ',', ]  
 ['Hebrew', 'Pronunciation', 'standard', 'Israeli', '(', 'Sephardi', ')',  
 ':', '\xca', '\x95', 'iv', '\xcb', '\x88', '\xc9', '\xbe', 'it', ',', ]  
 ['Hebrew', 'Pronunciation', 'Iraqi', ':', '\xca', '\x95', 'ib', '\xcb',  
 '\x88', 'ri', '\xcb', '\x90', '\xce', '\xb8', ',', ]  
 ['Hebrew', 'Regulated', 'by', 'Academy', 'of', 'the', 'Hebrew', 'Language']  
 ['Hebrew', 'Regulated', 'by', '\xd7', '\x94', '\xd7', '\x90', '\xd7',  
 '\xa7', '\xd7', '\x93', '\xd7', '\x9e', '\xd7', '\x99', '\xd7', '\x94',  
 '\xd7', '\x9c', '\xd7', '\x9c', '\xd7', '\xa9', '\xd7', '\x95', '\xd7',  
 '\x9f', '\xd7', '\x94', '\xd7', '\xa2', '\xd7', '\x91', '\xd7', '\xa8',  
 '\xd7', '\x99', '\xd7', '\xaa', '(', 'HaAkademia', 'LaLashon', 'Ha',  
 '\xe2', '\x80', '\x98', 'Ivrit', ')']  
 ['Hebrew', 'Language', 'codes', 'ISO', '639-3', ':', 'heb', '\xe2', '\x80',  
 '\x93', 'Modern', 'Hebrew']  
 ['Hebrew', 'Language', 'codes', 'ISO', '639-3', 'hbo', '\xe2', '\x80',  
 '\x93', 'Ancient', 'Hebrew']

['Hindi', 'Total', 'speakers', '550', 'million', 'in', '2011']  
['Hindi', 'Regulated', 'by', 'Central', 'Hindi', 'Directorate', '(',  
'India', ')']  
['Hindi', 'Language', 'codes', 'Linguasphere', '59-AAF-q', '(', 'with',  
'Urdu', ',']  
['Hindi', 'Language', 'codes', 'Linguasphere', 'including', '58',  
'varieties', ':', '59-AAF-qaa', 'to', '59-AAF-qil', ')']  
['Hindi', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains',  
'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
'conjuncts', '.', 'More']  
['Hungarian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Hungarian',  
'variant', ')']  
['Hungarian', 'Regulated', 'by', 'Research', 'Institute', 'for',  
'Linguistics', 'of', 'the', 'Hungarian', 'Academy', 'of', 'Sciences']  
['Hungarian', 'Language', 'codes', 'ISO', '639-3', 'Regions', 'of', 'the',  
'Carpathian', 'Basin', 'where', 'the', 'Hungarian', 'language', 'is',  
'spoken']  
['Hawaiian', 'Spoken', 'in', 'Hawai', '\xca', '\xbb', 'i', ':',  
'concentrated', 'on', 'Ni', '\xca', '\xbb', 'ihau', 'and', 'Hawai', '\xca',  
'\xbb', 'i', ', ', 'but', 'speakers', 'throughout', 'the', 'Hawaiian',  
'Islands', 'and', 'the', 'U', '.', 'S', '.', 'mainland']  
['Hawaiian', 'Total', 'speakers', '~', '2000', 'native']  
['Hawaiian', 'Total', 'speakers', '~', '27', ', ', '160']  
['Hawaiian', 'Writing', 'system', 'Latin', 'plus', 'Hawaiian', 'symbols']  
['Hawaiian', 'Official', 'language', 'in', 'Hawai', '\xca', '\xbb', 'i',  
'(', 'with', 'English', ')']  
['Hawaiian', 'Official', 'language', 'in', ':', 'U', '.', 'S', '.',  
'Mainland']  
['Italian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Italian',  
'variant', ')']  
['Ido', 'Total', 'speakers', '100', '\xe2', '\x80', '\x93', '200']  
['Ido', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
'language']  
['Ido', 'Created', 'by', 'A', 'group', 'of', 'reformist', 'Esperanto',  
'speakers']  
['Ido', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
'language']  
['Interlingua', 'Created', 'by', 'International', 'Auxiliary', 'Language',  
'Association']  
['Japanese', 'Writing', 'system', 'Japanese', 'logographs', 'and',  
'syllabaries']  
['Japanese', 'Regulated', 'by', 'Japanese', 'government', 'plays', 'major',  
'role']  
['Korean', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
['Korean', 'Total', 'speakers', '78', 'million', '(', '1999', 'est', '.',

')]  
 ['Korean', 'Official', 'language', 'in', 'Yanbian', ',', 'People's',  
 'Republic', 'of', 'China']  
 ['Korean', 'Regulated', 'by', ':', 'The', 'National', 'Institute', 'of',  
 'the', 'Korean', 'Language']  
 ['Korean', 'Regulated', 'by', 'The', 'Language', 'Research', 'Institute',  
 'of', 'Social', 'Science']  
 ['Klingon', 'Regulated', 'by', 'Klingon', 'Language', 'Institute']  
 ['Klingon', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Klingon',  
 'writing', 'systems', 'constructed', 'languages']  
 ['Klingon', 'Setting', 'and', 'usage', 'Star', 'Trek', 'films', 'and',  
 'television', 'series', '(', 'TNG', ',', 'DS9', ',', 'Voyager', ',',  
 'Enterprise', ')', 'and', 'the', 'opera', '"u"]  
 ['Kashubian', 'Total', 'speakers', '50', ',', '000']  
 ['Kashubian', 'Regulated', 'by', 'Kashubian', 'Language', 'Council']  
 ['Latin', 'Spoken', 'in', 'Roman', 'Monarchy', ',', 'Roman', 'Republic',  
 ',', 'Roman', 'Empire', ',', 'Medieval', 'and', 'Early', 'modern',  
 'Europe', ',', 'Armenian', 'Kingdom', 'of', 'Cilicia', '(', 'as', 'lingua',  
 'franca', ')', ',', ']  
 ['Latin', 'Language', 'codes', 'Linguasphere', 'Greatest', 'extent', 'of',  
 'the', 'Roman', 'Empire', '.', 'Latin', 'was', 'by', 'no', 'means',  
 'confined', 'to', 'these', 'regions']  
 ['Loglan', 'Category', '(', 'sources', ')', 'eight', 'of', 'the', 'most',  
 'common', 'languages', ':', 'English', ',', 'Chinese', '(', 'Beijing',  
 'dialect', ')', ',', 'Hindi', ',', 'Russian', ',', 'Spanish', ',',  
 'French', ',', 'Japanese', ',', 'German']  
 ['Luxembourgish', 'Total', 'speakers', '390', ',', '000']  
 ['Luxembourgish', 'Language', 'codes', 'ISO', '639-3', 'Area', 'where',  
 'Luxembourgisch', '(', 'hatched', ')', 'and', 'related', 'Moselle',  
 'Franconian', 'is', 'spoken']  
 ['Maltese', 'Total', 'speakers', '371', ',', '900', '(', '1975', ')']  
 ['Maltese', 'Regulated', 'by', 'National', 'Council', 'for', 'the',  
 'Maltese', 'Language']  
 ['Macedonian', 'Total', 'speakers', '1', '.', '6', '\xe2', '\x80', '\x93',  
 '3', '.', '0', 'million']  
 ['Macedonian', 'Regulated', 'by', 'Macedonian', 'Language', 'Institute',  
 '"', 'Krste', 'Misirkov', '"', 'at', 'the', 'Ss', '.', 'Cyril', 'and',  
 'Methodius', 'University', 'of', 'Skopje']  
 ['Macedonian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-  
 h']  
 ['Mandarin', 'Total', 'speakers', 'Overall', ':', '1', ',', '365', ',',  
 '053', ',', '177']  
 ['Mandarin', 'Dialects', '(', 'Jin', ',', 'disputed', ')']  
 ['Mandarin', 'Region', 'Most', 'of', 'northern', 'and', 'southwestern',  
 'China', ',', 'Taiwan', 'and', 'Singapore', '(', 'and', 'other',  
 'overseas', 'Chinese', 'communities', ')']



['Mandarin', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Mandarin', 'Spoken', 'in', 'United', 'States', '(', 'chiefly', 'New', 'York', 'City', ')']  
 ['Manx', 'Total', 'speakers', 'and', '1', ',', '689', 'people', '(', '2', '.', '2', '%', 'total', 'population', ')', 'in', 'the', 'Isle', 'of', 'Man', 'professing', 'some', 'knowledge', 'of', 'the', 'language', '(', '2001', ')']  
 ['Manx', 'Regulated', 'by', 'Coonseil', 'ny', 'Gaelgey', '(', 'Manx', 'Gaelic', 'Council', ')']  
 ['Marathi', 'Region', 'Dadra', '&', 'amp', ';', 'Nagar', 'Haveli']  
 ['Marathi', 'Region', 'and', 'Daman', '&', 'amp', ';', 'Diu']  
 ['Marathi', 'Spoken', 'in', 'Australia', '&', 'amp', ';', 'New', 'Zealand']  
 ['Marathi', 'Regulated', 'by', 'Maharashtra', 'Sahitya', 'Parishad', '&', 'amp', ';', 'various', 'other', 'institutions']  
 ['Marathi', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Marathi', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Mobilian', 'Language', 'extinction', 'active', 'through', '1960s']  
 ['Mobilian', 'Spoken', 'in', 'Gulf', 'coast', 'and', 'Mississippi', 'Valley']  
 ['Norwegian', 'Spoken', 'in', 'Norway', '(', '4', '.', '8', 'million', ')', ',', '']  
 ['Norwegian', 'Spoken', 'in', 'United', 'States', '(', '55', ',', '311', ')']  
 ['Norwegian', 'Spoken', 'in', 'Denmark', '(', '15', ',', '000', ')']  
 ['Norwegian', 'Spoken', 'in', 'Canada', '(', '7', ',', '710', ')']  
 ['Norwegian', 'Standard', 'forms', 'Nynorsk', '(', 'official', ')']  
 ['Norwegian', 'Standard', 'forms', 'Bokm', '\xc3', '\xa5', 'l', '(', 'official', ')', '/', 'Riksm', '\xc3', '\xa5', 'l', '(', 'unofficial', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Language', 'Council', '(', 'Bokm', '\xc3', '\xa5', 'l', 'and', 'Nynorsk', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Academy', '(', 'Riksm', '\xc3', '\xa5', 'l', ')']  
 ['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-ba', 'to', '-be', '&', 'amp', ';']  
 ['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-cf', 'to', '-cg']  
 ['Northern', 'Sotho', 'Total', 'speakers', '4', ',', '208', ',', '940']  
 ['Northern', 'Regulated', 'by', 'Pan', 'South', 'African', 'Language', 'Board']  
 ['New', 'Latin', 'Language', 'extinction', 'developed', 'into', 'contemporary', 'Latin', 'by', 'the', '20th', 'century']  
 ['Novial', 'Category', '(', 'sources', ')', 'Romance', 'and', 'Germanic', 'languages', ';', 'also', 'Occidental', 'and', 'Ido']

['Norn', 'Language', 'extinction', 'by', 'the', '18th', 'century', '(', '19th', 'century', 'at', 'the', 'latest', ')', ';', 'much', 'earlier', 'in', 'Caithness']

['Occitan', 'Total', 'speakers', '500', ',', '000-1', ',', '000', ',', '000']

['Occitan', 'Official', 'language', 'in', 'Spain', '(', 'also', 'the', 'preferred', 'language', 'for', 'use', 'in', 'the', 'Val', 'd'Aran', ')']

['Occitan', 'Language', 'codes', 'Linguasphere', '51-AAA-g', '&', 'amp', ';', '51-AAA-f']

['Prussian', 'Language', 'extinction', 'Late', '17th', '/', 'Early', '18th', 'century']

['Old', 'Norse', 'Language', 'extinction', 'developed', 'into', 'the', 'various', 'North', 'Germanic', 'languages', 'by', 'the', '14th', 'century']

['Old', 'Writing', 'system', 'Runic', ',', 'later', 'Latin', 'alphabet', '(', 'Old', 'Norse', 'variant', ')']

['Old', 'English', 'Language', 'extinction', 'mostly', 'developed', 'into', 'Middle', 'English', 'by', 'the', '13th', 'century']

['Old', 'Writing', 'system', 'Runic', ',', 'later', 'Latin', 'alphabet', '(', 'Old', 'English', 'variant', ')']

['Occidental', 'Setting', 'and', 'usage', 'International', 'auxiliary', 'language']

['Occidental', 'Created', 'by', 'Edgar', 'de', 'Wahl', '(', '1922', ')']

['Polish', 'Official', 'language', 'in', ':', 'Czech', 'Republic']

['Polish', 'Regulated', 'by', 'Polish', 'Language', 'Council']

['Polish', 'Language', 'codes', 'Linguasphere', '53-AAA-cc', '&', 'lt', ';', '53-AAA-b', '.', '.', '.', '-d']

['Polish', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-AAA-cca', 'to', '53-AAA-ccu', ')']

['Portuguese', 'Total', 'speakers', 'Native', ':', '236', '.', '1', 'million']

['Portuguese', 'Total', 'speakers', 'Total', ':', '272', '.', '9', 'million']

['Portuguese', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Portuguese', 'variant', ')']

['Portuguese', 'Pronunciation', 'po', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'es', '(', 'BP-ga', '\xc3', '\xba', 'cho', ')', '1']

['Portuguese', 'Regulated', 'by', 'International', 'Portuguese', 'Language', 'Institute', ';', 'CPLP', ';', 'Academia', 'Brasileira', 'de', 'Letras', '(', 'Brazil', ')', ';', 'Academia', 'das', 'Ci', '\xc3', '\xaa', 'ncias', 'de', 'Lisboa', ',', 'Classe', 'de', 'Letras', '(', 'Portugal', ')']

['Standard', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']

['Standard', 'Official', 'language', 'in', 'People's', 'Republic', 'of', 'China']

['Standard', 'Official', 'language', 'in', 'Singapore', '(', 'official',

'but', 'not', 'national', 'language', ')']  
 ['Standard', 'Regulated', 'by', 'In', 'the', 'PRC', ':', 'National',  
 'Language', 'Regulating', 'Committee']  
 ['Standard', 'Regulated', 'by', 'In', 'Singapore', ':', 'Promote',  
 'Mandarin', 'Council', '/', 'Speak', 'Mandarin', 'Campaign']  
 ['Punjabi', 'Language', 'Total', 'speakers', '88', ',', '000', ',', '000',  
 '(', 'Ethnologue', '2005', 'estimate', ')']  
 ['Punjabi', 'Writing', 'system', 'Devanagari', 'in', 'Haryana', ',',  
 'Himachal', 'Pradesh', ',', 'Delhi', 'and', 'Punjabi', 'Hindu', 'diaspora',  
 ')']  
 ['Punjabi', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'native', 'Punjabi', 'speakers', 'in', 'India', 'and', 'Pakistan', 'This',  
 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering',  
 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and',  
 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Russian', 'Total', 'speakers', 'Secondary', 'language', ':', '114',  
 'million', '(', '2006', ')', 'to', '125', 'million']  
 ['Russian', 'Spoken', 'in', 'Australia', 'and', 'Latin', 'America']  
 ['Russian', 'Official', 'language', 'in', 'South', 'Ossetia', '(', 'co-  
 official', ')']  
 ['Russian', 'Regulated', 'by', 'Russian', 'Language', 'Institute', 'at',  
 'the', 'Russian', 'Academy', 'of', 'Sciences']  
 ['Russian', 'Language', 'codes', 'Linguasphere', '53-AAA-ea', '&', 'lt',  
 ';', '53-AAA-e']  
 ['Russian', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':',  
 '53-AAA-eaa', 'to', '53-AAA-eat', ')']  
 ['Romansh', 'Total', 'speakers', '35', ',', '095', '(', 'Swiss', 'federal',  
 'census', '2000', ')']  
 ['Romanian', 'Region', 'Southeastern', ',', 'Central', 'and', 'Eastern',  
 'Europe']  
 ['Romanian', 'Spoken', 'in', ':', 'North', 'and', 'South', 'America']  
 ['Romanian', 'Spoken', 'in', 'Western', 'and', 'Southern', 'Europe']  
 ['Romanian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Romanian',  
 'variant', ')']  
 ['Romanian', 'Language', 'codes', 'Linguasphere', '51-AAD-c', '(',  
 'varieties']  
 ['Romanian', 'Language', 'codes', 'Linguasphere', ':', '51-AAD-ca', 'to',  
 '-ck', ')', 'Map', 'of', 'the', 'Romanian-speaking', 'territories']  
 ['Swedish', 'Total', 'speakers', '~', '10', 'million']  
 ['Swedish', 'Region', 'Northern', 'Europe', ',', 'parts', 'of', 'USA',  
 'and', 'other', 'countries']  
 ['Swedish', 'Spoken', 'in', 'Sweden', '(', '9', '.', '4', 'million', ')']  
 ['Swedish', 'Spoken', 'in', 'Finland', '(', '290', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'USA', '(', '70', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'Spain', '(', '40', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'Great', 'Britain', '(', '30', ',', '000', ')']

['Swedish', 'Spoken', 'in', 'Canada', '(', '20', ',', '000', ')']  
 ['Swedish', 'Regulated', 'by', 'Swedish', 'Language', 'Council', '(', 'in', 'Sweden', ')']  
 ['Swedish', 'Language', 'codes', 'Linguasphere', '52-AAA-ck', 'to', '-cw']  
 ['Spanish', 'Official', 'language', 'in', 'Organization', 'of', 'American', 'States']  
 ['Spanish', 'Official', 'language', 'in', 'Union', 'of', 'South', 'American', 'Nations']  
 ['Spanish', 'Official', 'language', 'in', 'Central', 'American', 'Integration', 'System']  
 ['Spanish', 'Official', 'language', 'in', 'North', 'American', 'Free', 'Trade', 'Agreement']  
 ['Spanish', 'Regulated', 'by', 'Association', 'of', 'Spanish', 'Language', 'Academies', '(', 'Real', 'Academia', 'Espa', '\xc3', '\xb1', 'ola', 'and', '21', 'other', 'national', 'Spanish', 'language', 'academies', ')']  
 ['Logudorese', 'Total', 'speakers', '500', ',', '000-700', ',', '000']  
 ['Sardinian', 'Total', 'speakers', '1', '.', '85', 'million']  
 ['Sardinian', 'Regulated', 'by', 'Limba', 'Sarda', 'Comuna', '(', 'Common', 'Sardinian', 'Language', ')', 'code']  
 ['Sardinian', 'Language', 'codes', 'Linguasphere', '51-AAA-s', '+', '(', 'Corso-Sardinian', ')']  
 ['Sardinian', 'Language', 'codes', 'Linguasphere', '51-AAA-pd', '&', 'amp', ';', '-pe']  
 ['Swiss', 'German', 'Total', 'speakers', '4', ',', '500', ',', '000']  
 ['Swiss', 'Language', 'codes', 'Linguasphere', '52-ACB-f', '(', '45', 'varieties', ':', '52-ACB-faa', 'to', '-fkb', ')']  
 ['Sanskrit', 'Total', 'speakers', '14', ',', '135', 'native', 'speakers', 'in', 'India', '(', '2001', ')']  
 ['Sanskrit', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Serbo-Croatian', 'Total', 'speakers', '16', '.', '3', 'million']  
 ['Serbo-Croatian', 'Region', 'the', 'Western', 'Balkans', 'Ethnicity']  
 ['Serbo-Croatian', 'Standard', 'forms', 'Standard', 'Serbo-Croatian', '(', 'defunct', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Bosnia', 'and', 'Herzegovina', '(', 'as', 'Bosnian', ',', 'Croatian', ',', 'Serbian', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Croatia', '(', 'as', 'Croatian', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Kosovo', '(', 'as', 'Serbian', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Montenegro', '(', 'as', 'Montenegrin', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Serbia', '(', 'as', 'Serbian', ')']

['Serbo-Croatian', 'Regulated', 'by', 'Council', 'for', 'Standard',  
 'Croatian', 'Language', 'Norm']  
 ['Serbo-Croatian', 'Regulated', 'by', 'Board', 'for', 'Standardization',  
 'of', 'the', 'Serbian', 'Language']  
 ['Slovene/Slovenian', 'Total', 'speakers', '2', '.', '5', 'million']  
 ['Slovene/Slovenian', 'Region', 'Central', 'Southern', 'and',  
 'Southeastern', 'Europe']  
 ['Slovene/Slovenian', 'Writing', 'system', 'Latin', 'alphabet', '(',  
 'Slovene', 'variant', ')']  
 ['Slovene/Slovenian', 'Official', 'language', 'in', 'Regional', 'or',  
 'local', 'official', 'language', 'in']  
 ['Slovene/Slovenian', 'Language', 'codes', 'Linguasphere', '53-AAA-f', '(',  
 '51', 'varieties', ')']  
 ['Slovak', 'Spoken', 'in', 'Czech', 'Republic', ',']  
 ['Slovak', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Slovak',  
 'variant', ')']  
 ['Slovak', 'Regulated', 'by', 'Ministry', 'of', 'Culture', 'of', 'the',  
 'Slovak', 'Republic']  
 ['Slovak', 'Language', 'codes', 'Linguasphere', '53-AAA-db', '&', 'lt',  
 ';', '53-AAA-b', '.', '.', '.', '-d']  
 ['Slovak', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-  
 AAA-dba', 'to', '53-AAA-dbs', ')']  
 ['Scottish', 'Total', 'speakers', '92', ',', '400', 'people', 'aged',  
 'three', 'and', 'over', 'in', 'Scotland', 'had', 'some', 'Gaelic',  
 'language', 'ability', 'in', '2001', 'with', 'an', 'additional', '2', ',',  
 '000', 'in', 'Nova', 'Scotia', '.', '1', ',', '610', 'speakers', 'in',  
 'the', 'United', 'States', 'in', '2000', '.', '822', 'in', 'Australia',  
 'in', '2001', '.', '669', 'in', 'New', 'Zealand', 'in', '2006']  
 ['Swahili', 'Spoken', 'in', 'Comoros', '(', 'as', 'Comorian', ')']  
 ['Swahili', 'Spoken', 'in', 'Seychelles', '(', 'as', 'Shimaore', ')']  
 ['Swahili', 'Total', 'speakers', 'First', 'and', 'second', 'language', ':',  
 '50', '+', 'million']  
 ['Tamil', 'Total', 'speakers', '65', ',', '675', ',', '200']  
 ['Tamil', 'Official', 'language', 'in', 'India', '(', 'Tamil', 'Nadu', ',',  
 'Puducherry', ')', ',']  
 ['Tamil', 'Official', 'language', 'in', 'Sri', 'Lanka', ',', 'and']  
 ['Tamil', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'native', 'Tamil', 'speakers', 'in', 'India', 'and', 'Sri', 'Lanka',  
 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering',  
 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and',  
 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Turkish', 'Total', 'speakers', 'Native', ':', '+', '83', 'million']  
 ['Turkish', 'Total', 'speakers', 'Total', ':', '+', '91', 'million']  
 ['Turkish', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Turkish',  
 'variant', ')']  
 ['Turkish', 'Standard', 'forms', 'Ottoman', 'Turkish', '(', 'defunct', ')']

['Turkish', 'Regulated', 'by', 'Turkish', 'Language', 'Association']  
 ['Turkish', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Setswana', 'Total', 'speakers', '4', ',', '521', ',', '700']  
 ['Tocharian', 'languages', 'Language', 'extinction', '9th', 'century', 'AD']  
 ['Tocharian', 'Spoken', 'in', 'Tarim', 'Basin', 'in', 'Central', 'Asia']  
 ['Tocharian', 'Language', 'codes', 'ISO', '639-3', ':', 'xto', '\xe2', '\x80', '\x93', 'Tocharian', 'A']  
 ['Tocharian', 'Language', 'codes', 'ISO', '639-3', 'txb', '\xe2', '\x80', '\x93', 'Tocharian', 'B']  
 ['Tagalog', 'Region', 'Central', 'and', 'South', 'Luzon']  
 ['Tagalog', 'Writing', 'system', 'Latin', '(', 'Tagalog', 'or', 'Filipino', 'variant', ')', ';']  
 ['Tagalog', 'Official', 'language', 'in', 'Philippines', '(', 'in', 'the', 'form', 'of', 'Filipino', ')']  
 ['Tagalog', 'Official', 'language', 'in', 'New', 'Jersey', '(', 'minority', 'language', ')']  
 ['Tagalog', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Tagalog', 'Language', 'codes', 'Linguasphere', 'The', 'map', 'shows', 'the', 'areas', 'where', 'the', 'language', 'is', 'spoken', 'by', 'a', 'significant', 'population', '.', 'Overseas', 'Filipinos', 'use', 'the', 'languages', 'as', 'a', 'lingua', 'franca', '.', 'There', 'are', 'over', 'a', 'million', 'speakers', 'in', 'Saudi', 'Arabia', 'and', 'in', 'the', 'United', 'States', ',', 'plus', '90', 'million', 'in', 'the', 'Philippines', 'itself']  
 ['Urdu', 'Total', 'speakers', '60', '.', '6', 'million']  
 ['Urdu', 'Writing', 'system', 'Urdu', 'alphabet', '(', 'Nasta', '\xca', '\xbf', 'l', '\xc4', '\xab', 'q', 'script', ')', ',', 'see', 'other', 'less-used', 'writing', 'systems', 'below']  
 ['Urdu', 'Official', 'language', 'in', 'India', '(', 'in', '5', 'states', ')']  
 ['Urdu', 'Regulated', 'by', 'National', 'Language', 'Authority', '(', 'Pakistan', ')', ';']  
 ['Urdu', 'Regulated', 'by', 'National', 'Council', 'for', 'Promotion', 'of', 'Urdu', 'Language', '(', 'India', ')']  
 ['Urdu', 'Language', 'codes', 'Linguasphere', '59-AAF-q', '(', 'with', 'Hindi', ',']  
 ['Urdu', 'Language', 'codes', 'Linguasphere', 'including', '58', 'varieties', ':', '59-AAF-qaa', 'to', '59-AAF-qil', ')']  
 ['Vietnamese', 'Total', 'speakers', '70-73', 'million', 'native', '(', 'includes', '3', 'million', 'overseas', ')']  
 ['Vietnamese', 'Total', 'speakers', '80', '+', 'million', 'total']  
 ['Vietnamese', 'Pronunciation', 'ti', '\xc9', '\x99', '\xcc', '\x8c', '\xc5', '\x8b', 'vi', '\xc9', '\x99', '\xcc', '\x80', '\xcb', '\x80', 't', '(', 'Northern', ')']

['Vietnamese', 'Pronunciation', 'ti', '\xc9', '\x99', '\xcc', '\x8c', '\xc5', '\x8b', 'ji', '\xc9', '\x99', '\xcc', '\x80', 'k', '(', 'Southern', ')']

['Volap\xc3\xbck', 'Date', 'founded', '1879', '\xe2', '\x80', '\x93', '1880']

['Volap\xc3\xbck', 'Writing', 'system', 'Latin', 'vocabulary', 'from', 'English']

['Volap\xc3\xbck', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']

['Volap\xc3\xbck', 'Setting', 'and', 'usage', 'International', ':', 'mostly', 'in', 'Europe']

['Welsh', 'Total', 'speakers', '750', ',', '000', '+']

['Welsh', 'Total', 'speakers', ':', '\xe2', '\x80', '\x94', 'Wales', ':', '611', ',', '000']

['Welsh', 'Total', 'speakers', '\xe2', '\x80', '\x94', 'England', ':', '150', ',', '000']

['Welsh', 'Total', 'speakers', '\xe2', '\x80', '\x94', 'Chubut', 'Province', ',', 'Argentina', ':', '5', ',', '000']

['Welsh', 'Total', 'speakers', '\xe2', '\x80', '\x94', 'USA', ':', '2', ',', '500']

['Welsh', 'Total', 'speakers', '\xe2', '\x80', '\x94', 'Canada', ':', '2', ',', '200']

['Welsh', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Welsh', 'variant', ')']

['Welsh', 'Regulated', 'by', 'Bwrdd', 'yr', 'Iaith', 'Gymraeg', '(', 'Welsh', 'Language', 'Board', ')']

['Welsh', 'Language', 'codes', 'Linguasphere', '50-ABA', 'Percentage', 'of', 'Welsh', 'speakers', 'by', 'principal', 'area']

['Yiddish', 'Total', 'speakers', '1', ',', '762', ',', '320']

['Yiddish', 'Official', 'language', 'in', 'Official', 'minority', 'language', 'in']

['Yiddish', 'Official', 'language', 'in', 'Jewish', 'Autonomous', 'Oblast']

['Yiddish', 'Language', 'codes', 'ISO', '639-3', ':', 'ydd', '\xe2', '\x80', '\x93', 'Eastern', 'Yiddish']

['Yiddish', 'Language', 'codes', 'ISO', '639-3', 'yih', '\xe2', '\x80', '\x93', 'Western', 'Yiddish']

['Yiddish', 'Language', 'codes', 'Linguasphere', '52-ACB-g', '=', '52-ACB-ga', '(', 'West', ')', '+', '52-ACB-gb', '(', 'East', ')', ';', 'totalling', '11', 'varieties']

['Indonesian', 'Spoken', 'in', 'East', 'Timor', '(', 'as', 'a', '"', 'working', 'language', '"', ')']

['Malayalam', 'Total', 'speakers', '35', ',', '893', ',', '990']

['Malayalam', 'Total', 'speakers', '33', ',', '015', ',', '420', 'in', 'India', '(', '2001', ')', ',', '']

['Malayalam', 'Total', 'speakers', '1', ',', '847', ',', '902', 'in', 'other', 'countries', '(', '2007', ')']

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '447', ',', '440', 'in', 'Saudi', 'Arabia']

['Malayalam', 'Language', 'codes', 'ISO', '639-3', 'Number', 'of', 'Malayalam', 'speakers', 'by', 'Countries', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Language', 'Total', 'speakers', '38', 'million', 'native', '(', '2001', ',', 'only', 'India', ')', ',', '13', 'million', 'as', 'a', 'second', 'language', '(', '2001', ')', '=', '51', 'million']

['Language', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Kannada', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Telugu', 'Total', 'speakers', '130', 'million', 'native', 'speakers', 'as', 'of', '2011']

['Telugu', 'Region', 'Andhra', 'Pradesh', ',', 'Tamil', 'Nadu']

['Telugu', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Telugu', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Dari', '(', 'Afghan', 'Persian', ')', 'Total', 'speakers', '12-15', 'million', ',', 'as', 'first', 'language', '(', 'Tajiks', '11-13', 'm', ';', 'Hazaras', '2', '.', '5-3', 'm', ';', 'Aym', '\xc4', '\x81', 'qs', '3-3', '.', '5', 'm', ';', 'Pashtuns', '(', 'unknown', ')', ')]

['Dari', 'Language', 'codes', 'ISO', '639-3', ':', 'prs', '\xe2', '\x80', '\x93', 'Dari', ',', 'Afghan', 'Persian']

['Dari', 'Language', 'codes', 'Linguasphere', '58-AAC-ce', '(', 'Dari', ')', '+', '58-AAC-cdo', '&', 'amp', ';', 'cdp', '(', 'Hazaragi', ')', '+', '58-AAC-ck', '(', 'Aimaq', ')]

['Kurdish', 'Writing', 'system', 'Kurdish', 'alphabet', '(', 'modified', 'Perso-Arabic', 'alphabet', 'in', 'Iraq', 'and', 'Iran', ',', 'Latin', 'alphabet', 'in', 'Turkey', ',', 'Syria', 'and', 'Armenia', ',', 'modified', 'Cyrillic', 'in', 'the', 'former', 'USSR', ')]

['Kurdish', 'Official', 'language', 'in', 'Iraq', ':', 'status', 'as', 'official', 'language', 'alongside', 'Arabic']

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'sdh', '\xe2', '\x80', '\x93', 'Southern', 'Kurdish']

['Kurdish', 'Language', 'codes', 'Linguasphere', '58-AAA-a', '(', 'North', 'Kurdish', 'incl', '.', 'Kurmanji', '&', 'amp', ';', 'Kurmanjiki', ')', '+', '58-AAA-b', '(', 'Central', 'Kurdish', 'incl', '.', 'Dimli', '/', 'Zaza', '&', 'amp', ';', 'Gurani', ')', '+', '58-AAA-c', '(', 'South', 'Kurdish', 'incl', '.', 'Kurdi', ')]

['Sudovian', 'Language', 'extinction', '17th', 'c', '.', '?']



['Old', 'Church', 'Slavonic', 'Spoken', 'in', 'formerly', 'in', 'Slavic',  
 'areas', ',', 'under', 'the', 'influence', 'of', 'Byzantium', '(', 'both',  
 'Catholic', 'and', 'Orthodox', ')']  
 ['Old', 'Language', 'extinction', 'in', 'use', 'only', 'as', 'a',  
 'liturgical', 'language']  
 ['Latvian', 'Total', 'speakers', 'native', '1', '.', '39', 'million', '(',  
 'Latvia', ')']  
 ['Latvian', 'Total', 'speakers', '1', '.', '5', 'million', '(',  
 'Worldwide', ')']  
 ['Latvian', 'Total', 'speakers', 'second', 'language', ':', '400', ',',  
 '000']  
 ['Latvian', 'Regulated', 'by', 'Latvian', 'State', 'Language', 'Center']  
 ['Latvian', 'Language', 'codes', 'ISO', '639-3', 'lvs', '\xe2', '\x80',  
 '\x93', 'Standard', 'Latvian']  
 ['Mon', 'Total', 'speakers', 'Myanmar', ':', '742', ',', '900', ',',  
 'Total', ':', '850', ',', '530']  
 ['Mon', 'Spoken', 'in', 'Burma', '(', 'Myanmar', ')', ',', 'Thailand']  
 ['Mon', 'Writing', 'system', 'Burmese', 'alphabet', '(', 'itself',  
 'derived', 'from', 'the', 'Old', 'Mon', 'Indic-based', 'script', ')']  
 ['Michif', 'Region', 'Turtle', 'Mountain', 'Indian', 'Reservation', 'in',  
 'North', 'Dakota']  
 ['Belarusian', 'Language', 'codes', 'Linguasphere', '53-AAA-eb', '&', 'lt',  
 ';', '53-AAA-e']  
 ['Belarusian', 'Language', 'codes', 'Linguasphere', '(', 'varieties']  
 ['Belarusian', 'Language', 'codes', 'Linguasphere', ':', '53-AAA-eba',  
 'to', '53-AAA-ebg', ')']  
 ['Belarusian', 'Language', 'codes', 'Linguasphere', 'Belarusian-speaking',  
 'world', 'Legend', ':', 'Dark', 'blue', '-', 'territory', ',', 'where',  
 'Belarusian', 'language', 'is', 'used', 'chiefly']  
 ['Ladin', 'Total', 'speakers', '30', ',', '000']  
 ['Ladin', 'Region', 'Province', 'of', 'Belluno', ',', 'South', 'Tyrol',  
 'and', 'Trentino']  
 ['Ladin', 'Regulated', 'by', 'The', 'office', 'for', 'Ladin', 'language',  
 'planning']  
 ['Ladin', 'Regulated', 'by', 'Ladin', 'Cultural', 'Centre', 'Majon', 'di',  
 'Fascegn']  
 ['Ladin', 'Regulated', 'by', 'Istitut', 'Ladin', 'Micur', '\xc3', '\xa0',  
 'de', 'R', '\xc3', '\xbc']  
 ['Ladin', 'Language', 'codes', 'Linguasphere', 'Distribution', 'of',  
 'Ladin', 'in', 'Trentino', ',', 'South', 'Tyrol', 'and', 'Belluno', 'in',  
 'northeastern', 'Italy']  
 ['Ukrainian', 'Language', 'codes', 'Linguasphere', '53-AAA-ed', '&', 'lt',  
 ';', '53-AAA-e']  
 ['Ukrainian', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':',  
 '53-AAA-eda', 'to', '53-AAA-edq', ')']  
 ['Ukrainian', 'Language', 'codes', 'Linguasphere', 'Spread', 'of',

'Ukrainian', 'language', 'in', 'the', 'first', 'half', 'of', '20th', 'century']

['International', 'Language', 'codes', 'ISO', '639-3', 'ils', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['British', 'Language', 'codes', 'ISO', '639-3', 'bfi', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Thai', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Duala', 'Total', 'speakers', '90', ',', '000', '(', '1982', ')']

['Akkadian', 'Language', 'extinction', '100', 'AD']

['Akkadian', 'Official', 'language', 'in', 'initially', 'Akkad', '(', 'central', 'Mesopotamia', ')', ';', 'lingua', 'franca', 'of', 'the', 'Middle', 'East', 'and', 'Egypt', 'in', 'the', 'late', 'Bronze', 'and', 'early', 'Iron', 'Ages']

['Middle', 'English', 'Language', 'extinction', 'developed', 'into', 'Early', 'Modern', 'English', ',', 'Scots', 'and', 'Yola', 'in', 'Wexford', 'by', 'the', '16th', 'century']

['Middle', 'Spoken', 'in', 'south', 'east', 'Scotland', 'and', 'in', 'Scottish', 'burghs']

['M\xc4\x81ori', 'Total', 'speakers', '157', ',', '110', 'New', 'Zealand', 'residents', 'claimed', 'they', 'could', 'converse', 'in', 'M', '\xc4', '\x81', 'ori', 'about', 'everyday', 'things', ',', 'in', 'the', '2006', 'Census']

['M\xc4\x81ori', 'Regulated', 'by', 'M', '\xc4', '\x81', 'ori', 'Language', 'Commission']

['Vulgar', 'Latin', 'Language', 'extinction', 'developed', 'into', 'Romance', 'languages', '6th', 'to', '9th', 'centuries']

['Cree', 'Total', 'speakers', '117', ',', '410', '(', 'including', 'Montagnais-Naskapi', 'and', 'Atikamekw', ')']

['Cree', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Cree', 'syllabics', '(', 'variation', 'of', 'Canadian', 'Aboriginal', 'syllabics', ')']

['Cree', 'Language', 'codes', 'ISO', '639-3', 'crm', '\xe2', '\x80', '\x93', 'Moose', 'Cree']

['Cree', 'Language', 'codes', 'ISO', '639-3', 'crl', '\xe2', '\x80', '\x93', 'Northern', 'East', 'Cree']

['Cree', 'Language', 'codes', 'ISO', '639-3', 'crj', '\xe2', '\x80', '\x93', 'Southern', 'East', 'Cree']

['Cree', 'Language', 'codes', 'ISO', '639-3', 'crw', '\xe2', '\x80', '\x93', 'Swampy', 'Cree']

['Cree', 'Language', 'codes', 'ISO', '639-3', 'cwd', '\xe2', '\x80', '\x93', 'Woods', 'Cree']

['Cree', 'Language', 'codes', 'ISO', '639-3', 'crk', '\xe2', '\x80', '\x93', 'Plains', 'Cree']

['Tetum', 'Total', 'speakers', '800', ',', '000']

['Igbo', 'Region', 'West', 'and', 'West', 'Central', 'Africa']

['Igbo', 'Writing', 'system', 'Latin', 'alphabet', '(', '\xc3', '\x96', 'nwu', 'script', ')', ',', 'Nsibidi']

['Igbo', 'Language', 'codes', 'ISO', '639-3', 'ibo', 'Lingusitic', 'map', 'of', 'Nigerian', 'and', 'Cameroon', 'with', 'Igbo', 'bottom', 'center']

['Hakka', 'Region', 'in', 'China', ':', 'Eastern', 'Guangdong', 'province', ';', 'adjoining', 'regions', 'of', 'Fujian', 'and', 'Jiangxi', 'provinces']

['Hakka', 'Spoken', 'in', 'Japan', '(', 'due', 'to', 'presence', 'of', 'Taiwanese', 'community', 'in', 'Tokyo-Yokohama', 'Metropolitan', 'Area', ')']

['Hakka', 'Spoken', 'in', 'India', 'and', 'other', 'countries', 'where', 'Hakka', 'Chinese', 'migrants', 'have', 'settled']

['Hakka', 'Official', 'language', 'in', 'none', '(', 'legislative', 'bills', 'have', 'been', 'proposed', 'for', 'it', 'to', 'be', 'one', 'of', 'the', '"national", "languages"', 'in', 'the', 'Republic', 'of', 'China', ')', ';', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public', 'transport', 'announcements', 'in', 'the', 'ROC', ';', 'ROC', 'government', 'sponsors', 'Hakka', 'language', 'television', 'station', 'to', 'preserve', 'language']

['Hakka', 'Regulated', 'by', 'The', 'Guangdong', 'Provincial', 'Education', 'Department', 'created', 'an', 'official', 'romanisation', 'of', 'Meixian', 'Hakka', 'dialect', 'in', '1960', ',', 'one', 'of', 'four', 'languages', 'receiving', 'this', 'status', 'in', 'Guangdong', '.', 'It', 'is', 'called', 'Kejiahua', 'Pinyin', 'Fang'an']

['Hakka', 'Language', 'codes', 'ISO', '639-1', 'zh', '(', 'Chinese', ')']

['Nyabwa', 'Total', 'speakers', '43', ',', '000']

['Quebec', 'Total', 'speakers', '(', 'mother', 'tongue', ')', '6', 'million', 'in', 'Quebec', ',', '700', ',', '000', 'elsewhere', 'in', 'Canada']

['Quebec', 'Official', 'language', 'in', 'Quebec', '(', 'as', 'French', ')']

['Quebec', 'Language', 'codes', 'Linguasphere', '51-AAA-hq', '&', 'amp', ';', '51-AAA-ii']

['Dida', 'Total', 'speakers', '200', ',', '000', '(', '1993', ')']

['Dida', 'Language', 'codes', 'ISO', '639-3', 'dic', '\xe2', '\x80', '\x93', 'Lakota', 'Dida']

['Taiwanese', 'Hokkien', 'Total', 'speakers', 'About', '15', 'million', 'in', 'Taiwan', ';', '49', 'million', '(', 'Min', 'Nan', 'as', 'a', 'group', ')']

['Taiwanese', 'Writing', 'system', 'Latin', '(', 'pe', '\xcc', '\x8d', 'h-', '\xc5', '\x8d', 'e-j', '\xc4', '\xab', ')', ',', 'Hanzi']

['Syriac', 'Language', 'extinction', 'marginalized', 'by', 'Arabic', 'from', 'the', '8th', 'century', ',', 'evolved', 'into', 'Neo-Aramaic',

'vernaculars', 'around', 'AD', '1200', '(', 'conventional', 'date', ')']  
 ['Syriac', 'Language', 'codes', 'ISO', '639-3', 'syc', '\xe2', '\x80',  
 '\x93', 'Syriac', '(', 'classical', ')']  
 ['Balochi', 'Total', 'speakers', '7', '\xe2', '\x80', '\x93', '8',  
 'million', '(', '1998', ',', 'Ethnologue', ')']  
 ['Balochi', 'Official', 'language', 'in', 'Pakistan', '(', 'Balochistan',  
 ')', '(', 'Regional', 'Language', ')']  
 ['Balochi', 'Regulated', 'by', 'Balochi', 'Academy', '(', 'Pakistan', ')']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', ':', 'bgp', '\xe2',  
 '\x80', '\x93', 'Eastern', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bgn', '\xe2', '\x80',  
 '\x93', 'Western', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bcc', '\xe2', '\x80',  
 '\x93', 'Southern', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'Linguasphere', '58-AAB-a', '&', 'gt',  
 ';', '58-AAB-aa', '(', 'East', 'Balochi', ')', '+', '58-AAB-ab', '(',  
 'West', 'Balochi', ')', '+', '58-AAB-ac', '(', 'South', 'Balochi', ')',  
 '+', '58-AAB-ad', '(', 'Bashkardi', ')']  
 ['Balochi', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Zulu', 'Total', 'speakers', 'First', 'language', '-', '10', 'million']  
 ['Zulu', 'Total', 'speakers', 'Second', 'language', '-', '16', 'million']  
 ['Zulu', 'Regulated', 'by', 'Pan', 'South', 'African', 'Language', 'Board']  
 ['Nepali', 'Official', 'language', 'in', 'West', 'Bengal', '(', 'India',  
 ')']  
 ['Nepali', 'Regulated', 'by', 'Language', 'Academy', 'of', 'Nepal']  
 ['Nepali', 'Language', 'codes', 'ISO', '639-3', 'World', 'map', 'with',  
 'significant', 'Nepali', 'language', 'speakers', 'Dark', 'Blue', ':',  
 'Main', 'official', 'language']  
 ['Nepali', 'Language', 'codes', 'ISO', '639-3', 'Light', 'blue', ':',  
 'One', 'of', 'the', 'official', 'languages']  
 ['Nepali', 'Language', 'codes', 'ISO', '639-3', 'Red', ':', 'Places',  
 'with', 'significant', 'population', 'or', 'greater', 'than', '20', '%',  
 'but', 'without', 'official', 'recognition', '.', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Scots', 'Total', 'speakers', 'est', '.', '200', ',', '000', '(',  
 'ethnologue', ')', 'to', 'over', '1', '.', '5', 'million', '(', 'General',  
 'Register', 'Office', 'for', 'Scotland', ',', '1996', ')']  
 ['Scots', 'Total', 'speakers', 'A', 'Scottish', 'Government', 'study',  
 'found', 'that', '85', '%', 'of', 'respondents', '(', 'being', 'a',  
 'representative', 'sample', 'of', 'Scotland's', 'adult', 'population', ')',  
 'claim', 'to', 'speak', 'Scots', 'to', 'varying', 'degrees']

['Scots', 'Region', 'Scotland', ':', 'Scottish', 'Lowlands']

['Scots', 'Official', 'language', 'in', '\xe2', '\x80', '\x94', 'Classified', 'as', 'a', '"', 'traditional', 'language', '"', 'by', 'the', 'Scottish', 'Government']

['Scots', 'Official', 'language', 'in', '\xe2', '\x80', '\x94', 'Classified', 'as', 'a', '"', 'regional', 'or', 'minority', 'language', '"', 'under', 'the', 'European', 'Charter', 'for', 'Regional', 'or', 'Minority', 'Languages', ',', 'ratified', 'by', 'the', 'United', 'Kingdom', 'in', '2001']

['Scots', 'Official', 'language', 'in', '\xe2', '\x80', '\x94', 'Classified', 'as', 'a', '"', 'traditional', 'language', '"', 'by', 'The', 'North', '/', 'South', 'Language', 'Body']

['Scots', 'Regulated', 'by', '\xe2', '\x80', '\x94', 'Scotland', ':', 'None', ',', 'although', 'the', 'Dictionary', 'of', 'the', 'Scots', 'Language', 'carries', 'great', 'authority', '(', 'the', 'Scottish', 'Government's', 'Partnership', 'for', 'a', 'Better', 'Scotland', 'coalition', 'agreement', '(', '2003', ')', 'promises', '"', 'support', '"', ')']

['Scots', 'Language', 'codes', 'Linguasphere', '52-ABA-aa', '(', 'varieties']

['Scots', 'Language', 'codes', 'Linguasphere', ':', '52-ABA-aaa', 'to', '-aav', ')']

['Sumerian', 'Language', 'extinction', 'Effectively', 'extinct', 'from', 'about', '1800', 'BC', ';', 'used', 'as', 'classical', 'language', 'until', 'about', 'AD', '100']

['Hindi-Urdu', 'Total', 'speakers', 'Native', ':', '570', 'million', '(', '2011', ')']

['Hindi-Urdu', 'Total', 'speakers', 'Second', 'language', ':', '490', 'million', '(', '2011', ')']

['Hindi-Urdu', 'Total', 'speakers', 'Total', ':', '970', 'million', '(', '2011', ')']

['Hindi-Urdu', 'Spoken', 'in', 'India', 'and', 'Pakistan', '.', 'Also', 'various', 'other', 'countries', 'due', 'to', 'South', 'Asian', 'immigration']

['Hindi-Urdu', 'Standard', 'forms', 'Standard', 'Hindi']

['Hindi-Urdu', 'Standard', 'forms', 'Standard', 'Urdu']

['Hindi-Urdu', 'Official', 'language', 'in', 'India', '(', 'as', 'Hindi', 'and', 'Urdu', ')']

['Hindi-Urdu', 'Official', 'language', 'in', 'Pakistan', '(', 'as', 'Urdu', ')']

['Hindi-Urdu', 'Regulated', 'by', 'Central', 'Hindi', 'Directorate', '(', 'Hindi', ',', 'India', ')', ',']

['Hindi-Urdu', 'Regulated', 'by', 'National', 'Language', 'Authority', ',', '(', 'Urdu', ',', 'Pakistan', ')', ';']

['Hindi-Urdu', 'Regulated', 'by', 'National', 'Council', 'for', 'Promotion', 'of', 'Urdu', 'language', '(', 'Urdu', ',', 'India', ')']

["Ge'ez", 'Language', 'extinction', 'Extinct', '.', 'Ceased', 'to', 'be',  
 'a', 'spoken', 'tongue', '(', 'in', '4th', 'century', 'CE', 'according',  
 'to', ')', '(', 'sometime', 'before', 'the', '10th', 'century', 'CE',  
 'according', 'to', ')', ',', 'remains', 'in', 'use', 'as', 'a',  
 'liturgical', 'language']  
 ['Ainu', 'Total', 'speakers', '~', '100s']  
 ['Kazakh', 'Total', 'speakers', '8-8', '.', '3', 'million']  
 ['Uzbek', 'Total', 'speakers', '23', '.', '5', 'million']  
 ['Uzbek', 'Language', 'codes', 'ISO', '639-3', ':', 'uzn', '\xe2', '\x80',  
 '\x93', 'Northern', 'Uzbek']  
 ['Uzbek', 'Language', 'codes', 'ISO', '639-3', 'uzs', '\xe2', '\x80',  
 '\x93', 'Southern', 'Uzbek']  
 ['Tok', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Tok', 'Official', 'language', 'in', 'Papua', 'New', 'Guinea']  
 ['Serbian', 'Regulated', 'by', 'Board', 'for', 'Standardization', 'of',  
 'the', 'Serbian', 'Language']  
 ['Serbian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-g']  
 ['(Paraguayan)', 'Guaran', '\xc3', '\xad', 'Total', 'speakers', '4', ',',  
 '650', ',', '000', 'in', 'Paraguay', '(', 'SIL', '1995', ')', ',', '200',  
 ',', '000', 'in', 'Argentina']  
 ['(Paraguayan)', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Guaran',  
 '\xc3', '\xad', 'variant', ')']  
 ['Lingala', 'Region', 'Central', 'and', 'Eastern', 'Africa']  
 ['Lingala', 'Writing', 'system', 'African', 'reference', 'alphabet', '(',  
 'Latin', 'alphabet', ')', ',', 'Mandombe']  
 ['Lingala', 'Language', 'codes', 'Linguasphere', 'Geographic',  
 'distribution', 'of', 'Lingala', 'speakers', ',', 'showing', 'regions',  
 'of', 'native', 'speakers', '(', 'dark', 'green', ')', 'and', 'other',  
 'regions', 'of', 'use']  
 ['Lithuanian', 'Total', 'speakers', '3', '.', '5', 'million', '(',  
 'Lithuania', ')']  
 ['Lithuanian', 'Regulated', 'by', 'Commission', 'of', 'the', 'Lithuanian',  
 'Language']  
 ['Astur-Leonese', 'Total', 'speakers', '300', ',', '000-450', ',', '000']  
 ['Astur-Leonese', 'Official', 'language', 'in', 'In', 'Portugal', ':', 'Co-  
 official', '(', 'as', 'Mirandese', ')', 'in', 'Miranda', 'do', 'Douro']  
 ['Astur-Leonese', 'Official', 'language', 'in', 'In', 'Spain', ':',  
 'Given', 'special', 'protection', '(', 'as', 'Leonese', ')', 'in',  
 'Castile', 'and', 'Leon', ',', 'and', '(', 'as', 'Asturian', ')', 'in',  
 'Asturias']  
 ['Astur-Leonese', 'Regulated', 'by', 'Academy', 'of', 'the', 'Asturian',  
 'Language', '(', 'Asturian', ')', ',', 'Institute', 'of', 'Mirandese',  
 'Language', '(', 'Mirandese', ')']  
 ['Astur-Leonese', 'Language', 'codes', 'Linguasphere', '51-AAA-c', 'Astur-  
 Leonese', 'area']  
 ['Alsatian', 'Total', 'speakers', 'more', 'than', '700', ',', '000']

['Kyrgyz', 'Total', 'speakers', 'Approx', '.', '4', '.', '5', 'million']  
 ['Kyrgyz', 'Writing', 'system', 'Cyrillic', 'alphabet', '(', 'Kyrgyz',  
 'variant', ')', ';', 'Arabic', 'alphabet', '(', 'Kyrgyz', 'variant', ')']  
 ['Ganda', 'Total', 'speakers', 'First', 'language', ':', '6', 'million',  
 '(', '2008', ')']  
 ['Ganda', 'Total', 'speakers', 'Second', 'language', ':', '4', 'million',  
 '(', '2008', ')']  
 ['Sabine', 'Language', 'extinction', 'Only', 'trace', 'the', 'vocabulary',  
 'of', 'mainly', 'Marcus', 'Terentius', 'Varro', ',', '1st', 'century',  
 'BC']  
 ['Tigrinya', 'Total', 'speakers', '6', '.', '7', 'million']  
 ['Wolof', 'Total', 'speakers', '5', '.', '5', 'million', '(', 'mother',  
 'tongue', ')']  
 ['Wolof', 'Total', 'speakers', '12', '.', '3', 'million', '(', 'second',  
 'language', ')']  
 ['Xhosa', 'Total', 'speakers', '7', '.', '9', 'million']  
 ['Xhosa', 'Region', 'Eastern', 'Cape', ',', 'Western', 'Cape']  
 ['Oscan', 'Language', 'extinction', 'Latest', 'inscriptions', '1st',  
 'century', 'BC']  
 ['Oscan', 'Language', 'codes', 'ISO', '639-3', 'Approximate',  
 'distribution', 'of', 'languages', 'in', 'Iron', 'Age', 'Italy', 'during',  
 'the', '6th', 'century', 'BCE']  
 ['Umbrian', 'Language', 'extinction', 'Latest', 'inscriptions', '1st',  
 'century', 'BC']  
 ['Umbrian', 'Language', 'codes', 'ISO', '639-3', 'xum', 'Approximate',  
 'distribution', 'of', 'languages', 'in', 'Iron', 'Age', 'Italy', 'during',  
 'the', '6th', 'century', 'BC']  
 ['Ribagorçà', 'Dialects', 'Catalan', 'dialect', 'transition']  
 ['Classical', 'Latin', 'Spoken', 'in', 'Roman', 'republic', ',', 'Roman',  
 'empire']  
 ['Classical', 'Language', 'extinction', 'developed', 'into', 'Late',  
 'Latin']  
 ['Classical', 'Language', 'codes', 'ISO', '639-3', 'The', 'range', 'of',  
 'Latin', ',', '60', 'AD']  
 ['Pashto', 'Language', 'codes', 'ISO', '639-3', ':', 'pst', '\xe2', '\x80',  
 '\x93', 'Central', 'Pashto']  
 ['Pashto', 'Language', 'codes', 'ISO', '639-3', 'pbu', '\xe2', '\x80',  
 '\x93', 'Northern', 'Pashto']  
 ['Pashto', 'Language', 'codes', 'ISO', '639-3', 'pbt', '\xe2', '\x80',  
 '\x93', 'Southern', 'Pashto']  
 ['Kanuri', 'Total', 'speakers', '4', 'million', '(', '1987', 'figures',  
 ')']  
 ['Kanuri', 'Language', 'codes', 'ISO', '639-3', ':', 'knc', '\xe2', '\x80',  
 '\x93', 'Central', 'Kanuri']  
 ['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'Map', 'of', 'the',  
 'majority', 'usage', 'of', 'the', 'five', 'major', 'languages', 'of',

'the', 'Kanuri', 'language', 'group', '.', '\*', 'BMS', 'Kanuri', ',', 'Bilma']

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'KNC', 'Kanuri', ',', 'Central']

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'KNY', 'Kanuri', ',', 'Manga']

['Kanuri', 'Language', 'codes', 'ISO', '639-3', 'KRT', 'Kanuri', ',', 'Tumari']

['Malay', 'Spoken', 'in', 'Malaysia', '(', 'as', 'Malaysian', ')']

['Malay', 'Spoken', 'in', 'Indonesia', '(', 'as', 'Indonesian', ')']

['Malay', 'Spoken', 'in', 'East', 'Timor', '(', 'as', 'Indonesian', ')']

['Malay', 'Total', 'speakers', '40', 'million', 'native', '(', 'incl', '.', '23', 'million', 'Indonesian', ')', ';', 'approx', '.', '180', 'million', 'total', '(', '90', '%', 'Indonesian', ')']

['Malay', 'Writing', 'system', 'Rumi', '(', 'Latin', 'alphabet', ')', ',', 'Malay', 'Variant', '(', 'official', 'in', 'Malaysia', ',', 'Singapore', 'and', 'Indonesia', ';', 'co-official', 'in', 'Brunei', ')']

['Malay', 'Writing', 'system', 'Jawi', '(', 'Arabic', 'script', ')', '(', 'co-official', 'in', 'Brunei', 'and', 'Malaysia', ')']

['Malay', 'Writing', 'system', 'Historically', 'written', 'in', 'Pallava', ',', 'Kawi', 'and', 'Rencong']

['Malay', 'Regulated', 'by', 'Dewan', 'Bahasa', 'dan', 'Pustaka', '(', 'Institute', 'of', 'Language', 'and', 'Literature', ';']

['Malay', 'Regulated', 'by', 'Majlis', 'Bahasa', 'Brunei', '-', 'Indonesia', '-', 'Malaysia', '(', 'Brunei', '-', 'Indonesia', '-', 'Malaysia', 'Language', 'Council', '\xe2', '\x80', '\x94', 'MABBIM', ')', '(', 'a', 'trilateral', 'joint-venture', ')']

['Malay', 'Recognised', 'minority', 'language', 'in', 'Indonesia', '(', 'Malay', 'enjoys', 'a', 'status', 'as', 'a', 'regional', 'language', 'in', 'Indonesia', ',', 'separated', 'from', 'Indonesian', 'which', 'receives', 'a', 'national', 'language', 'status', ')']

['Malay', 'Recognised', 'minority', 'language', 'in', 'Thailand', '(', 'Malay', 'is', 'the', 'language', 'of', 'the', 'Muslim', 'community', 'in', 'Southern', 'Thailand', ')']

['Malay', 'Language', 'codes', 'ISO', '639-3', 'mhp', '\xe2', '\x80', '\x93', 'Balinese', 'Malay']

['Malay', 'Language', 'codes', 'ISO', '639-3', 'lrt', '\xe2', '\x80', '\x93', 'Larantuka', 'Malay']

['Malay', 'Language', 'codes', 'ISO', '639-3', 'mfp', '\xe2', '\x80', '\x93', 'Makassar', 'Malay']

['Malay', 'Language', 'codes', 'ISO', '639-3', 'max', '\xe2', '\x80', '\x93', 'North', 'Moluccan', 'Malay']

['Valencian', 'Total', 'speakers', '2', '.', '4', 'million', '(', '2004', ')']

['Valencian', 'Total', 'speakers', '(', '11', '.', '5', 'milion', 'in', 'the', 'whole', 'dominion', ')']



['Valencian', 'Official', 'language', 'in', ':', 'Valencian', 'Community']  
 ['Tutchone', 'Total', 'speakers', '400', 'in', '1995']  
 ['Tutchone', 'Region', '1400', 'Southern', 'Tutchone', 'people', ', ',  
 '1000', 'Northern', 'Tutchone', 'people']  
 ['Lao', 'Total', 'speakers', '5', ', ', '225', ', ', '552', '(', '2006', ')',  
 ', ', 'roughly', '20', 'million', 'if', 'Isan', 'speakers', 'are',  
 'included']  
 ['Lao', 'Spoken', 'in', 'U', '. ', 'S']  
 ['Gujarati', 'Total', 'speakers', '46', '. ', '1', 'million']  
 ['Gujarati', 'Spoken', 'in', 'U', '. ', 'S']  
 ['Gujarati', 'Writing', 'system', 'former', 'use', 'of', 'Devanagari',  
 'before', 'invention', 'of', 'Gujarati', 'Script']  
 ['Gujarati', 'Writing', 'system', 'also', 'use', 'of', 'Arabic', 'script',  
 'by', 'the', 'Ismaili', 'community', 'and', 'other', 'Gujarati',  
 'communities']  
 ['Gujarati', 'Official', 'language', 'in', 'Daman', 'and', 'Diu', '(',  
 'India', ')']  
 ['Gujarati', 'Language', 'codes', 'ISO', '639-3', 'Countries', 'where',  
 'Gujarati', 'is', 'spoken', '. ', 'This', 'page', 'contains', 'Indic',  
 'text', '. ', 'Without', 'rendering', 'support', 'you', 'may', 'see',  
 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts',  
 '. ', 'More']  
 ['Southern', 'Ndebele', 'Total', 'speakers', '586', ', ', '961', '(',  
 '1996', 'census', ')']  
 ['Dalmatian', 'Language', 'extinction', '10', 'June', '1898', ', ', 'when',  
 'Tuone', 'Udaina', 'was', 'killed']  
 ['Vandalic', 'Language', 'extinction', '6th', 'century', 'AD']  
 ['Vandalic', 'Spoken', 'in', 'Spain', ', ', 'North', 'Africa']  
 ['Romani', 'Spoken', 'in', 'Central', 'and', 'Eastern', 'Europe', ', ',  
 'worldwide']  
 ['Romani', 'Total', 'speakers', '2', '. ', '5', 'million', '(', 'SIL',  
 'estimate', ')']  
 ['Romani', 'Official', 'language', 'in', '(', 'not', 'official',  
 'language', ')', 'recognised', 'as', 'minority', 'language', 'only', 'in',  
 'parts', 'of']  
 ['Romani', 'Language', 'codes', 'ISO', '639-3', ':', 'rmn', '\xe2', '\x80',  
 '\x93', 'Balkan', 'Romani']  
 ['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmf', '\xe2', '\x80',  
 '\x93', 'Finnish', 'Kalo']  
 ['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmy', '\xe2', '\x80',  
 '\x93', 'Vlax', 'Romani']  
 ['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmw', '\xe2', '\x80',  
 '\x93', 'Welsh', 'Romani']  
 ['Yue', 'Total', 'speakers', '56', 'million', 'in', '1984']  
 ['Yue', 'Region', 'in', 'China', ':', 'the', 'Pearl', 'River', 'Delta',  
 '(', 'central', 'Guangdong', ';', 'Hong', 'Kong', ', ', 'Macau', ')', ';',

'the', 'eastern', 'and', 'southern', 'Guangxi', ';', 'parts', 'of',  
'Hainan', ';']

['Yue', 'Spoken', 'in', 'China', ',', 'Malaysia', '(', 'Kuala', 'Lumpur',  
',', 'Ipoh', ',', 'Kota', 'Kinabalu', ')', ',', 'Vietnam', ',',  
'Indonesia', '(', 'Medan', ')', ',', 'United', 'States', ',', 'Canada',  
',', 'United', 'Kingdom', 'and', 'other', 'countries', 'where',  
'Cantonese', 'migrants', 'have', 'settled']

['Yue', 'Official', 'language', 'in', 'Hong', 'Kong', 'and', 'Macau', '(',  
'de', 'facto', ',', 'though', 'officially', 'referred', 'to', 'as', '"',  
'Chinese', '"', ';', 'Cantonese', 'and', 'occasionally', 'Mandarin', 'are',  
'used', 'in', 'government', ')', '.', 'Recognised', 'regional', 'language',  
'in', 'Suriname']

['Siwi', 'Total', 'speakers', 'about', '15', ',', '000']

['Zenaga', 'Total', 'speakers', 'about', '200']

['Tuareg', 'language', '(', 's', ')', 'Total', 'speakers', '1', '.', '2',  
'million', '(', 'Ethnologue', ')']

['Ancient', 'Greek', 'Language', 'extinction', 'developed', 'into', 'Koin',  
'\xc3', '\xa9', 'Greek', 'by', 'the', '4th', 'century', 'BC']

['Black', 'Category', '(', 'purpose', ')', 'Black', 'Speech']

['Black', 'Date', 'founded', 'c', '.', '1945', '\xe2', '\x80', '\x93',  
'1973']

['Polabian', 'Language', 'extinction', '18th', 'century']

['Polabian', 'Language', 'codes', 'Linguasphere', 'Former', 'settlement',  
'area', 'of', 'the', 'Polabian', 'Slavs']

['Slovincian', 'Language', 'extinction', '20th', 'century']

['Rusyn', 'Total', 'speakers', 'Estimated', ':', 'At', 'least', '600', ',', '000']

['Rusyn', 'Total', 'speakers', 'Census', 'population', ':', '50', ',', '000', '.', 'These', 'are', 'numbers', 'from', 'national', 'official',  
'bureaus', 'for', 'statistics']

['Rusyn', 'Total', 'speakers', ':', 'Slovakia', '-', '24', ',', '201']

['Rusyn', 'Total', 'speakers', 'Serbia', '-', '15', ',', '626']

['Rusyn', 'Total', 'speakers', 'Ukraine', '-', '6', ',', '725']

['Rusyn', 'Total', 'speakers', 'Croatia', '-', '2', ',', '337']

['Rusyn', 'Total', 'speakers', 'Poland', '-', '5', ',', '800']

['Rusyn', 'Total', 'speakers', 'Hungary', '-', '1', ',', '098']

['Rusyn', 'Total', 'speakers', 'Czech', 'Republic', '-', '1', ',', '106']

['Rusyn', 'Language', 'codes', 'Linguasphere', '53-AAA-ec', '&', 'lt', ';',  
'53-AAA-e']

['Rusyn', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-  
AAA-eca', 'to', '53-AAA-ecc', ')']

['Romano-Serbian', 'Total', 'speakers', '172', ',', '000', '(', 'in',  
'2006', ')']

['Manchu', 'Total', 'speakers', '60', '(', '1999', 'Zhao', 'Aping', ')']

['Cimbrian', 'Total', 'speakers', '2', ',', '230', '&', 'gt', ';']

['Austro-Bavarian', 'Spoken', 'in', 'Italy', ',', 'South', 'Tyrol']

['Galindan', 'Galindian', 'Language', 'extinction', '14th', 'century']  
 ['Galindan', 'Spoken', 'in', 'Northern', 'Poland', 'and', 'Russia']  
 ['Mongolian', 'Total', 'speakers', '5', '.', '2', 'million']  
 ['Mongolian', 'Spoken', 'in', 'Mongolia', 'People's', 'Republic', 'of', 'China']  
 ['Mongolian', 'Writing', 'system', 'Mongolian', 'script', 'Mongolian', 'Cyrillic', 'alphabet']  
 ['Mongolian', 'Official', 'language', 'in', 'People's', 'Republic', 'of', 'China', '(', 'Inner', 'Mongolia', ')']  
 ['Mongolian', 'Regulated', 'by', 'State', 'Language', 'Council', '(', 'Mongolia', ')', 'Council', 'for', 'Language', 'and', 'Literature', 'Work', '(', 'Inner', 'Mongolia', ')']  
 ['Mongolian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-BAA-b']  
 ['Marsian', 'Language', 'extinction', 'ca', '.', '150', 'BC']  
 ['Rotokas', 'Total', 'speakers', '4', '320']  
 ['Rotokas', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Luba', 'Total', 'speakers', '9', '500', '000', '(', '2009', ')']  
 ['Burushaski', 'Total', 'speakers', '87', '000', '(', '2000', ')']  
 ['Acadian', 'Total', 'speakers', '371', '614', '(', '1996', ')']  
 ['Acadian', 'Total', 'speakers', '371', '590', '(', '2006', ')', 'Canadian', 'census', 'ethnic', 'data']  
 ['Norwegian', 'Spoken', 'in', 'Norway', '(', '4', '.', '8', 'million', ')', ',']  
 ['Norwegian', 'Spoken', 'in', 'United', 'States', '(', '55', '311', ')']  
 ['Norwegian', 'Spoken', 'in', 'Canada', '(', '7', '710', ')']  
 ['Norwegian', 'Standard', 'forms', 'Nynorsk', '(', 'official', ')']  
 ['Norwegian', 'Standard', 'forms', 'Bokm', '\xc3', '\xa5', 'l', '(', 'official', ')', '/', 'Riksm', '\xc3', '\xa5', 'l', '(', 'unofficial', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Language', 'Council', '(', 'Bokm', '\xc3', '\xa5', 'l', 'and', 'Nynorsk', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Academy', '(', 'Riksm', '\xc3', '\xa5', 'l', ')']  
 ['Dungan', 'Total', 'speakers', '41', '400', '(', '2001', ')']  
 ['Dungan', 'Region', 'Fergana', 'Valley', 'Chu', 'Valley']  
 ['Norwegian', 'Spoken', 'in', 'Norway', '(', '4', '.', '8', 'million', ')', ',']  
 ['Norwegian', 'Spoken', 'in', 'United', 'States', '(', '55', '311', ')']  
 ['Norwegian', 'Spoken', 'in', 'Canada', '(', '7', '710', ')']  
 ['Norwegian', 'Standard', 'forms', 'Nynorsk', '(', 'official', ')']  
 ['Norwegian', 'Standard', 'forms', 'Bokm', '\xc3', '\xa5', 'l', '(', 'official', ')', '/', 'Riksm', '\xc3', '\xa5', 'l', '(', 'unofficial', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Language', 'Council', '(',

'Bokm', '\xc3', '\xa5', 'l', 'and', 'Nynorsk', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Academy', '(', 'Riksm',  
 '\xc3', '\xa5', 'l', ')']  
 ['Brandenburgisch', 'Language', 'codes', 'ISO', '639-3', 'Low', 'Saxon',  
 'and', 'Low', 'Franconian', 'dialects']  
 ['Haz\xc4\x81rag\xc4\xab', 'Total', 'speakers', '1', '.', '8', 'to',  
 'over', '2', '.', '2', 'million']  
 ['Avestan', 'Writing', 'system', 'Avestan', 'alphabet', '(', 'independent',  
 'ad-hoc', 'development', ')', ',', 'Gujarati', 'script', 'used', 'by',  
 'the', 'Indian', 'Zoroastrian', 'community']  
 ['Avestan', 'Language', 'codes', 'Linguasphere', 'Yasna', '28', '.', '1',  
 ',', 'Ahunavaiti', 'Gatha', '(', 'Bodleian', 'MS', 'J2', ')']  
 ['Navajo', 'Total', 'speakers', '170', ',', '717', '(', '2007', ')']  
 ['Bosnian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-g']  
 ['Eastphalian', 'Region', 'Lower', 'Saxony', ',', 'Saxony-Anhalt']  
 ['Eastphalian', 'Language', 'codes', 'ISO', '639-3', 'Eastphalian', '(',  
 'No', '.', '7', ')', 'within', 'the', 'Low', 'German', '(', 'yellow', ')',  
 'and', 'Low', 'Franconian', '(', 'red', ')', 'language', 'area']  
 ['Okinawan', 'Total', 'speakers', '984', ',', '000', '(', '2000', ')']  
 ['Okinawan', 'Writing', 'system', 'R', '\xc5', '\x8d', 'maji']  
 ['Texas', 'German', 'Total', 'speakers', '~', '3', ',', '000', ',',  
 'declining']  
 ['Hutterite', 'Spoken', 'in', 'North', 'and', 'South', 'Dakota']  
 ['Hutterite', 'Total', 'speakers', '34', ',', '000', '(', '2003',  
 'estimate', ')']  
 ['Pennsylvania', 'German', ',', 'Pennsylvania', 'Dutch', 'Total',  
 'speakers', '250', '000', '+']  
 ['Pennsylvania', 'Region', 'Pennsylvania', ';;', 'Ohio', ';;', 'Indiana',  
 ';;', 'Ontario', ';;', 'and', 'elsewhere']  
 ['Pennsylvania', 'Language', 'codes', 'Linguasphere', 'Blue', ':', 'The',  
 'counties', 'with', 'the', 'highest', 'proportion', 'of', 'Pennsylvania',  
 'German', 'speakers']  
 ['Pennsylvania', 'Language', 'codes', 'Linguasphere', 'Red', ':', 'The',  
 'counties', 'with', 'the', 'highest', 'number', 'of', 'Pennsylvania',  
 'German', 'speakers']  
 ['Plautdietsch', 'Total', 'speakers', '260', ',', '710', '\xe2', '\x80',  
 '\x93', '318', ',', '500']  
 ['Mohawk', 'Region', 'Ontario', ',', 'Quebec', 'and', 'northern', 'New',  
 'York']  
 ['Quechua', 'Language', 'codes', 'Linguasphere', '85-', '(', 'phylozone',  
 ')']  
 ['Chinook', 'Jargon', 'Total', 'speakers', 'Unknown', '.', 'Definitely',  
 'more', 'than', '10', '-', 'Also', 'a', 'small', ',', 'yet', 'present',  
 'revival', 'effort', 'in', 'place']  
 ['Chinook', 'Region', 'Pacific', 'Northwest', '(', 'Interior', 'and',  
 'Coast', ')']

['Chinook', 'Writing', 'system', 'de', 'facto', 'Latin', 'Alphabet']  
 ['Chinook', 'Official', 'language', 'in', 'De', 'facto', 'in', 'Pacific',  
 'Northwest', 'until', 'about', '1900']  
 ['Chechen', 'Total', 'speakers', 'circa', '1', ',', '500', ',', '000',  
 'self-reported', 'speakers', 'worldwide', '(', '2009', ')']  
 ['East', 'Frisian', 'Low', 'Saxon', 'Total', 'speakers', '230', ',', '000',  
 'in', 'Germany']  
 ['Taiwanese', 'Language', 'codes', 'ISO', '639-3', 'tss', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Gullah', 'Total', 'speakers', '250', ',', '000']  
 ['Gullah', 'Region', 'Coastal', 'low', 'country', 'region', 'of', 'South',  
 'Carolina', 'and', 'Georgia', 'including', 'the', 'Sea', 'Islands']  
 ['Arvanitika', 'Total', 'speakers', '30', ',', '000', '-', '150', ',', '000']  
 ['Arvanitika', 'Region', 'Saronic', 'Islands', ';;', 'Western', 'Thrace',  
 ';;', 'Peloponnese', ';;', 'some', 'villages', 'in', 'NW', 'of', 'Greece',  
 ';;', 'N', 'of', 'island', 'of', 'Andros', ';;', 'more', 'than', '500',  
 'villages', 'in', 'total']  
 ['Arvanitika', 'Writing', 'system', 'Greek', 'alphabet', '(', 'Arvanitic',  
 'variant', ')']  
 ['Arvanitika', 'Language', 'codes', 'Linguasphere', '55-AAA-aea', 'to',  
 '55-AAA-aed']  
 ['Modern', 'English', 'Total', 'speakers', 'First', 'language', ':', '309',  
 '\xe2', '\x80', '\x93', '380', 'million']  
 ['Modern', 'Total', 'speakers', 'Overall', ':', '\xe2', '\x89', '\x88',  
 '1', '.', '8', 'billion']  
 ['Modern', 'Pronunciation', '/', '\xcb', '\x88', '\xc9', '\xaa', '\xc5',  
 '\x8b', '\xc9', '\xa1', 'l', '\xc9', '\xaa', '\xca', '\x83', '/', '1']  
 ['Modern', 'Language', 'codes', 'ISO', '639-3', 'Countries', 'where',  
 'English', 'is', 'a', 'primary', 'language', 'are', 'dark', 'blue', ';;',  
 'countries', 'where', 'it', 'is', 'an', 'official', 'but', 'not', 'a',  
 'primary', 'language', 'are', 'light', 'blue', '.', 'English', 'is',  
 'also', 'one', 'of', 'the', 'official', 'languages', 'of', 'the',  
 'European', 'Union']  
 ['Assamese', 'Total', 'speakers', '13', ',', '079', ',', '696', '(', 'in',  
 '1991', ')']  
 ['Assamese', 'Spoken', 'in', 'India', ',', 'Bhutan', '&', 'amp', ';;',  
 'USA', '(', 'DE', ',', 'NJ', '&', 'amp', ';;', 'NY', ')']  
 ['Assamese', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Spanish', 'Language', 'codes', 'ISO', '639-3', 'Spanish', 'language',  
 'spread', 'in', 'the', 'United', 'States']  
 ['Lydian', 'Language', 'extinction', '1st', 'c', '.', 'BC']

['Lycian', 'Language', 'extinction', '1st', 'century', 'BC']  
 ['Lycian', 'Writing', 'system', 'Lycian', 'script', 'Names', 'and',  
 'inscriptions', ',', 'presumed', 'spoken']  
 ['Tatar', 'Total', 'speakers', '6', ',', '496', ',', '600']  
 ['Tatar', 'Writing', 'system', 'Cyrillic', 'alphabet', ',', 'Latin',  
 'alphabet', 'and', 'Arabic', 'alphabet']  
 ['Tatar', 'Regulated', 'by', 'Institute', 'of', 'Language', ',',  
 'Literature', 'and', 'Arts', 'of', 'the', 'Academy', 'of', 'Sciences',  
 'of', 'the', 'Republic', 'of', 'Tatarstan']  
 ['Azerbaijani', 'Writing', 'system', 'Latin', 'alphabet', 'for', 'North',  
 'Azeri', 'in', 'Azerbaijan', ',', 'Perso-Arabic', 'script', 'for', 'South',  
 'Azeri', 'in', 'Iran']  
 ['Azerbaijani', 'Official', 'language', 'in', 'Azerbaijan', '(', 'North',  
 'Azeri', ')']  
 ['Azerbaijani', 'Official', 'language', 'in', 'Iran', '(', 'South',  
 'Azeri', ')', '-', 'constitutional', 'status', 'as', 'a', 'regional',  
 'language']  
 ['Azerbaijani', 'Official', 'language', 'in', 'Iraq', '(', 'South',  
 'Azeri', ')', '-', 'constitutional', 'status', 'as', 'a', 'regional',  
 'language', '(', 'referred', 'to', 'as', 'Turkmen', 'in', 'the',  
 'constitution', ')']  
 ['Azerbaijani', 'Official', 'language', 'in', 'Russia', '-', 'One', 'of',  
 'the', 'official', 'languages', 'of', 'Dagestan', 'Republic']  
 ['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', ':', 'azj', '\xe2',  
 '\x80', '\x93', 'North', 'Azerbaijani']  
 ['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', 'azb', '\xe2', '\x80',  
 '\x93', 'South', 'Azerbaijani']  
 ['Azerbaijani', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-  
 a']  
 ['Azerbaijani', 'Language', 'codes', 'Linguasphere', 'Map', 'showing',  
 'locations', 'of', 'Azerbaijani', 'speakers', '(', 'until', '1991', ')']  
 ['Shelta', 'Total', 'speakers', '6', ',', '000', 'in', 'Ireland', ';',  
 '86', ',', '000', 'world-wide', 'Irish', 'Travelers']  
 ['Shelta', 'Region', 'Used', 'by', 'some', 'Irish', 'Travellers']  
 ['Southern', 'Region', 'Southern', 'Fujian', 'province', ';', 'the',  
 'Chaozhou-Shantou', '(', 'Chaoshan', ')', 'area', 'and', 'Leizhou',  
 'Peninsula', 'in', 'Guangdong', 'province', ';', 'extreme', 'south', 'of',  
 'Zhejiang', 'province', ';', 'much', 'of', 'Hainan', 'province', '(', 'if',  
 'Hainanese', 'or', 'Qiong', 'Wen', 'is', 'included', ')', ';', 'and',  
 'most', 'of', 'Taiwan', ';']  
 ['Southern', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Southern', 'Spoken', 'in', 'United', 'States', '(', 'New', 'York',  
 'City', ')']  
 ['Southern', 'Spoken', 'in', 'and', 'other', 'areas', 'of', 'Southern',  
 'Min', 'and', 'Hoklo', 'settlement']  
 ['Southern', 'Official', 'language', 'in', 'None', '(', 'Legislative',

'bills', 'have', 'been', 'proposed', 'for', 'Taiwanese', '(', 'Amoy', 'Southern', 'Min', ')', 'to', 'be', 'one', 'of', 'the', '"national", 'languages"', 'in', 'Taiwan', ')', ';', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public', 'transport', 'announcements', 'in', 'the', 'ROC']

['Southern', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'Southern', 'Min']

['Ubykh', 'Language', 'extinction', 'October', '1992', 'when', 'Tevfik', 'Esen', '\xc3', '\xa7', 'died']

['Chagatai', 'Spoken', 'in', 'Central', 'Asia', ',', 'Khorasan']

['Proto-Germanic', 'Language', 'extinction', 'evolved', 'into', 'Proto-Norse']

['Proto-Germanic', 'Language', 'extinction', 'Frankish', 'and', 'Ingvaeonic', 'by', 'the', '4th', 'century']

['Abkhaz', 'Total', 'speakers', '125', '000']

['Abkhaz', 'Writing', 'system', 'Cyrillic', 'alphabet', '(', 'Abkhaz', 'variant', ')']

['Northern', 'Thai', 'Language', 'Total', 'speakers', '6', 'million']

['Northern', 'Writing', 'system', 'Tai', 'Tham', 'script', ',', 'Thai', 'script']

['Northern', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Dyirbal', 'Total', 'speakers', '~', '5']

['Cheyenne', 'Total', 'speakers', '1', ',', '721']

['Ugaritic', 'Language', 'extinction', 'twelfth', 'century', 'BC']

['Tanacross', 'Total', 'speakers', '30', '(', '2000', ')']

['Tanacross', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan', 'variant', ')']

['Nux\xc3\xa1llk', '(', 'Bella', 'Coola', ')', 'Total', 'speakers', '20']

['Filipino', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']

['Filipino', 'Regulated', 'by', '(', 'Komisyon', 'sa', 'Wikang', 'Filipino', ')']

['Middle', 'Chinese', 'Spoken', 'in', 'China']

['Middle', 'Language', 'extinction', 'Evolved', 'into', 'Proto-Mandarin', 'and', 'other', 'Chinese', 'dialects', 'apart', 'from', 'Min']

['Middle', 'Official', 'language', 'in', 'Southern', 'and', 'Northern', 'Dynasties']

['Sona', 'Category', '(', 'sources', ')', 'The', 'language', 'has', '375', 'radicals', 'or', 'root', 'words', '--', 'based', 'on', 'the', 'terms', 'in', 'Roget's', 'original', 'thesaurus', '.', 'Ideas', 'and', 'sentences', 'are', 'formed', 'by', 'juxtaposing', 'the', 'radicals']

['Livonian', 'Total', 'speakers', 'About', '30']

['Old', 'Curonian', 'language', 'Language', 'extinction', '14-17th',

'century']  
 ['Nivkh', 'Total', 'speakers', '1', ',', '089', '(', '1989', 'census', ')']  
 ['Ga', 'Total', 'speakers', '600', ',', '000', '(', 'SIL', '2004', ')']  
 ['Ga', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Ga', 'variant', ')']  
 ['Tulu', 'Total', 'speakers', '1', '.', '95', 'million', '(', '1997', ')']  
 ['Tulu', 'Region', 'Tulu', 'Nadu', ':', 'of', 'Karnataka', 'and', 'Kerala', 'States']  
 ['Tulu', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Tulu', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Chukchi', 'Total', 'speakers', '7742', '(', '2002', 'Census', ')']  
 ['Karelian', 'Total', 'speakers', '118', ',', '000']  
 ['Karelian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Karelian', 'variant', ')']  
 ['Veps', 'language', 'Total', 'speakers', '6', ',', '355', '(', '1989', 'census', ')']  
 ['Veps', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Vepsian', 'variant', ')']  
 ['Luwian', 'Language', 'extinction', 'around', '600', 'BC']  
 ['Luwian', 'Region', 'Anatolia', ',', 'Northern', 'Syria']  
 ['Luwian', 'Language', 'codes', 'ISO', '639-3', 'hlu', '\xe2', '\x80', '\x93', 'Hieroglyphic', 'Luwian', 'Distribution', 'of', 'the', 'Luwian', 'language', '(', 'after', 'Melchert', '2003', ')']  
 ['Turkmen', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Proven\xc3\xa7al', 'Total', 'speakers', '362', ',', '000']  
 ['Kodava', 'Total', 'speakers', '122', ',', '000', '(', '1997', ')']  
 ['Kodava', 'Writing', 'system', 'Kannada', 'script', '&', 'amp', ';', 'Coorgi-Cox', 'alphabet']  
 ['Kodava', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Mirandese', 'Total', 'speakers', '10', ',', '000-15', ',', '000']  
 ['Mirandese', 'Region', 'Northeastern', 'Portugal', '(', 'municipalities', 'of', 'Miranda', 'do', 'Douro', ',', 'Vimioso', 'and', 'Mogadouro', ')']  
 ['Modern', 'Language', 'codes', 'ISO', '639-3', 'ell', '\xe2', '\x80', '\x93', 'Modern', 'Greek']  
 ['Modern', 'Language', 'codes', 'Linguasphere', 'part', 'of', '56-AAA-a']  
 ['Upper', 'Sorbian', 'Total', 'speakers', '40', ',', '000']  
 ['Upper', 'Language', 'codes', 'Linguasphere', '53-AAA-bb', '&', 'lt', ';', '53-AAA-b', '&', 'lt', ';', '53-AAA-b', '.', '.', '.', '-d', '(', 'varieties', ':', '53-AAA-bba', 'to', '53-AAA-bbf', ')']  
 ['Lower', 'Sorbian', 'Total', 'speakers', '14', ',', '000']  
 ['Lower', 'Language', 'codes', 'Linguasphere', '53-AAA-ba', '&', 'lt', ';',



'53-AAA-b', '&', 'lt', ';', '53-AAA-b', '.', '.', '.', '-d', '(',  
 'varieties', ':', '53-AAA-baa', 'to', '53-AAA-bah', ')']

['Abaza', 'Total', 'speakers', '48', ',', ', '280']

['Abaza', 'Official', 'language', 'in', 'Russian', 'autonomous',  
 'republic', 'of', 'Karachay-Cherkess']

['Sundanese', 'Region', 'parts', 'of', 'western', 'Central', 'Java']

['Walloon', 'Total', 'speakers', 'est', '.', '600', ',', ', '000']

['Orokolo', 'Total', 'speakers', '13', ',', ', '000']

['Orokolo', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Tiwa', 'Total', 'speakers', '&', 'lt', ';', '3000']

['Tiwa', 'Region', 'New', 'Mexico', ',', ', 'Arizona', 'Ethnicity']

['Tiwa', 'Language', 'codes', 'ISO', '639-3', ':', 'tix', '\xe2', '\x80',  
 '\x93', 'Southern', 'Tiwa']

['Tiwa', 'Language', 'codes', 'ISO', '639-3', 'twf', '\xe2', '\x80',  
 '\x93', 'Northern', 'Tiwa']

['Cumbric', 'Language', 'extinction', '11th', '-', '12th', 'century']

['Cumbric', 'Spoken', 'in', 'Lancashire', 'and', 'possibly', 'North',  
 'Yorkshire']

['Camuno', '(', '?', ')', 'Language', 'extinction', 'first', 'millennium',  
 'AD']

['Camuno', 'Spoken', 'in', 'the', 'south', 'side', 'of', 'the', 'Central',  
 'Alps', '(', 'Valcamonica', ',', ', 'Valtellina', ')', ',', ', 'Italy']

['Western', 'Total', 'speakers', '879', ',', ', '612', '-', 'Out', 'of', '6',  
 '.', '7', 'million', 'total', 'speakers', 'of', 'Armenian', '(', 'both',  
 'Western', 'and', 'Eastern', 'Armenian', ')']

['Western', 'Language', 'codes', 'ISO', '639-3', ':', 'hye', '\xe2',  
 '\x80', '\x93', 'Modern', 'Armenian']

['Western', 'Language', 'codes', 'ISO', '639-3', 'xcl', '\xe2', '\x80',  
 '\x93', 'Classical', 'Armenian']

['Western', 'Language', 'codes', 'ISO', '639-3', 'axm', '\xe2', '\x80',  
 '\x93', 'Middle', 'Armenian']

['Western', 'Language', 'codes', 'Linguasphere', '57-AAA-aca', 'to', '57-  
 AAA-act']

['Friulan', 'Total', 'speakers', '794', ',', ', '000']

['Swabian', 'Total', 'speakers', '819', ',', ', '000', '(', '2006', ')']

['Old', 'French', 'Language', 'extinction', 'evolved', 'into', 'Middle',  
 'French', 'by', 'the', '14th', 'century']

['Chamorro', 'Total', 'speakers', 'First', 'language', ':', 'more', 'than',  
 '60', ',', ', '000']

['Chamorro', 'Spoken', 'in', 'Northern', 'Mariana', 'Islands']

['Chamorro', 'Official', 'language', 'in', 'Northern', 'Mariana',  
 'Islands']

['Chamorro', 'Language', 'codes', 'ISO', '639-3', 'cha', 'Chamorro',  
 'language', 'spread', 'in', 'the', 'United', 'States']

['Sotho', 'Regulated', 'by', 'Pan', 'South', 'African', 'Language',  
 'Board']

['Northern', 'Writing', 'system', 'Latin', '(', 'German', 'variant', ')',  
,',']

['Northern', 'Writing', 'system', 'historically', 'the', 'Gothic',  
'script']

['Northern', 'Language', 'codes', 'ISO', '639-3', 'Austro-Bavarian',  
'languages']

['Sindhi', 'Dialects', 'of', 'Sindhi', 'Spoken', 'in']

['Sindhi', 'Dialects', 'South', 'Asia', 'Total', 'speakers']

['Sindhi', 'Dialects', '35', 'million', 'Language', 'family']

['Sindhi', 'Dialects', 'Sri', 'Lanka', 'Region']

['Sindhi', 'Regulated', 'by', 'Sindhi', 'Language', 'Authority', '(',  
'Pakistan', ')', ',']

['Sindhi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
'conjuncts', '.', 'More']

['Tup\\xc3\\xad', 'Language', 'extinction', '(', 'survives', 'as',  
'Nheengatu', ')']

['Tup\\xc3\\xad', 'Dialects', 'L', '\\xc3', '\\xad', 'ngua', 'Geral', 'of',  
'S', '\\xc3', '\\xa3', 'o', 'Paulo', '(', 'Tup', '\\xc3', '\\xad', 'Austral',  
)']

['Tup\\xc3\\xad', 'Language', 'codes', 'ISO', '639-3', ':', 'tpw', '\\xe2',  
'\\x80', '\\x93', 'Tup', '\\xc3', '\\xad', '(', 'Old', 'Tup', '\\xc3', '\\xad',  
)']

['Myanmar', 'Regulated', 'by', 'Myanmar', 'Language', 'Commission']

['Myanmar', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
'of', 'conjuncts', '.', 'More']

['Tajik', 'Total', 'speakers', 'approximately', '24', ',', '380', ',',  
'000', '(', '1991', ')']

['Samoan', 'Total', 'speakers', '369', ',', '957', 'total', 'speakers',  
'according', 'to', 'Ethnologue', '(', '2009', 'edition', ',', 'figures',  
'from', '1999', ')']

['Samoan', 'Region', 'with', 'substantial', 'communities', 'of',  
'speakers', 'in', 'New', 'Zealand']

['Samoan', 'Official', 'language', 'in', 'Samoa', '(', '199', ',', '000',  
'speakers', ')']

['Samoan', 'Official', 'language', 'in', 'American', 'Samoa', '(', '56',  
',', '700', 'speakers', ')']

['Uyghur', 'Official', 'language', 'in', 'Xinjiang', 'Uyghur',  
'Autonomous', 'Region']

['Uyghur', 'Regulated', 'by', 'Working', 'Committee', 'of', 'Ethnic',  
'Language', 'and', 'Writing', 'of', 'Xinjiang', 'Uyghur', 'Autonomous',  
'Region']

['Oromo', 'Total', 'speakers', 'From', 'more', 'than', 'the', '70', ',',

'000', ',', '000', 'Ethiopians', ',', 'the', '2007', 'census', 'lists',  
'25', ',', '488', ',', '344', 'ethnic', 'Oromos', ',', 'with', 'a',  
'number', 'of', 'others', 'also', 'speaking', 'Oromo', '.', 'The',  
'Ethnologue', 'lists', 'about', '263', ',', '000', 'speakers', 'of',  
'Oromo', 'languages', 'in', 'neighbouring', 'Somalia', 'and', 'Kenya']  
['Oromo', 'Writing', 'system', 'Latin', 'Qubee', ',', 'formerly', 'also',  
'Ethiopic', 'syllabary']  
['Oromo', 'Language', 'codes', 'ISO', '639-3', 'hae', '\xe2', '\x80',  
'\x93', 'Eastern', 'Oromo']  
['Oromo', 'Language', 'codes', 'ISO', '639-3', 'gaz', '\xe2', '\x80',  
'\x93', 'West', 'Central', 'Oromo']  
['Aymara', 'Language', 'codes', 'ISO', '639-3', ':', 'ayr', '\xe2', '\x80',  
'\x93', 'Central', 'Aymara']  
['Aymara', 'Language', 'codes', 'ISO', '639-3', 'ayc', '\xe2', '\x80',  
'\x93', 'Southern', 'Aymara']  
['Aymara', 'Language', 'codes', 'ISO', '639-3', 'Geographic',  
'Distribution', 'of', 'the', 'Aymara', 'language']  
['Oriya', 'Total', 'speakers', '31', 'million', '(', '1996', ')']  
['Oriya', 'Language', 'codes', 'Linguasphere', 'Distribution', 'of',  
'Oriya', 'speakers', 'in', 'World', 'This', 'page', 'contains', 'Indic',  
'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see',  
'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts',  
'.', 'More']  
['Shona', 'Total', 'speakers', '11', ',', '000', ',', '000']  
['Aleut', 'Total', 'speakers', 'ca', '.', '150', 'in', '2007']  
['Hindko', 'Total', 'speakers', 'approx', '.', '2', '.', '2', '\xe2',  
'\x80', '\x93', '4', 'million']  
['Hindko', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
'conjuncts', '.', 'More']  
['Lower', 'Silesian', 'language', 'Total', 'speakers', '22', ',', '900']  
['Lower', 'Region', 'Lower', 'Silesia', '.', 'Also', 'spoken', 'in',  
'Czech', 'Republic', ',', 'eastern', 'Germany']  
['Belgian', 'Total', 'speakers', '6', '.', '3', 'million', '(', 'Belgium',  
)']  
['Belgian', 'Total', 'speakers', '24', '.', '3', 'million', '(', 'DRC',  
)']  
['Belgian', 'Total', 'speakers', '783', ',', '000', '(', 'Rwanda', ')']  
['Belgian', 'Official', 'language', 'in', 'Burundi', '(', 'as', 'French',  
)']  
['Belgian', 'Language', 'codes', 'ISO', '639-3', 'Belgian', 'French', 'is',  
'primarily', 'spoken', 'in', 'the', 'French', 'Community', 'of', 'Belgium',  
'', 'highlighted', 'in', 'red']  
['Belgian', 'Language', 'codes', 'ISO', '639-3', 'The', 'French', 'spoken',  
'in', 'the', 'Democratic', 'Republic', 'of', 'the', 'Congo', ',', 'Rwanda',

', 'and', 'Burundi', ', ', 'which', 'were', 'formerly', 'Belgian',  
 'colonies', ', ', 'can', 'also', 'be', 'considered', 'an', 'offshoot', 'of',  
 'Belgian', 'French', '(', 'red', 'on', 'the', 'map', ')', '.', 'Pink', ':',  
 'Other', 'French-speaking', 'countries', ', ', 'see', 'African', 'French']  
 ['Mari', 'Standard', 'forms', 'Meadow', 'Mari']  
 ['Mari', 'Standard', 'forms', 'Hill', 'Mari']  
 ['Mari', 'Spoken', 'in', 'Russian', 'Federation', ':', 'autonomous',  
 'republics', 'Mari', 'El']  
 ['Mari', 'Total', 'speakers', 'more', 'than', '600', ', ', '000']  
 ['Mari', 'Language', 'codes', 'ISO', '639-3', ':', 'mhr', '\xe2', '\x80',  
 '\x93', 'Meadow', 'Mari', '(', 'Eastern', 'Mari', ')']  
 ['Mari', 'Language', 'codes', 'ISO', '639-3', 'mrj', '\xe2', '\x80',  
 '\x93', 'Hill', 'Mari', '(', 'Western', 'Mari', ')']  
 ['Sassarese', 'Total', 'speakers', '\xe2', '\x89', '\x88', '125', '.',  
 '000']  
 ['Sassarese', 'Writing', 'system', 'Latin', 'Alphabet', '(', 'Italian',  
 'variant', ')']  
 ['Hittite', 'Language', 'extinction', 'records', 'cease', 'after', '1200',  
 'BC', ', ', 'extinction', 'likely', 'by', '1100', 'BC']  
 ['Istriot', 'Total', 'speakers', '1', ', ', '000-2', ', ', '000']  
 ['Cebuano', 'Region', '(', 'Central', 'Visayas', 'and', 'Mindanao', ')']  
 ['Huilliche', 'Total', 'speakers', '2', ', ', '000']  
 ['Mapuche', 'Total', 'speakers', '240', ', ', '000']  
 ['Mapuche', 'Language', 'codes', 'ISO', '639-3', 'arn', 'Graffiti', 'in',  
 'Mapudungun', 'meaning', '"', 'War', 'Council', '"']  
 ['Mam', 'Total', 'speakers', '477', ', ', '717', '(', '2002', ')']  
 ['Lingua', 'Franca', 'Nova', 'Total', 'speakers', '&', 'gt', ';', '100']  
 ['Lingua', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Lingua', 'Category', '(', 'purpose', ')', 'Lingua', 'Franca', 'Nova']  
 ['Lingua', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Silesian', 'Total', 'speakers', '(', 'about', '60', ', ', '000', 'people',  
 'declared', 'Silesian', 'as', 'their', 'native', 'language', 'in',  
 'Polish', 'census', ')']  
 ['Silesian', 'Spoken', 'in', 'Silesian', 'Voivodeship', 'and', 'Opole',  
 'Voivodeship', 'in', 'Poland', ', ', 'Moravian-Silesian']  
 ['Silesian', 'Region', 'and', 'Olomouc', '(', 'only', 'Jesen', '\xc3',  
 '\xad', 'k', 'District', ')', 'in', 'Czech', 'Republic', ', ', 'Germany',  
 'and', 'USA']  
 ['Silesian', 'Region', 'Upper', 'Silesia', '/', 'Silesia']  
 ['Cherokee', 'Total', 'speakers', '12', ', ', '000', 'to', '22', ', ', '000']  
 ['Cherokee', 'Region', 'Oklahoma', 'and', 'the', 'Qualla', 'Boundary', ', ',  
 'North', 'Carolina']  
 ['Yor\xc3\xb9b\xc3\xa1', 'Total', 'speakers', '19', 'million', '(',  
 'Johnstone', '1993', 'as', 'cited', 'in', 'Ethnologue', ')']

['Taishanese', 'Total', 'speakers', '~', '1-2', 'million']  
 ['Taishanese', 'Spoken', 'in', 'United', 'States', '(', 'mostly',  
 'California', 'and', 'New', 'York', 'City', ')']  
 ['Baoule', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]  
 ['Baoule', 'Total', 'speakers', '2', ',', '130', ',', '000']  
 ['J\xc3\xa8rriais', 'Total', 'speakers', '2', ',', '874', '(', '113', 'as',  
 'main', 'language', ';', '2', ',', '761', 'as', 'secondary', 'language',  
 ')', '(', '2001', 'Census', ')']  
 ['Papiamentu', 'Total', 'speakers', '329', ',', '002']  
 ['Taa', 'Total', 'speakers', '4', ',', '200']  
 ['Taa', 'Region', 'western', 'Southern', 'and', 'western', 'Kweneng',  
 'districts', 'in', 'Botswana', ';', 'southern', 'Omaheke', 'and',  
 'northeastern', 'Hardap', 'regions', 'in', 'Namibia']  
 ['Vedic', 'Sanskrit', 'Language', 'extinction', 'evolved', 'into',  
 'Classical', 'Sanskrit', 'by', 'the', 'sixth', 'century', 'BC']  
 ['Sinhala', 'Total', 'speakers', '15', '.', '6', 'million']  
 ['Sinhala', 'Writing', 'system', 'Sinhala', 'abugida', '(', 'developed',  
 'from', 'the', 'Brahmi', ')']  
 ['Sinhala', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Balinese', 'Total', 'speakers', '3', '.', '9', 'million', '(', 'as',  
 'of', '2001', ')']  
 ['Buginese', 'Spoken', 'in', 'South', 'Sulawesi', ',']  
 ['Buginese', 'Total', 'speakers', '3', '.', '5', 'to', '4', 'million']  
 ['Gayo', 'Total', 'speakers', '550', ',', '000']  
 ['Iban', 'Total', 'speakers', '700', ',', '000']  
 ['Javanese', 'Language', 'codes', 'ISO', '639-3', 'jas', '\xe2', '\x80',  
 '\x93', 'New', 'Caledonian', 'Javanese']  
 ['Javanese', 'Language', 'codes', 'ISO', '639-3', 'osi', '\xe2', '\x80',  
 '\x93', 'Osing', 'language']  
 ['Javanese', 'Language', 'codes', 'ISO', '639-3', 'kaw', '\xe2', '\x80',  
 '\x93', 'Old', 'Javanese']  
 ['Kawi', 'Language', 'extinction', 'literary', 'language', ',', 'archaic',  
 'by', '14th', 'century']  
 ['Madurese', 'Total', 'speakers', '14', 'million', '(', '1995', ')']  
 ['Makassar', 'Language', 'Total', 'speakers', '1', ',', '600', ',', '000',  
 '(', '1989', ')']  
 ['Makassar', 'Region', 'southern', 'tip', 'of', 'South', 'Sulawesi', ',',  
 'Sulawesi', '(', 'Celebes', ')']  
 ['Minangkabau', 'Total', 'speakers', '6', ',', '500', ',', '000']  
 ['Minangkabau', 'Region', 'Negeri', 'Sembilan', '(', 'Malaysia', ')']  
 ['Nicaraguan', 'Total', 'signers', '3', ',', '000', '(', 'estimated', 'in',  
 '1997', ';', 'no', 'official', 'census', 'data', 'available', ')']  
 ['Nicaraguan', 'Language', 'codes', 'ISO', '639-3', 'ncs', 'sign',

'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
'\xe2', '\x80', '\x94', 'legal', 'recognition']  
['Afar', 'Total', 'speakers', '1', '.', '4', '\xe2', '\x80', '\x93', '1',  
'.', '5', 'million']  
['Afar', 'Region', 'Afar', 'Triangle', ',', 'North-Eastern', 'Africa']  
['Dzongkha', 'Total', 'speakers', 'First', 'language', ':', '171', ',',  
'000']  
['Dzongkha', 'Total', 'speakers', 'Second', 'language', '~', '470', ',',  
'000']  
['Dzongkha', 'Regulated', 'by', 'Dzongkha', 'Development', 'Commission']  
['Dzongkha', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
'of', 'conjuncts', '.', 'More']  
['Fijian', 'Region', 'Ogea', 'and', 'Vatoka', ';', 'spoken', 'as', 'second',  
'language', 'in', 'the', 'rest', 'of', 'Fiji']  
['Kashmiri', 'Total', 'speakers', '4', '.', '6', 'million']  
['Kashmiri', 'Region', 'Northwestern', 'region', 'of', 'the', 'Indian',  
'subcontinent']  
['Rundi', 'Total', 'speakers', '9', 'million', '(', '2011', ')']  
['Latino', 'Category', '(', 'purpose', ')', 'Latino', 'sine', 'flexione']  
['Latino', 'Category', '(', 'sources', ')', 'Completely', 'based', 'on',  
'Latin', ',', 'but', 'influenced', 'by', 'ideas', 'in', 'other',  
'auxiliary', 'languages']  
['Ligurian', '(', 'ancient', ')', 'Language', 'extinction', 'after', '210',  
'BC']  
['Ligurian', 'Region', 'Northern', 'Mediterranean', 'Coast', 'straddling',  
'South-east', 'French', 'and', 'North-east', 'Italian', 'coasts']  
['Sango', 'Total', 'speakers', 'an', 'estimated', '404', ',', '000',  
'native', 'speakers', ',', '1', '.', '6', 'million', 'second-language',  
'speakers', 'as', 'of', '1988']  
['Sango', 'Official', 'language', 'in', 'Central', 'African', 'Republic']  
['Tongan', 'Total', 'speakers', '200', ',', '000', '(', 'as', 'of', '2010',  
')']  
['Tsonga', 'Total', 'speakers', '3', ',', '275', ',', '105']  
['Zhuang', 'Language', 'codes', 'ISO', '639-3', ':', 'zch', '\xe2', '\x80',  
'\x93', 'Central', 'Hongshuihe', 'Zhuang']  
['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zeh', '\xe2', '\x80',  
'\x93', 'Eastern', 'Hongshuihe', 'Zhuang']  
['Auslan', 'Region', 'Northern', 'dialect', '(', 'NSW', ',', 'QLD', ')',  
',', 'Victorian', 'dialect', 'and', 'South-Eastern', 'dialect', '(', 'all',  
'other', 'states', ')']  
['Auslan', 'Total', 'signers', '14', ',', '000']  
['Auslan', 'Language', 'codes', 'ISO', '639-3', 'asf', 'sign', 'language',  
'\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
'\x94', 'legal', 'recognition']

['Algonquin', 'Total', 'speakers', '2', ',', '680']  
 ['Algonquin', 'Region', 'Quebec', 'and', 'into', 'Ontario']  
 ['French', 'Total', 'signers', '50', ',', '000', '\xe2', '\x80', '\x93',  
 '100', ',', '000', '(', '1986', ')']  
 ['French', 'Language', 'codes', 'ISO', '639-3', 'fsl', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Dacian', 'Language', 'extinction', 'probably', 'by', 'the', 'sixth',  
 'century', 'AD']  
 ['Dacian', 'Spoken', 'in', 'and', 'Northern', 'Bulgaria']  
 ['Galatian', 'Language', 'extinction', '4th', 'century', 'AD']  
 ['Eyak', 'Language', 'extinction', 'January', '21']  
 ['Eyak', 'Language', 'extinction', 'upon', 'the', 'death', 'of', 'Marie',  
 'Smith', 'Jones']  
 ['Eyak', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact',  
 'distribution', 'of', 'Eyak', 'highlighted']  
 ['Wenedyk', 'Category', '(', 'sources', ')', '(', 'Romance', 'language',  
 'based', 'on', 'Polish', ')']  
 ['Wenedyk', 'Setting', 'and', 'usage', 'if', 'Latin', 'had', 'replaced',  
 'Polish's', 'ancestor']  
 ['Sranan', 'Total', 'speakers', '300', ',', '000']  
 ["Martha's", 'Vineyard', 'Sign', 'Language', 'Language', 'extinction',  
 'with', 'the', 'death', 'of', 'Katie', 'West', '(', '1952', ')']  
 ["Martha's", 'Region', "Martha's", 'Vineyard', ',', 'Massachusetts']  
 ["Martha's", 'Language', 'codes', 'ISO', '639-3', 'mre', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Khoekhoegowab', 'Total', 'speakers', '270', ',', '000']  
 ['Khoekhoegowab', 'Total', 'speakers', 'Namibia', ':', '200', ',', '000',  
 '(', '2006', ')']  
 ['Khoekhoegowab', 'Total', 'speakers', 'plus', '18', ',', '000', 'Hai',  
 '\xc7', '\x81', 'om', '(', '2006', ')']  
 ['Khoekhoegowab', 'Total', 'speakers', 'South', 'Africa', ':', '51', ',', '000',  
 '(', '2006', ')']  
 ['Khoekhoegowab', 'Total', 'speakers', 'Botswana', ':', '200', '(', '2006',  
 ')']  
 ['Khoekhoegowab', 'Spoken', 'in', 'Namibia', ',', 'Botswana', 'and',  
 'South', 'Africa']  
 ['Avar', 'Total', 'speakers', '788', ',', '960']  
 ['Ruthenian', 'Language', 'extinction', 'developed', 'into', 'Belarusian',  
 ',', 'Ukrainian', 'and', 'Rusyn']  
 ['Ruthenian', 'Spoken', 'in', 'Polish-Lithuanian', 'Commonwealth', ',',  
 '(', 'language', 'of', 'administration', 'of', 'Grand', 'Duchy', 'of',  
 'Lithuania', 'until', '1699', ')']  
 ['Old', 'Persian', 'Language', 'extinction', 'Ancestor', 'of', 'Middle',  
 'Persian']

['Middle', 'Dutch', 'Language', 'extinction', 'developed', 'into',  
'modern', 'Dutch', 'by', 'the', 'middle', 'of', 'the', '16th', 'century']  
['Middle', 'Spoken', 'in', 'the', 'Low', 'Countries']  
['Middle', 'High', 'German', 'Language', 'extinction', 'developed', 'into',  
'Early', 'New', 'High', 'German', 'from', 'the', '14th', 'century']  
['Early', 'New', 'High', 'German', 'Language', 'extinction', 'developed',  
'into', 'Modern', 'German', 'from', 'the', '1650s']  
["Dena'ina", 'Region', 'Alaska', '(', 'Cook', 'Inlet', 'region', ',',  
'Lake', 'Clark', ',', 'Lake', 'Iliamna', ')']  
['Ilocano', 'Total', 'speakers', '7', '.', '7', 'million', ',', '2', '.',  
'3', 'million', '2nd', 'language', '=', '10', 'million', 'total', ';',  
'3rd', 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']  
['Ilocano', 'Writing', 'system', 'Latin', '(', 'Ilocano', 'or', 'Filipino',  
'variant', ')', ';']  
['Ilocano', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
'Language']  
['Phoenician', 'Language', 'extinction', 'continued', 'in', 'its', 'Punic',  
'form', 'perhaps', 'as', 'late', 'as', '7th', 'century', 'AD']  
['Phoenician', 'Spoken', 'in', 'Southern', 'Mediterranean', 'Iberia']  
['Zauruan', 'Total', 'speakers', '7', ',', '000']  
['Aequian', 'Language', 'extinction', '3rd', 'century', 'BC']  
['Aequian', 'Language', 'codes', 'ISO', '639-3', 'xae', 'Abraham',  
"Ortelius's", 'map', 'of', 'ancient', 'Latium', ',', 'published', 'in',  
'1595']  
['Alemannic', 'German', 'Total', 'speakers', 'about', '10', 'million']  
['Alemannic', 'Language', 'codes', 'ISO', '639-3', 'gsw', '\xe2', '\x80',  
'\x93', 'Swiss', 'German']  
['Alemannic', 'Language', 'codes', 'ISO', '639-3', 'swg', '\xe2', '\x80',  
'\x93', 'Swabian', 'German']  
['Alemannic', 'Language', 'codes', 'ISO', '639-3', 'wae', '\xe2', '\x80',  
'\x93', 'Walser', 'German', 'The', 'traditional', 'distribution', 'area',  
'of', 'Western', 'Upper', 'German', '(', '=', 'Alemannic', ')', 'dialect',  
'features', 'in', 'the', '19th', 'and', '20th', 'century']  
['Brithenig', 'Setting', 'and', 'usage', 'if', 'Latin', 'had', 'replaced',  
'Celtic']  
['Zimakani', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
['Old', 'High', 'German', 'Language', 'extinction', 'developed', 'into',  
'Middle', 'High', 'German', 'from', 'the', '11th', 'century']  
['Old', 'Spoken', 'in', 'Sporadic', 'communities', 'in', 'Eastern', 'Gaul']  
['Old', 'Latin', 'Language', 'extinction', 'Developed', 'into',  
'Classical', 'Latin', 'in', '1st', 'century', 'BC']  
['Old', 'Language', 'codes', 'ISO', '639-3', '2nd', 'century', 'BC', '.',  
'It', 'is', 'unlikely', 'Latin', 'was', 'spoken', 'much', 'beyond', 'the',  
'green', 'area']  
['Xavante', 'Total', 'speakers', '\xe2', '\x89', '\x88', '10', ',', '000']  
['Marwari', 'Language', 'codes', 'ISO', '639-3', 'rwr', '\xe2', '\x80',



'\x93', 'Marwari', '(', 'India', ')']  
 ['Marwari', 'Language', 'codes', 'ISO', '639-3', 'mve', '\xe2', '\x80',  
 '\x93', 'Marwari', '(', 'Pakistan', ')']  
 ['Mozarabic', 'Language', 'extinction', 'by', 'the', 'Late', 'Middle',  
 'Ages']  
 ['Hattic', 'Language', 'extinction', '1100', 'BC', '?']  
 ['Somali', 'Spoken', 'in', 'Somali', 'communities', 'in', 'the', 'Middle',  
 'East', ',', 'Europe', 'and', 'North', 'America']  
 ['Gutnish', 'Total', 'speakers', '~', '500', '?']  
 ['Venda', 'Total', 'speakers', '875', '000']  
 ['Bemba', 'Total', 'speakers', '4', ',', '208', ',', '940']  
 ['Hopi', 'Total', 'speakers', '\xe2', '\x89', '\x88', '5000']  
 ['Kaonde', 'Total', 'speakers', '240', ',', '000']  
 ['Zizilivakan', 'Total', 'speakers', '2', ',', '800', '(', '2002', ')']  
 ['Zizilivakan', 'Region', 'Far', 'North', 'Province', ';', 'Adamawa',  
 'State']  
 ['Lozi', 'Total', 'speakers', 'around', '500', ',', '000']  
 ['Lozi', 'Region', 'Western', 'Zambia', ',', 'Caprivi']  
 ['Ekajuk', 'Total', 'speakers', '30', ',', '000', '(', '1986', ')']  
 ['Ekajuk', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Nsibidi',  
 'ideographs', ')']  
 ['Western', 'Panjabi', 'Total', 'speakers', '60', ',', '600', ',', '000',  
 'in', 'Pakistan', '(', '2000', ')', ',', '1', ',', '910', ',', '000', 'in',  
 'India', '(', '2007', ')', '.', 'Population', 'total', 'all', 'countries',  
 ':', '62', ',', '648', ',', '000']  
 ['Western', 'Language', 'codes', 'ISO', '639-3', 'Punjabi', 'Dialects']  
 ['Lunda', 'Total', 'speakers', '\xce', '\xa3', '400', ',', '000', '(',  
 '2001', ')', '(', '222', ',', '000', 'in', 'Zambia', ',', '178', ',',  
 '000', 'in', 'Angola', ')']  
 ['Masaba', 'Total', 'speakers', '1', ',', '500', ',', '000']  
 ['Yorkshire', 'Dialect', 'Total', 'speakers', 'unknown']  
 ['Mauritian', 'Total', 'speakers', '1', '.', '2', 'million']  
 ['Mauritian', 'Language', 'codes', 'Linguasphere', '(', 'to', '51-AAC-cee',  
 ')']  
 ['Tahitian', 'Total', 'speakers', '120', ',', '000']  
 ['Palauan', 'Spoken', 'in', 'Northern', 'Mariana', 'Islands']  
 ['Palauan', 'Total', 'speakers', 'about', '15', ',', '000']  
 ['Franco-Proven\xe3\xa7al', 'Arpitan', 'Total', 'speakers', '(', 'est',  
 '.', ')', '113', ',', '400']  
 ['Franco-Proven\xe3\xa7al', 'Language', 'codes', 'Linguasphere', 'Map',  
 'of', 'the', 'Franco-Proven', '\xc3', '\xa7', 'al', 'Language', 'Area']  
 ['Sicilian', 'Spoken', 'in', 'Calabria', '(', 'Central', 'and', 'Southern',  
 ')']  
 ['Sicilian', 'Spoken', 'in', 'United', 'States', '(', 'New', 'York',  
 'City', ',', 'Boston', ',', 'Baltimore', ')']  
 ['Sicilian', 'Total', 'speakers', 'est', '.', '8', 'million', '(',

'Gordon', ',', '2005', ')']  
 ['Sicilian', 'Language', 'codes', 'Linguasphere', '51-AAA-re', '&', 'amp',  
 ';', '-rf']  
 ['Sicilian', 'Language', 'codes', 'Linguasphere', '(', 'mainland', '51-AAA-  
 rc', '&', 'amp', ';', '-rd', ')']  
 ['Comorian', 'Total', 'speakers', '350', ',', '702', '(', '2004', ')']  
 ['Maithili', 'Total', 'speakers', '34', ',', '700', ',', '000']  
 ['Maithili', 'Region', 'parts', 'of', 'West', 'Bengal', 'in', 'India']  
 ['Maithili', 'Official', 'language', 'in', 'In', '8th', 'schedule', 'of',  
 'Constitution', 'of', 'India', ',', 'Bihar', 'state', 'in', 'India', 'and',  
 'Nepal']  
 ['Maithili', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Haitian', 'Total', 'speakers', '12', ',', '000', ',', '000']  
 ['Haitian', 'Regulated', 'by', 'Minist', '\xc3', '\xa8', 're', 'de', "l'",  
 '\xc3', '\xa9', 'ducation', 'nationale', 'et', 'de', 'la', 'formation',  
 'professionnelle']  
 ['Old', 'East', 'Slavic', 'Language', 'extinction', 'developed', 'into',  
 'the', 'various', 'East', 'Slavic', 'languages']  
 ['Bhojpuri', 'Total', 'speakers', '38', ',', '546', ',', '000']  
 ['Bhojpuri', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Cajun', 'Total', 'speakers', 'over', '250', ',', '000']  
 ['Cajun', 'Language', 'codes', 'Linguasphere', 'French', 'spread', 'in',  
 'Louisiana', '.', 'Parishes', 'marked', 'in', 'yellow', 'are', 'those',  
 'where', '4', '\xe2', '\x80', '\x93', '10', '%', 'of', 'the', 'population',  
 'speak', 'French', 'or', 'Cajun', 'French', 'at', 'home']  
 ['Cajun', 'Language', 'codes', 'Linguasphere', 'orange', '10', '\xe2',  
 '\x80', '\x93', '15', '%']  
 ['Cajun', 'Language', 'codes', 'Linguasphere', 'red', '15', '\xe2', '\x80',  
 '\x93', '20', '%']  
 ['Cajun', 'Language', 'codes', 'Linguasphere', 'brown', '20', '\xe2',  
 '\x80', '\x93', '30', '%']  
 ['Picard', 'Total', 'speakers', '~', '700', '000']  
 ['Picard', 'Official', 'language', 'in', '(', 'official', 'recognition',  
 'as', 'regional', 'language', 'by', 'Belgium', ')']  
 ['Crimean', 'Gothic', 'Language', 'extinction', 'by', 'the', '18th',  
 'century', '(', '?', ')']  
 ['Selonian', 'Selian', 'Language', 'extinction', '16th', 'century']  
 ['Old', 'Chinese', 'Spoken', 'in', 'Shang', 'Dynasty']  
 ['Old', 'Language', 'extinction', 'Evolved', 'into', 'Middle', 'Chinese',  
 'and', 'proto-Min']

['Thracian', 'Language', 'extinction', 'Fifth', 'century']  
 ['Thracian', 'Spoken', 'in', 'Eastern', 'and', 'Central', 'Balkans', ';']  
 ['Chewa', 'Total', 'speakers', '9', '.', '3', 'million']  
 ['Slavey', 'Total', 'speakers', 'Total', ':', '3', ',', '545']  
 ['Slavey', 'Total', 'speakers', 'North', 'Slavey', ':', '1', ',', '235']  
 ['Slavey', 'Total', 'speakers', 'South', 'Slavey', ':', '2', ',', '310']  
 ['Slavey', 'Language', 'codes', 'ISO', '639-3', ':', 'scs', '\xe2', '\x80', '\x93', 'North', 'Slavey']  
 ['Slavey', 'Language', 'codes', 'ISO', '639-3', 'xsl', '\xe2', '\x80', '\x93', 'South', 'Slavey']  
 ['Wu', 'Total', 'speakers', '~', '90', 'million']  
 ['Wu', 'Region', 'Fujian', 'province', ';', 'North', 'Point']  
 ['Wu', 'Spoken', 'in', 'China', ';', 'and', 'overseas', 'Chinese', 'in', 'urban', 'emigrant', 'communities', 'originating', 'from', 'Wu', 'speaking', 'areas', 'of', 'China', '-', 'particularly', 'United', 'States', '(', 'New', 'York', 'City', ')']  
 ['Tuvan', 'Total', 'speakers', '264', ',', '400']  
 ['Southern', 'Language', 'codes', 'ISO', '639-3', 'Southern', 'Sami', 'is', '1', 'on', 'this', 'map']  
 ['Auvergnat', 'Total', 'speakers', '1', '.', '3', 'million']  
 ['Soqotri', 'Total', 'speakers', '64', ',', '000']  
 ['L\x3\xa1adan', 'Category', '(', 'sources', ')', 'a', 'priori', 'language', ',', 'with', 'influences', 'from', 'Navajo', 'and', 'English']  
 ['L\x3\xa1adan', 'Setting', 'and', 'usage', 'experiment', 'in', 'feminist', 'linguistics', ',', 'and', 'featured', 'in', 'Elgin's', 'novel', 'Native', 'Tongue']  
 ['Piedmontese', 'Total', 'speakers', '~', '2', ',', '000', ',', '000']  
 ['Suzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Suzhou', 'Region', 'Suzhou', ',', 'Jiangsu', 'province']  
 ['K\x3\xb6lsch', 'Total', 'speakers', '250', ',', '000']  
 ['Low', 'Language', 'codes', 'ISO', '639-3', 'nds', '\xe2', '\x80', '\x93', 'Low', 'German', '(', 'generic', ')']  
 ['Wyandot', 'Language', 'extinction', 'Spoken', 'until', 'recently', 'near', 'Sandwich']  
 ['Wyandot', 'Language', 'extinction', 'Ontario', 'and', 'Wyandotte']  
 ['Wyandot', 'Language', 'extinction', 'Oklahoma', '.', 'There', 'were', 'two', 'older', 'adult', 'speakers', 'still', 'alive', 'in', '1961']  
 ['Elamite', 'Language', 'extinction', 'by', 'the', 'end', 'of', 'the', '4th', 'century', 'BC']  
 ['Elamite', 'Spoken', 'in', 'Elamite', 'Empire', '(', 'extinct', ')']  
 ['Chiac', 'Region', 'Acadian', 'communities', 'throughout', 'the', 'Maritime', 'provinces']  
 ['Norfolk', 'Total', 'speakers', '580', 'on', 'Norfolk', '(', '1989', ')']  
 ['Norfolk', 'Total', 'speakers', '36', 'on', 'Pitcairn', '(', '2002', ')', ',', 'Australia', ',', 'New', 'Zealand']  
 ['Bundeli', 'Total', 'speakers', '.', '5-1', 'Million', '(', '2001', ')']

['Bagheli', 'Total', 'speakers', '.', '5-1', 'Million', '(', '2001', ')']  
 ['Bagheli', 'Official', 'language', 'in', 'India', 'Madhya', 'Pradesh',  
 ')']  
 ['Ewe', 'Region', 'Southern', 'Ghana', 'east', 'of', 'the', 'Volta',  
 'River', ',', 'southern', 'Togo']  
 ['Khwarezmian', 'Language', 'extinction', 'ca', '.', '13th', 'century']  
 ['Hawaii', 'Total', 'speakers', '600', '000']  
 ['Torres', 'Language', 'codes', 'Linguasphere', '52-ABB-cb', '(',  
 'varieties']  
 ['Torres', 'Language', 'codes', 'Linguasphere', ':', '52-ABB-cba', 'to', '-  
 cbd', ')']  
 ['Zarma', 'Total', 'speakers', '2', '.', '4', 'million']  
 ['Zyrian', 'Total', 'speakers', '293', ',', '000', '(', 'census', '2002',  
 '.', ')']  
 ['Saraiki', 'Total', 'speakers', '13', '.', '9', 'million', 'in',  
 'Pakistan', '(', '1998', 'Population', 'and', 'Housing', 'Census', ',',  
 'Pakistan', ')', ',', ]  
 ['Saraiki', 'Total', 'speakers', '68', ',', '000', 'in', 'India', '(',  
 'Census', 'of', 'India', ',', '2001', ')', '(', 'combined', 'figure',  
 'for', 'persons', 'claiming', 'either', 'the', 'Multani', 'dialect', 'or',  
 'the', 'Bahawalpuri', 'dialect', ')']  
 ['Saraiki', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Shanghainese', 'Region', 'Shanghai', ';', 'also', 'spoken', 'in',  
 'overseas', 'emigrant', 'communities', ',', 'including', 'in', 'United',  
 'States', '(', 'New', 'York', 'City', ')']  
 ['Shanghainese', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Lakota', 'Total', 'speakers', '6', ',', '000']  
 ['Lakota', 'Region', 'Primarily', 'North', 'Dakota', 'and', 'South',  
 'Dakota']  
 ['Komi-Permyak', 'Total', 'speakers', '116', ',', '000']  
 ['Kildin', 'Language', 'codes', 'ISO', '639-3', 'Kildin', 'Sami', 'is',  
 '8', 'on', 'this', 'map']  
 ['Inari', 'Total', 'speakers', '~', '300']  
 ['Inari', 'Language', 'codes', 'ISO', '639-3', 'smn', 'Inari', 'Sami',  
 'is', '7', 'on', 'this', 'map']  
 ['Erzya', 'Total', 'speakers', ':', 'in', 'Russia', '440', ',', '000']  
 ['Erzya', 'Total', 'speakers', 'worldwide', '517', ',', '575']  
 ['Evenki', 'Total', 'speakers', '29', ',', '000']  
 ['Shoshoni', 'Total', 'speakers', 'USA', ':', '2905', '(', '2000', ')']  
 ['Balti', 'Total', 'speakers', '337', ',', '000', ',', '(', '450', ',',  
 '000', ')', 'Source', ':', 'Khowar', 'Academy']  
 ['Balti', 'Region', 'Pakistan', '(', 'Kashmir', ',', 'Gilgit-Baltistan',  
 ')', ',', 'Kargil', ',', 'India', 'and', 'small', 'pockets', 'in',

'Karachi', ',', 'Rawalpindi', ',', 'Islamabad', '&', 'amp', ';', 'Lahore']  
 ['Balti', 'Writing', 'system', 'Urdu', 'script', 'and', 'modified',  
 'Tibetan', 'script']  
 ['Kalmyk', 'Total', 'speakers', '174', ',', '000']  
 ['Kalmyk', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-BAA-b']  
 ['Beja', 'Total', 'speakers', '1', ',', '178', ',', '000']  
 ['Kabardian', 'Total', 'speakers', '1', ',', '632', ',', '500']  
 ['Blackfoot', 'Total', 'speakers', '5', ',', '000', 'to', '&', 'lt', ';',  
 '8', ',', '000']  
 ['Kongo', 'Language', 'codes', 'ISO', '639-3', 'kwy', '\xe2', '\x80',  
 '\x93', 'San', 'Salvador', 'Kongo']  
 ['Lombard', 'Total', 'speakers', '3', '.', '5', 'million']  
 ['Lombard', 'Language', 'codes', 'Linguasphere', '51-AAA-oc', '&', 'amp',  
 ';', '51-AAA-od']  
 ['Moabite', 'Language', 'extinction', '5th', 'century', 'BC']  
 ['O'odham', 'Total', 'speakers', 'USA', ':', '9595', '(', '2000', ')', ';',  
 'Mexico', ':', '153', '(', '2000', ')']  
 ['Phrygian', 'Language', 'extinction', '5th', 'century']  
 ['Phrygian', 'Spoken', 'in', 'Central', 'Asia', 'Minor']  
 ['Adyghe', 'Total', 'speakers', '500', ',', '000']  
 ['H\xc3\xa4n', 'Writing', 'system', 'Latin', '(', 'Den', '\xc3', '\xa9',  
 'variant', ')']  
 ['Ojibwe', 'Language', 'Total', 'speakers', '56', ',', '531', '(', '47',  
 ',', '740', 'in', 'Canada', ';', '8', ',', '791', 'in', 'the', 'United',  
 'States', ')']  
 ['Ojibwe', 'Region', 'British', 'Columbia', ';', 'United', 'States', ':',  
 'Michigan']  
 ['Ojibwe', 'Region', 'groups', 'in', 'North', 'Dakota']  
 ['Ojibwe', 'Writing', 'system', 'Latin', 'alphabet', ',', 'various',  
 'orthographies', 'in', 'Canada', 'and', 'the', 'United', 'States', ';',  
 'Ojibwe', 'syllabics', 'in', 'Canada', ';', 'Great', 'Lakes', 'Algonquian',  
 'syllabary', 'in', 'the', 'United', 'States']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', ':', 'ojs', '\xe2', '\x80',  
 '\x93', 'Severn', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ojg', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ojc', '\xe2', '\x80',  
 '\x93', 'Central', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'obj', '\xe2', '\x80',  
 '\x93', 'Northwestern', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ojw', '\xe2', '\x80',  
 '\x93', 'Western', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'Linguasphere', '62-ADA-d', '(', 'Ojibwa',  
 '+', 'Anissinapek', ')']  
 ['Ojibwe', 'Language', 'codes', 'Linguasphere', 'Location', 'of', 'all',  
 'Ojibwe', 'Reservations', '/', 'Reserves', 'and', 'cities', 'with', 'an',

'Ojibwe', 'population', 'in', 'North', 'America', ',', 'with', 'diffusion', 'rings', 'about', 'communities', 'speaking', 'the', 'Ojibwe', 'language']

['Venetian', 'Total', 'speakers', '2', ',', '280', ',', '387', '(', 'as', 'of', '2000', ')', '(', 'some', 'estimate', 'goes', 'up', 'to', '5', ',', '000', ',', 'people', 'in', 'Triveneto', 'and', 'Istria', 'only', ')']

['Venetian', 'Spoken', 'in', 'Italy', ',', 'Slovenia', ',', 'Croatia', ',', 'Brazil', '(', 'States', 'of', 'Esp', '\xc3', '\xad', 'rito', 'Santo', ',', 'S', '\xc3', '\xa3', 'o', 'Paulo', ',', 'Paran', '\xc3', '\xa1', ',', 'Rio', 'Grande', 'do', 'Sul', 'and', 'Santa', 'Catarina', 'under', 'the', 'name', 'of', 'Tali', '\xc3', '\xa0', 'n', 'with', 'some', 'influence', 'of', 'Portuguese', 'and', 'other', 'Northern', 'Italian', 'languages', ')', ',', 'Mexico', '(', 'in', 'the', 'town', 'of', 'Chipilo', 'near', 'Puebla', 'a', 'northern', 'Venetian', 'variety', ',', 'Trevisan-Bellunese', ',', 'is', 'spoken', ')']

['Gen', 'Total', 'speakers', '358', ',', '900']

['Stadsfries', 'Total', 'speakers', '45', ',', '000']

['Stadsfries', 'Language', 'codes', 'ISO', '639-3', 'Map', 'of', 'the', 'areas', 'in', 'which', 'Stadsfries', 'is', 'spoken']

['West', 'Frisian', 'Total', 'speakers', '360', ',', '000', '\xe2', '\x80', '\x93', '700', ',', '000']

['Saterland', 'Frisian', 'Total', 'speakers', '2', ',', '000']

['North', 'Frisian', 'Total', 'speakers', '10', ',', '000']

['North', 'Language', 'codes', 'Linguasphere', '52-ACA-e', '(', 'varieties']

['North', 'Language', 'codes', 'Linguasphere', ':', '52-ACA-aaa', 'to', '-eak', '&', 'amp', ';']

['North', 'Language', 'codes', 'Linguasphere', 'extinct', '-eba', '&', 'amp', ';', '-ebb', ')']

['North', 'Language', 'codes', 'Linguasphere', 'North', 'Frisian', 'dialects']

['Classical', 'Spoken', 'in', 'Historically', 'in', 'the', 'Middle', 'East', ',', 'now', 'used', 'as', 'a', 'liturgical', 'language', 'of', 'Islam']

['Classical', 'Language', 'codes', 'ISO', '639-3', 'Verses', 'from', 'the', 'Qur'an', 'in', 'Classical', 'Arabic', ',', 'written', 'in', 'the', 'cursive', 'Arabic', 'script']

['Sireniki', 'Eskimo', 'language', 'Spoken', 'in', 'Russian', 'Federation']

['Sireniki', 'Language', 'extinction', 'Last', 'speaker', 'died', 'in', '1997']

['Sireniki', 'Region', 'Bering', 'Strait', 'region', ',', 'mixed', 'populations', 'in', 'settlements', 'Sireniki', 'and', 'Imtuk']

['Limburgish', 'Spoken', 'in', 'Belgium', '(', 'in', 'the', 'province', 'of', 'Limburg', 'and', 'also', 'in', 'some', 'villages', 'in', 'the', 'northeast', 'of', 'the', 'Walloon', 'province', 'of', 'Liege', ')']

['Limburgish', 'Regulated', 'by', 'Veldeke', 'Limburg', ',', 'Raod',

'veur', "'t", 'Limburgs']  
 ['Limburgish', 'Language', 'codes', 'Linguasphere', '52-ACB-al',  
 'Position', 'of', 'Limburgish', '(', 'orange', ')', 'among', 'the',  
 'other', 'minority', 'languages', ',', 'regional', 'languages', 'and',  
 'dialects', 'in', 'the', 'Benelux']  
 ['Soga', 'Total', 'speakers', 'First', 'language', '(', 'native',  
 'speakers', ')', ':', '+', '3', 'million', '(', '2007', 'census', ')']  
 ['Soga', 'Total', 'speakers', 'Second', 'language', ':', '+', '100', '000',  
 '(', '2007', 'census', ')']  
 ['Soga', 'Official', 'language', 'in', 'Major', 'language', 'of', 'Uganda',  
 ',', 'official', 'status', 'unclear', '/', 'pending']  
 ['Soga', 'Regulated', 'by', 'Lusoga', 'Language', 'Authority', '(', 'LULA',  
 ')', '(', 'Uganda', ')']  
 ['Pangasinense', 'Total', 'speakers', '1', '.', '5', 'million', ';', '9th',  
 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']  
 ['Pangasinense', 'Region', 'Ilocos', 'and', 'Central', 'Luzon']  
 ['Pangasinense', 'Writing', 'system', 'Latin', ',', 'Alibata', ',',  
 'Pangasinense', 'or', 'Filipino', 'variant', ')', ';']  
 ['Pangasinense', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Ottoman', 'Turkish', 'Language', 'extinction', 'Reformed', 'into',  
 'Modern', 'Turkish', 'in', '1928']  
 ['Finnish', 'Total', 'signers', '5', ',', '000', '(', '1986', ')']  
 ['Finnish', 'Language', 'codes', 'ISO', '639-3', 'fse', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Bulgar', 'Language', 'extinction', 'between', 'the', '9th', 'century',  
 'on', 'the', 'Danube', 'and', 'the', '14th', 'century', 'on', 'the',  
 'Volga']  
 ['Bulgar', 'Spoken', 'in', 'from', 'Central', 'Asia', 'to', 'the',  
 'steppes', 'North', 'of', 'the', 'Caucasus', ',', 'the', 'Volga', ',',  
 'and', 'the', 'Danube', ',', 'and', 'Southern', 'Italy', '(', 'Molise',  
 ',', 'Campania', ')']  
 ['Hassaniyya', 'Total', 'speakers', '3', ',', '123', ',', '190', '(',  
 '2006', ')']  
 ['Limousin', 'Language', 'codes', 'Linguasphere', 'Approximate',  
 'distribution', 'of', 'Limousin', 'within', 'the', 'Occitan', 'area']  
 ['Meitei', 'Total', 'speakers', '1', ',', '500', ',', '000', '+']  
 ['Meitei', 'Total', 'speakers', '1', ',', '466', ',', '705', 'in', 'India',  
 '(', '2001', 'census', ')']  
 ['Sylheti', 'Total', 'speakers', '10', ',', '300', ',', '000']  
 ['Kemi', 'Sami', 'Language', 'extinction', 'ca', '.', '1900']  
 ['Papi\x3\xa1', 'Total', 'speakers', '5', ',', '000']  
 ['Niuean', 'Total', 'speakers', '8', ',', '000']  
 ["Mi'kmaq", 'Total', 'speakers', '10', ',', '850', '(', '8', ',', '750',  
 'in', 'Canada', ',', '2', ',', '100', 'in', 'the', 'United', 'States', ')']

["Mi'kmaq", 'Region', 'Gaspé', 'Peninsula', '(', 'Quebec', ')']  
 ['Mandaic', 'Region', 'Iraq', '-', 'Baghdad', ',', 'Basra']  
 ['Mandaic', 'Region', 'Iran', '-', 'Kh', '\xc5', '\xab', 'zest', '\xc4',  
 '\x81', 'n']  
 ['Mandaic', 'Language', 'codes', 'ISO', '639-3', ':', 'mid', '\xe2',  
 '\x80', '\x93', 'Modern', 'Mandaic']  
 ['Mandaic', 'Language', 'codes', 'ISO', '639-3', 'myz', '\xe2', '\x80',  
 '\x93', 'Classical', 'Mandaic']  
 ['Ammonite', 'Language', 'extinction', '5th', 'century', 'BC']  
 ['Meroitic', 'Language', 'extinction', '~', '400', 'CE']  
 ['Khasi', 'Total', 'speakers', '865', ',', '000', 'mainly', 'in',  
 'Meghalaya', ',', 'India', '(', '1997', ')']  
 ['Santhali', 'Total', 'speakers', '6', ',', '050', ',', '000']  
 ['Galician\xe2\x80\x93Portuguese', 'Language', 'extinction', 'Split',  
 'and', 'evolved', 'into', 'the', 'Galician', 'and', 'Portuguese',  
 'languages']  
 ['Carolinian', 'Total', 'speakers', '5', ',', '700']  
 ['Carolinian', 'Spoken', 'in', 'Northern', 'Mariana', 'Islands']  
 ['Carolinian', 'Official', 'language', 'in', 'Northern', 'Mariana',  
 'Islands']  
 ['Dene', 'Suline', 'Total', 'speakers', '11', ',', '895']  
 ['Forro', 'Total', 'speakers', '69', ',', '899']  
 ['Angolar', 'Total', 'speakers', '5', ',', '000']  
 ['Principense', 'Total', 'speakers', '200', ',', '4000']  
 ['Macanese', 'Total', 'speakers', 'Unknown', ';', 'perhaps', 'several',  
 'dozen', 'in', 'Macau', 'and', 'Hong', 'Kong', 'and', 'several',  
 'hundreds', 'or', 'more', 'than', 'a', 'thousand', 'among', 'the',  
 'Macanese', 'Diaspora', ';', 'virtually', 'all', 'speakers', 'at', 'least',  
 'bilingual']  
 ['Gwich\xe2\x80\x99in', 'Total', 'speakers', '770', '(', '1998',  
 'Statistics', 'Canada', ')']  
 ['Gwich\xe2\x80\x99in', 'Writing', 'system', 'Latin', '(', 'Northern',  
 'Athabaskan', 'variant', ')']  
 ['Penang', 'Hokkien', 'Total', 'speakers', '1', ',', '824', ',', '741']  
 ['Amorite', 'Language', 'extinction', '2nd', 'millennium', 'BC']  
 ['Dogrib', 'Total', 'speakers', '2', ',', '640']  
 ['Leponitic', 'Language', 'extinction', 'ca', '.', '400', 'BC', '?']  
 ['Messapian', 'Messapic', 'Language', 'extinction', 'around', '1st',  
 'century', 'BC']  
 ['Hurrian', 'Language', 'extinction', 'Ca', '1000', 'BC']  
 ['Ongota', 'Total', 'speakers', '6', '(', '2008', ')']  
 ['Ongota', 'Region', 'Southern', 'Omo', 'Zone', ',', 'Southern']  
 ['Menominee', 'Regulated', 'by', 'Menominee', 'Language', '&', 'amp', ';',  
 'Culture', 'Commission']  
 ['Urantian', 'Language', 'extinction', 'c', '.', '6th', 'century', 'BCE']  
 ['Inuvialuktun', 'Total', 'speakers', '400', '\xe2', '\x80', '\x93', '700']



['Inuvialuktun', 'Regulated', 'by', 'Inuvialuit', 'Cultural', 'Resource',  
 'Centre', 'and', 'Inuit', 'Tapiriit', 'Kanatami']  
 ['Inuinnaqtun', 'Total', 'speakers', 'approximately', '2', ',', '000']  
 ['Inuinnaqtun', 'Regulated', 'by', 'Inuit', 'Tapiriit', 'Kanatami']  
 ['Inuinnaqtun', 'Language', 'codes', 'ISO', '639-3', 'Communities',  
 'where', 'Inuinnaqtun', 'is', 'spoken']  
 ['Semigallian', 'Language', 'extinction', '16th', 'century']  
 ['Extremaduran', 'Total', 'speakers', '200', ',', '000']  
 ['Lezgian', 'Total', 'speakers', '783', ',', '720']  
 ['Lushootseed', 'Region', 'Southern', 'British', 'Columbia', 'into',  
 'northern', 'Washington']  
 ['Punic', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'North', 'Africa',  
 'Tunisia', 'Malta']  
 ["N'Ko", 'Total', 'speakers', 'c', '.', '27', 'million', '(', 'mother',  
 'tongue', 'and', 'second', 'language', ')', ';', 'see', 'Bambara', ',',  
 'Mandinka', ',', 'Maninka', ',', 'and', 'Dioula']  
 ["N'Ko", 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]  
 ["N'Ko", 'Regulated', 'by', 'Kurukan', 'fuwa', 'gbara', '(', 'Association',  
 'Kurukan', 'fuwa', 'gbara', '-', ')']  
 ['Fang', 'Total', 'speakers', '858', ',', '000']  
 ['Swiss', 'Total', 'speakers', '1', '.', '5', 'million']  
 ['Swiss', 'Official', 'language', 'in', 'Switzerland', '(', 'as', 'French',  
 ')']  
 ['Mingrelian', 'Total', 'speakers', '500', ',', '000', '(', '1989', ')']  
 ['Moksha', 'Total', 'speakers', '~', '500', ',', '000']  
 ['Ecclesiastical', 'Latin', 'Spoken', 'in', 'Never', 'spoken', 'as', 'a',  
 'native', 'language', ';', 'other', 'uses', 'vary', 'widely', 'by',  
 'period', 'and', 'location']  
 ['Ecclesiastical', 'Language', 'extinction', 'Still', 'used', 'for',  
 'many', 'purposes']  
 ['Ecclesiastical', 'Region', 'In', 'dioceses', 'of', 'the', 'Roman',  
 'church', 'before', 'Protestantism', ';', 'in', 'dioceses', 'of', 'the',  
 'Catholic', 'Church', 'subsequently']  
 ['Svan', 'Total', 'speakers', '15', ',', '000']  
 ['Svan', 'Writing', 'system', 'Seldom', '(', 'if', 'ever', ')', 'written',  
 ',', 'but', 'Georgian', 'alphabet', 'is', 'used']  
 ['Me\xc3\xa4nkieli', 'Total', 'speakers', '40', ',', '000-70', ',', '000']  
 ["Han'er", 'Language', 'extinction', 'Evolved', 'into', 'Chinese',  
 'dialects']  
 ['Arawak', 'Total', 'speakers', '2', ',', '500']  
 ['Khariboli', 'Region', 'Delhi', ',', 'Western', 'Uttar', 'Pradesh']  
 ['Khariboli', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Philippine', 'Creole', 'Spanish', 'Total', 'speakers', '2', ',', '500',

', ', '000', ';', '7th', 'most', 'spoken', 'native', 'language', 'in',  
 'the', 'Philippines']  
 ['Philippine', 'Region', 'Filipino', 'diaspora', 'and', 'other', 'regions',  
 'with', 'Chavacano', 'communities']  
 ['Massachusett', 'Language', 'extinction', 'Late', '19th', 'century', ', ',  
 'revival', 'efforts', 'underway']  
 ['Massachusett', 'Region', 'Wampanoag', 'people', ', ', 'Massachusett',  
 'people']  
 ['Greenlandic', 'Total', 'speakers', 'approximately', '57', ', ', '000']  
 ['Greenlandic', 'Regulated', 'by', 'Oqaasileriffik', '(', 'The',  
 'Greenland', 'Language', 'Secretariat', ')']  
 ['Seychellois', 'Total', 'speakers', '72', ', ', '700']  
 ['Seychellois', 'Language', 'codes', 'Linguasphere', '51-AAC-cea', '(',  
 '&', 'amp', ';', '-ceb', ')']  
 ['Abenaki', 'Total', 'speakers', '20', 'in', '1991']  
 ['Abenaki', 'Language', 'codes', 'ISO', '639-3', ':', 'aaq', '\xe2',  
 '\x80', '\x93', 'Eastern', 'Abenaki', '(', 'extinct', ')']  
 ['Abenaki', 'Language', 'codes', 'ISO', '639-3', 'abe', '\xe2', '\x80',  
 '\x93', 'Western', 'Abenaki']  
 ['Venetic', 'Language', 'extinction', '1st', 'century', 'AD']  
 ['Bouyei', 'Total', 'speakers', '2', ', ', '649', ', ', '205', '(', '2000',  
 ')']  
 ['Inuktitut', 'Spoken', 'in', 'Canada', '(', 'Nunavut', ', ', 'Quebec', '(',  
 'Nunavik', ')', ', ', 'Northwest', 'Territories', ', ', 'Newfoundland',  
 'and', 'Labrador', '(', 'Nunatsiavut', ')', ')']  
 ['Inuktitut', 'Regulated', 'by', 'Inuit', 'Tapiriit', 'Kanatami', 'and',  
 'various', 'other', 'local', 'institutions']  
 ['Inuktitut', 'Language', 'codes', 'ISO', '639-3', ':', 'ike', '\xe2',  
 '\x80', '\x93', 'Eastern', 'Canadian', 'Inuktitut']  
 ['Hiligaynon', 'Total', 'speakers', '11', 'million', 'total', '(', 'first',  
 'language', ':', '7', 'million', ', ', 'second', 'language', ':', '4',  
 'million', '(', 'est', '.', ')', ')', '3rd', 'most', 'spoken', 'native',  
 'language', 'in', 'the', 'Philippines']  
 ['Hiligaynon', 'Region', 'Western', 'Visayas', 'and', 'Mindanao']  
 ['Hiligaynon', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant',  
 ')', ';']  
 ['Hiligaynon', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Sedang', 'Total', 'speakers', '101', ', ', '434']  
 ['Kapampangan', 'Total', 'speakers', '2', '.', '9', 'million', ';', '6th',  
 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']  
 ['Kapampangan', 'Writing', 'system', 'Latin', '(', 'Kapampangan', 'or',  
 'Spanish', 'variant', ')', ';']  
 ['Kapampangan', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Norman', 'Spoken', 'in', 'England', '\xe2', '\x80', '\xa0', '(', 'See',

':' , 'Norman' , 'England' , ')']  
 ['Norman' , 'Spoken' , 'in' , 'Ireland' , '\xe2' , '\x80' , '\xa0' , '(' , 'See' , ':' , 'Norman' , 'Ireland' , ')']  
 ['Norman' , 'Language' , 'codes' , 'ISO' , '639-3' , 'xno' , '(' , 'Anglo-Norman' , ')']  
 ['Norman' , 'Language' , 'codes' , 'Linguasphere' , '51-AAA-hc' , '&' , 'amp' , ';' , '51-AAA-hd']  
 ['Norman' , 'Language' , 'codes' , 'Linguasphere' , 'Areas' , 'where' , 'the' , 'Norman' , 'language' , 'is' , 'strongest' , 'include' , 'Jersey']  
 ['Gallo' , 'Total' , 'speakers' , '28' , ',' , '000']  
 ['Gallo' , 'Language' , 'codes' , 'Linguasphere' , 'The' , 'Gallo' , 'language' , 'area' , '(' , 'in' , 'grey' , ')', 'corresponding' , 'to' , 'Upper' , 'Brittany' , '(' , 'in' , 'colours' , ',' , 'Lower' , 'Brittany' , ')']  
 ['Middle' , 'French' , 'Language' , 'extinction' , 'evolved' , 'into' , 'Modern' , 'French' , 'by' , 'the' , 'early' , '17th' , 'century']  
 ['Judeo-Arabic' , 'Language' , 'codes' , 'ISO' , '639-3' , ':' , 'yhd' , '\xe2' , '\x80' , '\x93' , 'Judeo-Iraqi' , 'Arabic']  
 ['Judeo-Arabic' , 'Language' , 'codes' , 'ISO' , '639-3' , 'aju' , '\xe2' , '\x80' , '\x93' , 'Judeo-Moroccan' , 'Arabic']  
 ['Judeo-Arabic' , 'Language' , 'codes' , 'ISO' , '639-3' , 'yud' , '\xe2' , '\x80' , '\x93' , 'Judeo-Tripolitanian' , 'Arabic']  
 ['Judeo-Arabic' , 'Language' , 'codes' , 'ISO' , '639-3' , 'ajt' , '\xe2' , '\x80' , '\x93' , 'Judeo-Tunisian' , 'Arabic']  
 ['Judeo-Arabic' , 'Language' , 'codes' , 'ISO' , '639-3' , 'jye' , '\xe2' , '\x80' , '\x93' , 'Judeo-Yemeni' , 'Arabic']  
 ['Zazaki' , 'Total' , 'speakers' , 'About' , '1' , '.' , '5' , '\xe2' , '\x80' , '\x93' , '2' , '.' , '5' , 'million' , 'in' , 'Turkey' , ',' , 'unknown' , 'numbers' , 'elsewhere']  
 ['Zazaki' , 'Language' , 'codes' , 'ISO' , '639-3' , ':' , 'zza' , '\xe2' , '\x80' , '\x93' , 'Zazaki' , '(' , 'generic' , ')']  
 ['Zazaki' , 'Language' , 'codes' , 'ISO' , '639-3' , 'diq' , '\xe2' , '\x80' , '\x93' , 'Dimli' , '(' , 'Southern' , 'Zazaki' , ')']  
 ['Zazaki' , 'Language' , 'codes' , 'ISO' , '639-3' , 'kiu' , '\xe2' , '\x80' , '\x93' , 'Kirmanjki' , '(' , 'Northern' , 'Zazaki' , ')']  
 ['Zazaki' , 'Language' , 'codes' , 'Linguasphere' , '58-AAA-ba' , 'The' , 'regions' , 'where' , 'Zazaki' , 'is' , 'spoken' , 'in' , 'Turkey' , '.' , '(' , 'With' , 'three' , 'main' , 'dialect' , 'areas' , ';' , 'Dersim' , ',' , 'Palu-Bingol' , ',' , 'and' , 'Siverek' , ')', ',' , 'diasporic' , 'in' , 'Kars' , ',' , 'Sar' , '\xc4' , '\xb1' , 'z' , ',' , 'Aksaray' , ',' , 'and' , 'Taraz' , ')']  
 ['Mycenaean' , 'Greek' , 'Language' , 'extinction' , '12th' , 'century' , 'BC']  
 ['Mycenaean' , 'Language' , 'codes' , 'ISO' , '639-3' , 'Map' , 'of' , 'Greece' , 'as' , 'described' , 'in' , 'Homer's' , 'Iliad' , '.' , 'The' , 'geographical' , 'data' , 'is' , 'believed' , 'to' , 'refer' , 'primarily' , 'to' , 'Bronze' , 'Age' , 'Greece']  
 ['Mycenaean' , 'Language' , 'codes' , 'ISO' , '639-3' , 'when' , 'Mycenaean' , 'Greek' , 'would' , 'have' , 'been' , 'spoken']

['Guern\x3\xa9siais', 'Total', 'speakers', '1', ',', '327', '(', '2', '%', ')', 'fluent']  
 ['Elfdalian\x3\x96vdalian', 'Total', 'speakers', 'Around', '3000']  
 ['Elfdalian\x3\x96vdalian', 'Region', 'Elfdalia', '(', 'Dalecarlia', 'County', ')', '(', 'Northern', 'Europe', ')']  
 ['Elfdalian\x3\x96vdalian', 'Writing', 'system', 'Latin', 'Alphabet', '(', 'Elfdalian', 'variant', ')', ',', 'Historically', 'Dalecarlian', 'runes']  
 ['Bats', 'Total', 'speakers', '3', ',', '420', '(', '2000', 'WCD', ')']  
 ['Andalusian', 'Arabic', 'Language', 'extinction', '17th', 'century']  
 ['Andalusian', 'Language', 'extinction', '(', 'after', 'the', 'expulsion', 'from', 'Spain', 'and', 'Portugal', ')']  
 ['Egyptian', 'Arabic', 'Total', 'speakers', '54', ',', '000', ',', '000', '(', 'native', 'speakers', 'only', ')']  
 ['Yemeni', 'Language', 'codes', 'ISO', '639-3', ':', 'ayh', '\xe2', '\x80', '\x93', 'Hadrami', 'Arabic']  
 ['Yemeni', 'Language', 'codes', 'ISO', '639-3', 'ayn', '\xe2', '\x80', '\x93', 'Sanaani', 'Arabic']  
 ['Yemeni', 'Language', 'codes', 'ISO', '639-3', 'acq', '\xe2', '\x80', '\x93', 'Ta'izzi-Adeni', 'Arabic']  
 ['Washo', 'Total', 'speakers', 'USA', ':', '252', '(', '2000', ')']  
 ['Washo', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Washo', 'language']  
 ['Levantine', 'Total', 'speakers', '35', ',', '000', ',', '000']  
 ['Levantine', 'Language', 'codes', 'ISO', '639-3', ':', 'apc', '\xe2', '\x80', '\x93', 'North', 'Levantine', 'Arabic']  
 ['Levantine', 'Language', 'codes', 'ISO', '639-3', 'ajp', '\xe2', '\x80', '\x93', 'South', 'Levantine', 'Arabic']  
 ['(Eastern)', 'Total', 'speakers', 'less', 'than', '60']  
 ['\xc7\x80Xam', 'Language', 'extinction', '19', '?', '?']  
 ['Seroa', 'Language', 'extinction', '19', '?', '?']  
 ['Ju\xc7\x80\xca\xbcchoan', 'Total', 'speakers', '35', ',', '000']  
 ['K'iche'', 'Total', 'speakers', 'approx', '.', '1', ',', '000', ',', '000']  
 ['Jiwarli', 'Language', 'extinction', 'April', '1986', 'with', 'the', 'death', 'of', 'Jack', 'Butler']  
 ['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Total', 'speakers', '6', ',', '000']  
 ['\xc7\x82Kx\xca\xbc0\xc7\x81\xca\xbccein', 'Total', 'speakers', '7', ',', '000']  
 ['Naro', 'Total', 'speakers', '14', ',', '000']  
 ['Kxoe', 'Total', 'speakers', '11', ',', '000']  
 ['A'tong', 'Total', 'speakers', '15', ',', '000']  
 ['A'tong', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Aari', 'Total', 'speakers', '285', ',', ',', '473']  
 ['Abanyom', 'Total', 'speakers', '12', ',', ',', '500']  
 ['Abau', 'Total', 'speakers', '4', ',', ',', '500', '\xe2', '\x80', '\x93', '5',  
 ',', ',', '000']  
 ['Abau', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Hixkary\x3\xa1na', 'Total', 'speakers', '500', '\xe2', '\x80', '\x93',  
 '600']  
 ['Hixkary\x3\xa1na', 'Region', 'Upper', 'Nhamund', '\xc3', '\xa1',  
 'River', ',', ',', 'Amazonas']  
 ['Xiri', 'Spoken', 'in', 'South', 'Africa', ',', ',', 'Namibia']  
 ['Tama', 'Total', 'speakers', 'approximately', '63', ',', ',', '000', '(',  
 '1993', ')']  
 ['Korana', 'Spoken', 'in', 'South', 'Africa', ',', ',', 'maybe', 'Botswana']  
 ['Kwadi', 'Language', 'extinction', 'ca', '.', '1980']  
 ['Sandawe', 'Total', 'speakers', 'c', '.', '40', ',', ',', '000']  
 ['Hadza', 'Total', 'speakers', '975', '(', 'ethnicity', ',', ',', '2005', ')']  
 ['Proto-Mandarin', 'Language', 'extinction', 'Evolved', 'into', 'Mandarin',  
 'Chinese', 'and', 'possibly', 'other', 'and', 'Chinese', 'dialects']  
 ['Sudanese', 'Arabic', 'Total', 'speakers', '18', ',', ',', '986', ',', ',', '000']  
 ['Ab\x3\xa9', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']  
 ['Ab\x3\xa9', 'Total', 'speakers', '170', ',', ',', '000', '(', '1995', ')']  
 ['Awjilah', 'Total', 'speakers', '2', ',', ',', '000', '(', '1993', ')']  
 ['Guanche', 'Language', 'extinction', '?', 'Possibly', 'around', 'the',  
 '18th', 'century']  
 ['Gilbertese', 'Total', 'speakers', '102', ',', ',', '000']  
 ['Gilbertese', 'Regulated', 'by', 'Kiribati', 'Language', 'Board']  
 ['Montenegrin', 'Regulated', 'by', 'Board', 'for', 'Standardization', 'of',  
 'the', 'Montenegrin', 'Language']  
 ['Montenegrin', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-  
 g']  
 ['Hui', 'Total', 'speakers', '4', '.', '6', 'million']  
 ['Walser', 'Total', 'speakers', 'about', '22', ',', ',', '000']  
 ['Marshallese', 'Total', 'speakers', '43', ',', ',', '900', '(', '1979', ')']  
 ['Marshallese', 'Official', 'language', 'in', 'Marshall', 'Islands', '(',  
 'with', 'English', ')']  
 ['Tlingit', 'Total', 'speakers', '845', '(', 'Krauss', '1995', ')']  
 ['M\x3\xb3cheno', 'Total', 'speakers', '2', ',', ',', '200', '&', 'gt', ';']  
 ['Ingush', 'Total', 'speakers', '413', ',', ',', '000', '(', '2002', ')']  
 ['Altay', 'Total', 'speakers', '67', ',', ',', '900', '(', '2002', 'census',  
 ')']  
 ['Altay', 'Region', 'Altai', 'Republic', '(', 'Southern', 'Altay', ')',  
 ',', ',', 'Altai', 'Krai', '(', 'Northern', 'Altay', ')']  
 ['Altay', 'Language', 'codes', 'ISO', '639-3', ':', 'atv', '\xe2', '\x80',  
 '\x93', 'Northern', 'Altai']  
 ['Altay', 'Language', 'codes', 'ISO', '639-3', 'alt', '\xe2', '\x80',  
 '\x93', 'Southern', 'Altai']

['Samaritan', 'Spoken', 'in', 'Israel', 'and', 'Palestinian', 'Authority',  
 'territories', ',', 'predominantly', 'in', 'Nablus', 'and', 'Holon']  
 ['Fur', 'Total', 'speakers', '744', ',', '000', '(', '2004', ')']  
 ['Fon', 'Total', 'speakers', '1', '.', '7', 'million']  
 ['Chuvash', 'Total', 'speakers', '1', ',', '640', ',', '000', '(', 'in',  
 'Russia', ')', '(', 'Russian', 'Census', '(', '2002', ')', ',', 'self-  
 reported', 'speakers', ')']  
 ['Jakaltek', 'Total', 'speakers', 'approx', '.', '88', ',', '800']  
 ['Judeo-Italian', 'Language', 'codes', 'Linguasphere', '51-AAB-be', '&',  
 'amp', ';', '-bf']  
 ['Chan\xc3\xa9', 'Total', 'speakers', 'Extinct', 'since', '1700s']  
 ['Biblical', 'Total', 'speakers', 'Extinct', 'as', 'a', 'regularly',  
 'spoken', 'language', 'by', 'the', '4th', 'century', 'CE', ',', 'but',  
 'survived', 'as', 'a', 'liturgical', 'and', 'literary', 'language']  
 ['Biblical', 'Writing', 'system', 'Proto-Canaanite', '/', 'Phoenician',  
 'alphabet']  
 ['\xc3\x84ynu', 'Total', 'speakers', '6', ',', '570', '(', '2000', ')']  
 ['Santiam', 'Language', 'extinction', 'mid', '19th', 'century']  
 ['Santiam', 'Language', 'codes', 'ISO', '639-3', 'kyl', '\xe2', '\x80',  
 '\x93', '(', 'Kalapuya', ')']  
 ['Mossi', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Mossi', 'Total', 'speakers', '5', ',', '036', ',', '700']  
 ['Idiom', 'Category', '(', 'purpose', ')', 'Idiom', 'Neutral']  
 ['Lacandon', 'language', 'Total', 'speakers', 'approx', '.', '1000']  
 ['Neapolitan', 'Total', 'speakers', '7', '.', '0', 'million', '(', 'est',  
 '.', ')']  
 ['Noric', 'Language', 'extinction', 'unknown', ';', 'later', 'than', 'the',  
 '2nd', 'century', 'AD']  
 ['Shawia', 'Total', 'speakers', '2', ',', '000', ',', '000']  
 ['Shawia', 'Region', 'Aur', '\xc3', '\xa8', 's', '(', 'Batna', ',',  
 'Khenchela', ',', 'Oum', 'El', 'Bouaghi', ',', 'Souk', 'Ahras', ',', 'T',  
 '\xc3', '\xa9', 'bessa', ')']  
 ['Sened', 'Language', 'extinction', '20th', 'century']  
 ['Mandan', 'Region', 'Fort', 'Berthold', 'Reservation', ',', 'North',  
 'Dakota']  
 ['Crow', 'Total', 'speakers', '4', ',', '280']  
 ['Konkani', 'Total', 'speakers', '3', '.', '6', 'million']  
 ['Konkani', 'Region', 'Konkani', 'is', 'also', 'spoken', 'in', 'Kenya']  
 ['Konkani', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'native', 'Konkani', 'speakers', 'in', 'India', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Nuosu', 'Total', 'speakers', '2', 'million', '(', '2000', 'census', ')']  
 ['Tujia', 'Spoken', 'in', '8', 'million', 'Tujia', 'people']  
 ['Tujia', 'Total', 'speakers', '200', ',', '000']

['Kabyle', 'Total', 'speakers', '3', ',', '126', ',', '000', '(', '1995', ')', ',', '2', '.', '5', 'million', 'in', 'Algeria']  
 ['Kabyle', 'Language', 'codes', 'ISO', '639-3', 'kab', 'Kabyle-speaking', 'areas', '(', 'in', '1936', ')']  
 ['Crimean', 'Tatar', 'Total', 'speakers', '483', ',', '990']  
 ['Crimean', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Crimean', 'Language', 'codes', 'Linguasphere', 'Crimean', 'Tatar-speaking', 'world']  
 ['Ripuarian', 'Total', 'speakers', '~', '900', ',', '000']  
 ['Ripuarian', 'Language', 'codes', 'ISO', '639-3', 'Area', 'where', 'Ripuarian', 'is', 'spoken', '.', 'Green', '=', 'sparsely', 'populated', 'forest']  
 ['Ripuarian', 'Language', 'codes', 'ISO', '639-3', 'Central', 'German', 'language', 'area', '.', '1', '=', 'Ripuarian']  
 ['Gorani', 'Total', 'speakers', 'Between', '2', '-', '5', 'million']  
 ['Raetic', 'Language', 'extinction', '3rd', 'century']  
 ['Palaic', 'Language', 'extinction', 'around', '1300', 'BC']  
 ['Aja', 'Spoken', 'in', 'Sudan', ',', 'Central', 'African', 'Republic']  
 ['Aja', 'Total', 'speakers', '200', '(', '1993', ')']  
 ['Bantik', 'Total', 'speakers', '~', '10', ',', '000']  
 ['Manado', 'Malay', 'Total', 'speakers', '850', ',', '000', '(', '2001', ')']  
 ['Qashqai', 'Total', 'speakers', '1', ',', '500', ',', '000', '(', '1997', ')']  
 ['Qashqai', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Nogai', 'Total', 'speakers', '90', ',', '000']  
 ['Late', 'Latin', 'Spoken', 'in', 'Roman', 'Empire']  
 ['Late', 'Language', 'extinction', 'developed', 'into', 'Medieval', 'Latin']  
 ['Late', 'Writing', 'system', 'inscriptions', '.', 'Spoken', 'and', 'written', 'Lingua', 'Franca']  
 ['Late', 'Official', 'language', 'in', 'Western', 'Roman', 'Empire']  
 ['Lori', 'Total', 'speakers', 'ca', '.', '3', '.', '3', 'million']  
 ['Lori', 'Region', 'Southern', 'Zagros', '(', 'Mainly', ':', 'Lorestan', 'province', '.', ')']  
 ['Lori', 'Language', 'codes', 'ISO', '639-3', ':', 'lrc', '\xe2', '\x80', '\x93', 'Northern', 'Luri']  
 ['Lori', 'Language', 'codes', 'ISO', '639-3', 'luz', '\xe2', '\x80', '\x93', 'Southern', 'Luri']  
 ['Old', 'Saxon', 'Language', 'extinction', 'developed', 'into', 'Middle', 'Low', 'German', 'in', 'the', '12th', 'century']  
 ['Shenwa', 'Total', 'speakers', '4', ',', '764']  
 ['Warlpiri', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Pitkern', 'Total', 'speakers', '&', 'lt', ';', '100']

['Pitkern', 'Spoken', 'in', 'Pitcairn', 'Islands', ',', 'New', 'Zealand',  
 'and', 'Norfolk', 'Island']  
 ['Bikya', 'Total', 'speakers', '1', '(', '1986', 'Breton', ')']  
 ['Wakhi', 'Spoken', 'in', 'Afghanistan', '(', '9', ',', '600', ')', ',',  
 'China', '(', '6', ',', '000', ')', ',', 'Pakistan', '(', '9', ',', '100',  
 ')', ',', 'Tajikistan', '(', '7', ',', '000', ')']  
 ['Wakhi', 'Total', 'speakers', 'approximately', '31', ',', '700']  
 ['Siberian', 'Yupik', ',', 'Yuit', 'Total', 'speakers', 'approximately',  
 '1', ',', '350']  
 ['Siberian', 'Spoken', 'in', 'United', 'States', ',', 'Russian',  
 'Federation']  
 ['Spanish', 'Total', 'speakers', '2', ',', '660', 'native', 'Spanish',  
 'speakers', '(', '1990', 'census', ')']  
 ['Spanish', 'Total', 'speakers', '439', ',', '000', 'native', 'Spanish',  
 'speakers', '(', 'Moreno', 'and', 'Otero', '2007', ')']  
 ['Spanish', 'Total', 'speakers', '3', ',', '016', ',', '773', 'total',  
 'speakers', '(', 'Spanish', 'and', 'Chavacano', 'as', 'a', '1st', ',',  
 '2nd', ',', '3rd', ',', '4th', 'or', '5th', 'language', ')']  
 ['Soddo', 'Total', 'speakers', '250', ',', '000']  
 ['Iberian', 'Language', 'extinction', '1st', '\xe2', '\x80', '\x93', '2nd',  
 'century', 'AD']  
 ['Iberian', 'Region', 'Mediterranean', 'coast', 'of', 'the', 'Iberian',  
 'Peninsula']  
 ['Middle', 'Persian', 'Language', 'extinction', 'evolved', 'into',  
 'Modern', 'Persian', 'by', 'the', '9th', 'century']  
 ['Bashkardi', 'Total', 'speakers', '7', ',', '000']  
 ['Medieval', 'Latin', 'Spoken', 'in', 'Numerous', 'small', 'states']  
 ['Medieval', 'Language', 'extinction', 'replaced', 'by', 'Renaissance',  
 'Latin']  
 ['Medieval', 'Writing', 'system', 'Latin', 'alphabet', 'Spoken', 'and',  
 'written', 'lingua', 'franca']  
 ['Bench', 'Total', 'speakers', '347', ',', '636', 'of', 'Bench', 'dialect']  
 ['Bench', 'Total', 'speakers', '8', ',', '159', 'of', 'Mer', 'dialect']  
 ['Bench', 'Total', 'speakers', '492', 'of', 'She', 'dialect']  
 ['Bench', 'Region', 'Bench', 'Maji', 'Zone', ',', 'SNNPR']  
 ['Kwasio', 'Total', 'speakers', '18', ',', '000']  
 ['Kwasio', 'Spoken', 'in', 'Cameroon', ',', 'Equatorial', 'Guinea']  
 ['Kwasio', 'Region', 'along', 'the', 'coast', 'at', 'the', 'border',  
 'between', 'Cameroon', 'and', 'Equatorial', 'Guinea']  
 ['Waray-Waray', 'Total', 'speakers', '3', '.', '4', 'million', 'total',  
 '(', 'first', 'language', ':', '3', '.', '1', 'million', ',', 'second',  
 'language', ':', '250', ',', '000-300', ',', '000', '(', 'est', '.', ')',  
 ')', ';', '5th', 'most', 'spoken', 'native', 'language', 'in', 'the',  
 'Philippines']  
 ['Waray-Waray', 'Region', 'Eastern', 'Visayas', ',', 'some', 'parts', 'of',  
 'Masbate']



['Waray-Waray', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')', ';']  
 ['Waray-Waray', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Kinaray-a', 'Total', 'speakers', '1', ',', '051', ',', '968']  
 ['Kinaray-a', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Taus\xc5\xabg', 'Total', 'speakers', '1', ',', '022', ',', '000']  
 ['Taus\xc5\xabg', 'Total', 'speakers', '(', '900', ',', '000', 'in', 'the', 'Philippines', ')']  
 ['Taus\xc5\xabg', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Maranao', 'Total', 'speakers', '1', '.', '15', 'million']  
 ['Maranao', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')', ';']  
 ['Maranao', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Maguindanao', 'Total', 'speakers', '1', '.', '1', 'million']  
 ['Maguindanao', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')', ';']  
 ['Maguindanao', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Laal', 'Total', 'speakers', '749', '(', '2000', ')']  
 ['Laal', 'Language', 'codes', 'ISO', '639-3', 'Location', 'within', 'Chad', 'where', 'the', 'Laal', 'language', 'is', 'spoken']  
 ['Gulf', 'Spoken', 'in', 'Oman', '&', 'amp', ';', 'south-western', 'Iran']  
 ['Shabo', 'Total', 'speakers', '400', '\xe2', '\x80', '\x93', '500', '(', '2000', ')']  
 ['Nepal', 'Bhasa', 'Total', 'speakers', '0', '.', '8', 'million', 'approx']  
 ['Nepal', 'Regulated', 'by', 'Nepal', 'Bhasa', 'Academy']  
 ['Nepal', 'Regulated', 'by', 'Nepal', 'Bhasa', 'Parishad']  
 ['Scythian', 'Language', 'extinction', 'mostly', 'extinct', 'by', 'AD', '1000', ',', 'remnants', 'evolved', 'into', 'Ossetic']  
 ['Hunnic', 'Language', 'extinction', 'after', '453', 'CE']  
 ['Nuer', 'Region', 'Upper', 'Nile', 'state', ',', 'Gambela']  
 ['Beothuk', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'Beothuk', 'language']  
 ['Shambala', 'Total', 'speakers', '664', ',', '000']  
 ['Bagirmi', 'Total', 'speakers', '44', ',', '761']  
 ['Faliscan', 'Language', 'extinction', 'about', '150', 'BC']  
 ['Faliscan', 'Writing', 'system', 'Etruscan', 'and', 'Roman', 'alphabets']  
 ['Yola', 'Language', 'extinction', 'Mid-19th', 'century']  
 ['New', 'Signed', 'in', 'New', 'Zealand']  
 ['New', 'Total', 'signers', '24', ',', '000', 'as', 'of', '2006']  
 ['New', 'Language', 'codes', 'ISO', '639-3', 'nzs', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',

'\x94', 'legal', 'recognition']  
 ['Yaghnobi', 'Total', 'speakers', '12', ',', '500']  
 ['Yaghnobi', 'Region', 'in', '1970s', 'relocated', 'to', 'Zafarobod']  
 ['Yaghnobi', 'Region', 'in', '1990s', 'some', 'speakers', 'returned',  
 'back', 'to', 'Yaghnob']  
 ['Livvi-Karelian', 'language', 'Total', 'speakers', '14', ',', '100', '(',  
 '2000', ')']  
 ['Livvi-Karelian', 'Writing', 'system', 'Latin', 'alphabet', '(',  
 'Karelian', 'variant', ')']  
 ['Shan', 'Total', 'speakers', '3', '.', '3', 'million']  
 ['Kujarg\xc3\xa9', 'Total', 'speakers', '1', ',', '000', '(', '1983', ')']  
 ['Ibanag', 'Total', 'speakers', '300', ',', '000-500', ',', '000']  
 ['Amami', 'Total', 'speakers', '30', ',', '000']  
 ['Amami', 'Language', 'codes', 'ISO', '639-3', ':', 'ryn', '\xe2', '\x80',  
 '\x93', 'Northern', 'Amami']  
 ['Amami', 'Language', 'codes', 'ISO', '639-3', 'ams', '\xe2', '\x80',  
 '\x93', 'Southern', 'Amami']  
 ['Miyako', 'Total', 'speakers', '67', ',', '700']  
 ['Yaeyama', 'Total', 'speakers', '44', ',', '650']  
 ['Yucatec', 'Maya', 'Total', 'speakers', '800', ',', '000', '-', '1', ',',  
 '200', ',', '000']  
 ['Yucatec', 'Region', 'Belize', '5', ',', '000']  
 ['Yucatec', 'Region', 'Yucat', '\xc3', '\xa1', 'n', '547']  
 ['Griko', 'Total', 'speakers', '20', ',', '000']  
 ['Griko', 'Language', 'codes', 'Linguasphere', '56-AAA-ai', 'Location',  
 'map', 'of', 'the', 'Griko-speaking', 'areas', 'in', 'Salento', 'and',  
 'Calabria']  
 ['Kham,', 'Khamkura', ',', 'Kamkura', 'Total', 'speakers', '65', ',',  
 '000']  
 ['Kham,', 'Language', 'codes', 'ISO', '639-3', ':', 'kif', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Parbate', 'Kham']  
 ['Kham,', 'Language', 'codes', 'ISO', '639-3', 'kjl', '\xe2', '\x80',  
 '\x93', 'Western', 'Parbate', 'Kham']  
 ['Oropom', 'Language', 'extinction', 'ca', '.', '1950']  
 ['T\xc3\xafcuna', 'Total', 'speakers', '25', ',', '000', 'in', 'Brazil',  
 '(', '2000', 'SIL', ')', '.', 'Population', 'total', 'all', 'countries',  
 ':', '41', ',', '000']  
 ['T\xc3\xafcuna', 'Region', 'West', 'Amazonas', '.', 'Also', 'spoken',  
 'in', 'Colombia', ',', 'Peru', '.', 'Coordinates']  
 ['T\xc3\xafcuna', 'Region', '3', '\xc2', '\xb0', '15', '\xe2', '\x80',  
 '\xb2', 'S', '68', '\xc2', '\xb0', '35', '\xe2', '\x80', '\xb2', 'W',  
 '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '3', '.', '25',  
 '\xc2', '\xb0', 'S', '68', '.', '583', '\xc2', '\xb0', 'W', '\xef', '\xbb',  
 '\xbf', '/', '-3', '.', '25', ';', '-68', '.', '583']  
 ['Amdang', 'Total', 'speakers', '41', ',', '069']  
 ['Amdang', 'Language', 'codes', 'ISO', '639-3', 'Language', 'map', 'of',

'Amdang']  
 ['Inupiaq', 'Total', 'speakers', 'approximately', '2', ',', '100']  
 ['Inupiaq', 'Region', 'Alaska', ';', 'formerly', 'Big', 'Diomed', 'Island']  
 ['Inupiaq', 'Language', 'codes', 'ISO', '639-3', ':', 'esi', '\xe2', '\x80', '\x93', 'North', 'Alaskan', 'Inupiatun']  
 ['Manyika', 'Total', 'speakers', '1', '.', '0', 'million']  
 ['Komi', 'Total', 'speakers', '293', ',', '000', '(', 'Komi-Zyrian', ',', 'census', '2002', ')', ',', '94', ',', '300', '(', 'Komi-Permyak', ',', 'census', '2002', ')']  
 ['Komi', 'Region', 'Komi', 'Republic', ',', 'Perm', 'Krai', '(', 'Komi-Permyak', 'Okrug', ',', 'Krasnovishersky', 'Raion', ')']  
 ['Yevanic', 'Total', 'speakers', '&', 'lt', ';', '50']  
 ['Shtokavian', 'Total', 'speakers', '~', '13', 'million']  
 ['Shtokavian', 'Standard', 'forms', 'Standard', 'Serbo-Croatian', '(', 'defunct', ')']  
 ['Shtokavian', 'Language', 'codes', 'Linguasphere', '53-AAA-ga', 'to', '-gf', '&', 'amp', ';']  
 ['Shtokavian', 'Language', 'codes', 'Linguasphere', '53-AAA-gi', '(', '-gia', 'to', '-gii', ')']  
 ['Shtokavian', 'Language', 'codes', 'Linguasphere', 'Shtokavian', 'subdialects', '(', 'Pavle', 'Ivi', '\xc4', '\x87', '1988', ')']  
 ['Kajkavian', 'Total', 'speakers', '~', '1', '.', '3', 'million']  
 ['Chakavian', 'Total', 'speakers', '~', '660', ',', '000', '(', '2001', ')']  
 ['Jamtlandic', 'Total', 'speakers', '~', '50', ',', '000']  
 ['Bajan', 'Total', 'speakers', '259', ',', '000']  
 ['Early', 'Modern', 'English', 'Language', 'extinction', 'developed', 'into', 'Modern', 'English']  
 ['Early', 'Spoken', 'in', 'Wales', 'and', 'British', 'colonies']  
 ['Tianjin', 'Spoken', 'in', 'Urban', 'Tianjin', 'in', 'the', 'People's', 'Republic', 'of', 'China']  
 ['Old', 'Japanese', 'Language', 'extinction', 'Evolved', 'into', 'Early', 'Middle', 'Japanese', 'during', 'the', 'Heian', 'period']  
 ['Old', 'Writing', 'system', 'Chinese', 'characters', 'and', 'Man'y", '\xc5', '\x8d', 'gana']  
 ['Goguryeo', 'Language', 'extinction', '7th', 'century', '?']  
 ['Goguryeo', 'Language', 'codes', 'ISO', '639-3', 'The', 'Three', 'Kingdoms', 'of', 'Korea', ',', 'with', 'Goguryeo', 'and', 'Buyeo', 'in', 'blue']  
 ['Guugu', 'Total', 'speakers', '200', '\xe2', '\x80', '\x93', '300']  
 ['Kala', 'Spoken', 'in', 'Western', 'and', 'Central', 'Torres', 'Strait', 'Islands', ',', 'Queensland']  
 ['Antakarinya', 'Spoken', 'in', 'northeast', 'area', 'of', 'South', 'Australia']  
 ['Antakarinya', 'Total', 'speakers', 'possibly', '50']

['Yankunytjatjara', 'Total', 'speakers', '200', '\xe2', '\x80', '\x93', '300']

['Ghotuo', 'Total', 'speakers', '9', ',', '000']

['Roma', 'Total', 'speakers', '1', ',', '700', '(', 'as', 'of', '1991', ')']

['Herero', 'Total', 'speakers', '237', ',', '000']

['Choctaw', 'Total', 'speakers', '17', ',', '890']

['Choctaw', 'Region', 'Southeastern', 'Oklahoma', 'and', 'east', 'central', 'Mississippi', ',', 'and', 'into', 'Louisiana', 'and', 'Tennessee']

['Choctaw', 'Language', 'codes', 'ISO', '639-3', 'Choctaw', 'language', 'spread', 'in', 'the', 'United', 'States']

['Kutchi', 'Total', 'speakers', '866', ',', '000']

['Kutchi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Kumzari', 'Total', 'speakers', 'Less', 'than', '10', ',', '000', '(', 'Ethnologue', 'cites', '1', ',', '700', ')']

['Sogdian', 'Language', 'extinction', 'largely', 'extinct', 'by', 'the', '9th', 'century', ',', 'remnants', 'evolved', 'into', 'Yaghnobi']

['Kok-borok', 'Total', 'speakers', '950', ',', '000', '+']

['Kok-borok', 'Total', 'speakers', '854', ',', '023', 'in', 'India', '(', '2001', ')', ';', '105', ',', '000', 'in', 'Bangladesh', '(', '1993', ')']

['Gikuyu', 'Kikuyu', 'Total', 'speakers', 'about', '8', ',', '511', ',', '876', '(', 'Ethnologue', '2009', ')']

['Saurashtra', 'Total', 'speakers', '510', ',', '000', '(', '1997IMA', ')']

['Saurashtra', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Renaissance', 'Latin', 'Spoken', 'in', 'The', 'administrations', 'and', 'universities', 'of', 'numerous', 'countries']

['Renaissance', 'Language', 'extinction', 'developed', 'into', 'New', 'Latin', 'by', '16th', 'century']

['Renaissance', 'Writing', 'system', 'Latin', 'alphabet', 'Spoken', 'and', 'written', 'administrative', ',', 'liturgical', 'and', 'literary', 'language']

['Chol', 'Total', 'speakers', 'approx', '.', '100', ',', '000']

['Kalanga', 'Total', 'speakers', '850', ',', '000']

['Penang', 'Total', 'signers', '1000', 'older', 'signers', ',', 'most', 'bilingual', 'in', 'Malaysian', 'Sign']

['Penang', 'Language', 'codes', 'ISO', '639-3', 'psg', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Selangor', 'Total', 'signers', '500', ',', 'mostly', 'elderly', ',', 'all', 'bilingual', 'in', 'Malaysian', 'Sign']

['Selangor', 'Language', 'codes', 'ISO', '639-3', 'kgi', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Malaysian', 'Total', 'signers', '24', ',', '000']  
 ['Malaysian', 'Language', 'codes', 'ISO', '639-3', 'xml', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Beijing', 'Spoken', 'in', 'Beijing', 'and', 'surrounding', 'areas']  
 ['Beijing', 'Language', 'codes', 'ISO', '639-3', 'cmn-bei', '(', 'Beijing',  
 'area', ')']  
 ['Beijing', 'Language', 'codes', 'ISO', '639-3', 'cmn-bej', '(', 'Beijing',  
 'proper', ')']  
 ['Beijing', 'Language', 'codes', 'ISO', '639-3', 'Beijing', 'dialect',  
 'in', 'the', 'broad', 'sense']  
 ['Dongxiang', 'Total', 'speakers', '250', ',', '000']  
 ['Dongxiang', 'Region', 'mainly', 'in', 'Linxia', 'Hui', 'Autonomous',  
 'Prefecture']  
 ['Dongxiang', 'Region', 'and', 'Ili', 'Kazakh', 'Autonomous', 'Prefecture',  
 'in', 'Xinjiang', 'Uyghur', 'Autonomous']  
 ['Cantonese', 'Spoken', 'in', 'China', ':', 'central', 'and', 'western',  
 'Guangdong', '(', 'Zhongshan', ',', 'Foshan', ',', 'Shenzhen', ')', ',',  
 'the', 'Pearl', 'River', 'Delta', ',', 'Hainan', ',', 'and', 'the',  
 'eastern', 'and', 'southern', 'part', 'of', 'Guangxi', '(', 'Wuzhou', ')']  
 ['Cantonese', 'Spoken', 'in', 'Australia', ':', 'Sydney', ',', 'Melbourne']  
 ['Cantonese', 'Spoken', 'in', 'United', 'States', ':', 'New', 'York',  
 'City']  
 ['Cantonese', 'Spoken', 'in', ':', 'Ho', 'Chi', 'Minh', 'City']  
 ['Cantonese', 'Official', 'language', 'in', '(', 'de', 'facto', 'official',  
 'spoken', 'form', 'of', 'Chinese', 'in', 'government', ')']  
 ['Cantonese', 'Regulated', 'by', 'Official', 'Language', 'Division']  
 ['Neo-Mandaic', 'Total', 'speakers', 'ca', '.', '500']  
 ['Mbula', 'Total', 'speakers', '2', ',', '500']  
 ['Mbula', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Teochew', 'Total', 'speakers', 'About', '10', 'million', 'in',  
 'Chaoshan', '.', 'Approximately', '2-5', 'million', 'speakers', 'overseas',  
 '.', '(', '49', 'million', 'for', 'Min', 'Nan', 'as', 'a', 'group', ')']  
 ['Teochew', 'Spoken', 'in', 'France', 'and', 'other', 'countries', 'where',  
 'Teochew', 'migrants', 'have', 'settled']  
 ['Teochew', 'Region', 'and', 'the', 'southmost', 'Fujian', 'county', 'of',  
 'Zhao'an']  
 ['Teochew', 'Language', 'codes', 'ISO', '639-3', 'nan-teo', '(', 'Teochew',  
 '\xe2', '\x80', '\x93', 'ambiguous', ')']  
 ['Teochew', 'Language', 'codes', 'ISO', '639-3']  
 ['Masalit', 'Total', 'speakers', '240', ',', '000']  
 ['Masalit', 'Region', 'Dar', 'Masalit', '(', 'Darfur', ')']  
 ['Suba', 'Total', 'speakers', '174', ',', '000', '(', '2007', ')']

['Malaysian', 'Total', 'speakers', '10', '.', '3', 'million', 'native']  
 ['Malaysian', 'Writing', 'system', 'Rumi', '(', 'Latin', 'alphabet', ')',  
 'and', 'Jawi', '(', 'Arabic', 'script', ')', ')']  
 ['Malaysian', 'Regulated', 'by', 'Dewan', 'Bahasa', 'dan', 'Pustaka', '(',  
 'Institute', 'of', 'Language', 'and', 'Literature', ')']  
 ['Old', 'Belarusian', 'Language', 'extinction', 'redeveloped', 'from',  
 'vernacular', 'spoken', 'remnants', 'into', 'Belarusian']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'orv-olr', '(', 'Old',  
 'Ruthenian', ')']  
 ['Ghomara', 'Total', 'speakers', '\xe2', '\x89', '\x88', '10', ',', '000']  
 ['Udmurt', 'Total', 'speakers', '550', ',', '000', '(', '1989', 'census',  
 ')']  
 ['Shimaore', 'Total', 'speakers', '80', ',', '140', '(', '2002', ')']  
 ['Old', 'Nubian', 'Total', 'speakers', '0']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'A', 'page', 'from', 'an',  
 'Old', 'Nubian', 'translation', 'of', 'the', 'Instructions', 'of', 'the',  
 'Archangel', 'Michael']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'from', 'the', '9th-10th',  
 'century']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'now', 'at', 'the', 'British',  
 'Museum', '.', "Michael's", 'name', 'appears', 'in', 'red', ':', 'Nubians',  
 'during', 'the', 'period', 'frequently', 'used', 'Greek', 'personal',  
 'names']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'often', 'with', 'a',  
 'terminal', '-', '\xce', '\xb9', 'added']  
 ['Harauti', 'Total', 'speakers', '8', '.', '8', 'million', '(', '1997',  
 ')']  
 ['Ludic', 'language', 'Total', 'speakers', '3', ',', '000', '(', '2007',  
 ')']  
 ['Ludic', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Karelian',  
 'variant', ')']  
 ['Tagoi', 'Total', 'speakers', '\xe2', '\x89', '\x88', '15', ',', '000']  
 ['Taiap', 'Total', 'speakers', '80', '(', '2000', ')']  
 ['Taiap', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Taiap', 'Region', 'Gapun', 'village', '(', 'East', 'Sepik', 'Province',  
 ')']  
 ['Delaware', 'Total', 'speakers', '1', 'or', '2', 'speakers', 'of',  
 'Munsee', 'as', 'of', '2008', ';', 'Unami', 'extinct']  
 ['Delaware', 'Region', 'Around', 'the', 'lower', 'Delaware', 'and',  
 'Hudson', 'rivers', 'in', 'the', 'United', 'States', ';', 'one', 'or',  
 'two', 'Munsee', 'speakers', 'in', 'Canada', ';', 'Unami', 'groups', 'in',  
 'Oklahoma']  
 ['Northern', 'Language', 'codes', 'ISO', '639-3', 'Northern', 'Sami', 'is',  
 '5', 'on', 'this', 'map']  
 ['Louisiana', 'Spoken', 'in', 'Illinois', 'and', 'a', 'small', 'community',  
 'in', 'East', 'Texas', '.', 'Significant', 'community', 'in', 'California',

';' , 'chiefly' , 'in' , 'Northern' , 'California']  
 ['Louisiana' , 'Total' , 'speakers' , '~' , '70' , ',' , '000' , '(' , 'as' , 'of' ,  
 '1985' , ')']  
 ['Tokelauan' , 'Total' , 'speakers' , '3' , ',' , '200' , 'total' , ',' , 'with' ,  
 '1400' , 'in' , 'Tokelau' , '(' , '2004' , ')']  
 ['Ivatan/' , 'Ibatan' , 'Total' , 'speakers' , '35' , ',' , '000']  
 ['Ivatan/' , 'Total' , 'speakers' , '(' , 'Population' , 'includes' , '3' , ',' ,  
 '448' , 'Itbayatan' , ')']  
 ['Matumbi' , 'Total' , 'speakers' , '70' , ',' , '000']  
 ['Yuchi' , 'Total' , 'speakers' , '~' , '5']  
 ['Totonac' , 'Total' , 'speakers' , 'over' , '250' , ',' , '000']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , ':' , 'toc' , '\xe2' ,  
 '\x80' , '\x93' , 'Totonac' , 'of' , 'Coyutla']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'tlp' , '\xe2' , '\x80' ,  
 '\x93' , 'Totonac' , 'of' , 'Filomeno' , 'Mata-Coahuítl' , '\xc3' , '\xa1' , 'n']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'tos' , '\xe2' , '\x80' ,  
 '\x93' , 'Highland' , 'Totonac']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'tqt' , '\xe2' , '\x80' ,  
 '\x93' , 'Totonac' , 'of' , 'Ozumatl' , '\xc3' , '\xa1' , 'n']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'top' , '\xe2' , '\x80' ,  
 '\x93' , 'Papantla' , 'Totonac']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'tot' , '\xe2' , '\x80' ,  
 '\x93' , 'Upper' , 'Necaxa' , 'Totonac']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'too' , '\xe2' , '\x80' ,  
 '\x93' , 'Totonac' , 'de' , 'Xicotepec']  
 ['Totonac' , 'Language' , 'codes' , 'ISO' , '639-3' , 'tlc' , '\xe2' , '\x80' ,  
 '\x93' , 'Misantla' , 'Totonac']  
 ["Weyt'o" , 'Language' , 'extinction' , '19th' , 'century']  
 ['Bung' , 'Total' , 'speakers' , '3' , '(' , '1995' , ')']  
 ['Bete' , 'Spoken' , 'in' , 'All' , 'speakers' , 'located' , 'in' , 'Bete' ,  
 'Town']  
 ['Bete' , 'Total' , 'speakers' , 'a' , 'small' , 'minority' , 'of' , '3000']  
 ['Tigre' , 'Total' , 'speakers' , '800' , ',' , '000' , 'in' , 'Eritrea' , '(' ,  
 '1997' , ')']  
 ['Adamorobe' , 'Region' , 'eastern' , 'Ghana' , ',' , 'Adamorobe' , 'village']  
 ['Adamorobe' , 'Total' , 'signers' , '1' , ',' , '200']  
 ['Adamorobe' , 'Language' , 'codes' , 'ISO' , '639-3' , 'ads' , 'sign' ,  
 'language' , '\xe2' , '\x80' , '\x94' , 'list' , 'of' , 'sign' , 'languages' ,  
 '\xe2' , '\x80' , '\x94' , 'legal' , 'recognition']  
 ['Sabaic' , 'Region' , 'Horn' , 'of' , 'Africa' , '&' , 'amp' , ';' , 'Arabian' ,  
 'Peninsula']  
 ['Tzeltal' , 'Total' , 'speakers' , '371' , ',' , '730']  
 ['Tzotzil' , 'Total' , 'speakers' , '329' , ',' , '937' , '(' , '2005' , 'census' ,  
 ')']  
 ['Irish' , 'Signed' , 'in' , 'Republic' , 'of' , 'Ireland' , ',' , 'Northern' ,  
 'Ireland']

['Irish', 'Language', 'codes', 'ISO', '639-3', 'isg', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Nubi', 'Total', 'speakers', '36', ',', '220']  
 ['Kwavi', 'Total', 'speakers', '30', ',', '000']  
 ['Lufu', 'Total', 'speakers', '3', ',', '200', '(', '2006', ')']  
 ['Lengue', 'Total', 'speakers', '40', ',', '000']  
 ['Old', 'Frankish', 'Language', 'extinction', 'Evolved', 'into', 'Old',  
 'Low', 'Franconian', 'by', 'the', '8th', 'century']  
 ['F\xc3\xa1', 'Total', 'speakers', '2', ',', '500']  
 ['Pothohari', 'Total', 'speakers', '4', ',', '680', ',', '000']  
 ['Southern', 'Spoken', 'in', 'Austria', '(', 'Carinthia', ',', 'Tyrol',  
 ',', 'Styria', ')', ',', 'Italy', ',', '(', 'South', 'Tyrol', ')']  
 ['Tsat', 'Total', 'speakers', '4', ',', '500']  
 ['Samogitian', 'Total', 'speakers', '~', '500', ',', '000']  
 ['Antillean', 'Spoken', 'in', 'Dominica', ',', 'French', 'Antilles', '(',  
 'Guadeloupe', ',', 'Martinique', ')', ',', 'Saint', 'Lucia']  
 ['Antillean', 'Total', 'speakers', '1', ',', '205', ',', '585']  
 ['Antillean', 'Language', 'codes', 'ISO', '639-3', 'acf', '\xe2', '\x80',  
 '\x93', 'Saint', 'Lucian', 'Creole', 'French']  
 ['Antillean', 'Language', 'codes', 'Linguasphere', '51-AAC-cc', '(',  
 'varieties']  
 ['Antillean', 'Language', 'codes', 'Linguasphere', ':', '51-AAC-cca', 'to',  
 '-cck', ')']  
 ['Mandingo', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']  
 ['Mandingo', 'Total', 'speakers', '5', '.', '3', 'million']  
 ['T\xc3\xbbbr\xc3\xb4yo', 'Total', 'speakers', '350', '.', '000']  
 ['T\xc3\xbbbr\xc3\xb4yo', 'Writing', 'system', 'Syriac', 'abjad', '(',  
 'Serto', 'variant', ')', ',', 'Latin', 'alphabet', 'has', 'been',  
 'modified', 'for', 'writing', 'Turoyo', 'in', 'Sweden', 'from', 'Yusuf',  
 'Ishaq', 'and', 'Germany', 'from', 'Silas', '\xc3', '\x9c', 'zel']  
 ['Southern', 'Kurdish', 'Total', 'speakers', 'About', '3', ',', '000', ',',  
 '000', 'in', 'Iran', 'and', 'elsewhere']  
 ['Southern', 'Spoken', 'in', 'Eastern', 'Iraq', ',', 'Western', 'Iran']  
 ['Southern', 'Region', 'Main', 'in', 'Western', 'Iran']  
 ['Southern', 'Region', 'Eastern', 'Iraq', 'border', 'with', 'those',  
 'provinces', 'including', 'Xanaqin']  
 ['Southern', 'Language', 'codes', 'Linguasphere', 'Southern', 'Kurdish',  
 'Area']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', ':', 'mww', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Daw', '(', 'Laos', ',', 'China', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmq', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Qiandong', 'Miao']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'muq', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Xiangxi', 'Miao']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hnj', '\xe2', '\x80',



'\x93', 'Hmong', 'Njua']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmz', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Shua', '(', 'Vietnam', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmc', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Central', 'Huishui', '(', 'China', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmm', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Central', 'Mashan', '(', 'China', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmj', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Chongjiang', '(', 'China', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hme', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Eastern', 'Huishui', '(', 'China', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'huj', '\xe2', '\x80',  
 '\x93', 'Northern', 'Guiyang', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmi', '\xe2', '\x80',  
 '\x93', 'Northern', 'Huishui', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmp', '\xe2', '\x80',  
 '\x93', 'Northern', 'Mashan', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hea', '\xe2', '\x80',  
 '\x93', 'Northern', 'Qiandong', 'Miao']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmy', '\xe2', '\x80',  
 '\x93', 'Southern', 'Guiyang', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hma', '\xe2', '\x80',  
 '\x93', 'Southern', 'Mashan', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hms', '\xe2', '\x80',  
 '\x93', 'Southern', 'Qiandong', 'Miao']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmg', '\xe2', '\x80',  
 '\x93', 'Southwestern', 'Guiyang', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmh', '\xe2', '\x80',  
 '\x93', 'Southwestern', 'Huishui', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmw', '\xe2', '\x80',  
 '\x93', 'Western', 'Mashan', 'Hmong']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'mmr', '\xe2', '\x80',  
 '\x93', 'Western', 'Xiangxi', 'Miao']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmv', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Do', '(', 'Vietnam', ')']  
 ['Hmong', 'Language', 'codes', 'ISO', '639-3', 'hmf', '\xe2', '\x80',  
 '\x93', 'Hmong', 'Don', '(', 'Vietnam', ')']  
 ['Enets', 'Language', 'codes', 'ISO', '639-3', ':', 'enf', '\xe2', '\x80',  
 '\x93', 'Forest', 'Enets']  
 ['Enets', 'Language', 'codes', 'ISO', '639-3', 'enh', '\xe2', '\x80',  
 '\x93', 'Tundra', 'Enets']  
 ['Mlahs\xc3\xb4', 'Language', 'extinction', 'with', 'the', 'death', 'of',  
 'Ibrahim', 'Hanna', 'in', '1998']  
 ['Yauma', 'Total', 'speakers', '22', ',', '000']  
 ['Creek', 'Total', 'speakers', '6', ',', '213']  
 ['Gayle', 'Spoken', 'in', 'South', 'Africa', ':', 'mainly', 'in',

'Johannesburg']  
 ['Lanc-Patu\xc3\xa1', 'Total', 'speakers', '40', ',', '000', '~', '50', ',', '000']  
 ['Yaqui', 'Total', 'speakers', '16', ',', '000']  
 ['Yaqui', 'Spoken', 'in', 'Mexico', ',', 'U', '.', 'S']  
 ['Badaga', 'Total', 'speakers', '400', ',', '000']  
 ['Badaga', 'Region', 'Tamil', 'Nadu', ',', 'The', 'Nilgiris']  
 ['Angami', 'Total', 'speakers', '109', ',', '000', '(', '1997', ')']  
 ['Aimaq', 'Total', 'speakers', '1', ',', '620', ',', '000']  
 ['Alyutor', 'Total', 'speakers', '200', '(', '2002', 'Census', ')']  
 ['Askunu', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Awadhi', 'Total', 'speakers', '38', ',', '261', ',', '000']  
 ['Awadhi', 'Region', 'India', ':', 'Awadh', 'and', 'Lower', 'Doab', 'regions', 'of', 'Uttar', 'Pradesh', ',', 'as', 'well', 'as', 'Madhya', 'Pradesh', 'and', 'Delhi']  
 ['Awadhi', 'Region', 'Nepal', ':', 'Lumbini', 'Zone']  
 ['Awadhi', 'Spoken', 'in', 'Fiji', '(', 'as', 'Fiji', 'Hindi', ')']  
 ['Awadhi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Asi', 'Total', 'speakers', '65', ',', '000']  
 ['Surigaanon', 'Total', 'speakers', '345', ',', '000', '(', '1990', ')']  
 ['Rombloanon', 'Total', 'speakers', '100', ',', '000']  
 ['Nafaanra', 'Total', 'speakers', '61', ',', '000']  
 ['Nafaanra', 'Spoken', 'in', 'Ghana', ',', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Nafaanra', 'Region', 'North-west', 'corner', 'of', 'the', 'Brong-Ahafo', 'region', 'in', 'Ghana', ',', 'east', 'of', 'Bondouko', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Masbate\xc3\xb1o', 'Total', 'speakers', '600', ',', '000']  
 ['Ratagnon', 'Total', 'speakers', '2', 'to', '3']  
 ['Ratagnon', 'Region', 'Southern', 'tip', 'of', 'western', 'Mindoro']  
 ['Pontic', 'Greek', 'Total', 'speakers', '324', ',', '535']  
 ['Cappadocian', 'Spoken', 'in', 'Greece', ',', 'originally', 'Cappadocia', '(', 'Central', 'Turkey', ')']  
 ['Warsaw', 'Language', 'extinction', 'in', 'decline', 'following', 'the', 'Warsaw', 'Uprising', 'of', '1944']  
 ['Warsaw', 'Pronunciation', 'Polish', ':', '\xcb', '\x88', 'j', '\xc9', '\x9b', 'nz', '\xc9', '\xa8', 'k', 'var', '\xcb', '\x88', '\xc9', '\x95', 'aski']  
 ['Zuni', 'Total', 'speakers', '9', ',', '651']  
 ['Zuni', 'Spoken', 'in', 'U', '.', 'S']

['Carib', 'Total', 'speakers', '10', ',', '226']  
 ['Ligbi', 'Total', 'speakers', '10', ',', '000']  
 ['Ligbi', 'Region', 'Brong-Ahafo', 'region', ',', 'adjacent', 'part', 'of', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]  
 ['Laki', 'Total', 'speakers', 'ca', '.', '1', ',', '500', ',', '000']  
 ['Cuyonon', 'Total', 'speakers', 'Official', 'population', 'are', 'unknown']  
 ['Caluyanon', 'Total', 'speakers', '30', ',', '000']  
 ['Yugh', 'Language', 'extinction', '20th', 'century']  
 ['Chickasaw', 'Total', 'speakers', 'About', '1', ',', '000']  
 ['Cypriot', 'Total', 'speakers', '1', ',', '300', '(', '1995', ')']  
 ['Karakalpak', 'Total', 'speakers', '412', ',', '000', '(', '1993', ')']  
 ['Karakalpak', 'Language', 'codes', 'ISO', '639-3', 'Map', 'showing', 'locations', 'of', 'Karakalpak', '(', 'red', ')', 'within', 'Uzbekistan']  
 ['Gan', 'ua', 'Total', 'speakers', '20', 'million', '~', '50', 'million']  
 ['Tunisian', 'Total', 'speakers', '11', ',', '000', ',', '000', '+']  
 ['Eteocretan', 'Language', 'extinction', 'beginning', 'of', '1st', 'millennium', 'BC', ',']  
 ['Kumyk', 'Total', 'speakers', '422', ',', '550']  
 ['Halkomelem', 'Total', 'speakers', '225', '(', '200', 'in', 'Canada', ';', '25', 'in', 'United', 'States', ')']  
 ['Halkomelem', 'Region', 'Southwestern', 'British', 'Columbia', 'into', 'Northern', 'Washington']  
 ['Tawang', '(', 'Monpa', ')', 'Total', 'speakers', '10', ',', '000']  
 ['Oneida', 'Total', 'speakers', '160', '-', '514']  
 ['Oneida', 'Region', 'Ontario', 'and', 'central', 'New', 'York', 'and', 'around', 'Green', 'Bay']  
 ['V\xc3\xb5ro', 'Total', 'speakers', '70', ',', '000']  
 ['V\xc3\xb5ro', 'Language', 'codes', 'ISO', '639-3', 'V', '\xc3', '\xb5', 'ro', 'language', 'area', '\xe2', '\x80', '\x94', 'V', '\xc3', '\xb5', 'romaa', '(', 'V', '\xc3', '\xb5', 'ro', 'county', ')', 'in', 'its', 'historical', 'boundaries', 'between', 'Tartu', 'and', 'Seto', 'areas', ',', 'Russia', '(', 'Vinnemaa', ')', 'and', 'Latvia', '(', 'L', '\xc3', '\xa4', 'timaa', ')']  
 ['Supyire', 'Total', 'speakers', '364', ',', '000', '(', '1991', ')']  
 ['Sucite', 'Total', 'speakers', '35', ',', '000', 'in', 'Burkina', 'Faso', '(', '1999', 'SIL', ')']  
 ['Newfoundland', 'French', 'Language', 'extinction', 'moribund']  
 ['Jutlandic', 'Region', 'Jutland', '(', 'Denmark', ')', 'and', 'in', 'the', 'northern', 'parts', 'of', 'Southern', 'Schleswig', '(', 'Germany', ')']  
 ['Chaldean', 'Neo-Aramaic', 'Total', 'speakers', '200', ',', '000', '(', '1990s', 'estimate', ')']  
 ['Chaldean', 'Region', 'now', 'also', 'Baghdad', 'and', 'Basra']  
 ['Chaldean', 'Writing', 'system', 'Syriac', 'abjad', '(', 'Madenhaya', 'variant', ')']  
 ['Assyrian', 'Neo-Aramaic', 'Total', 'speakers', '220', ',', '000']

['Senaya', 'Writing', 'system', 'Syriac', 'abjad', '(', 'M', '\xc4', '\x81', '\xe1', '\xb8', '\x8f', 'nh', '\xc4', '\x81', 'y', '\xc4', '\x81', 'variant', ')']

['Minyaka', 'Total', 'speakers', '740', ',', '000']

['Dagbani', 'Total', 'speakers', '~', '800', ',', '000']

['Homshetsi', 'Spoken', 'in', 'and', 'Central', 'Asia']

['Homshetsi', 'Writing', 'system', 'Armenian', 'alphabet', ',', 'modified', 'Turkish', 'alphabet']

['Lishana', 'Deni', 'Total', 'speakers', '8', ',', '000']

['Lishana', 'Region', 'Jerusalem', 'and', 'Maoz', 'Tsiyon', ',', 'originally', 'from', 'Iraqi', 'Kurdistan']

['Hiri', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Hiri', 'Total', 'speakers', '120', ',', '000', '(', 'use', 'declining', 'since', '1965', ')']

['Hiri', 'Official', 'language', 'in', 'Papua', 'New', 'Guinea']

['Lish\xc3\xa1n', 'Did', '\xc3', '\xa1', 'n', 'Total', 'speakers', '4', ',', '378']

['Lishanid', 'Noshan', 'Total', 'speakers', '2', ',', '000', 'to', '2', ',', '500']

['Karachay-Balkar', 'Total', 'speakers', '400', ',', '000']

['Karachay-Balkar', 'Writing', 'system', 'Latin', 'alphabet', '(', 'non', 'official', ')']

['Saramaccan', 'Total', 'speakers', '26', ',', '000']

['Patois', 'Spoken', 'in', 'Costa', 'Rica', '55', ',', '100']

['Patois', 'Spoken', 'in', 'Panama', '268', ',', '435']

['Patois', 'Total', 'speakers', '3', '.', '1', 'million']

['Mon\xc3\xa9gasque', 'Total', 'speakers', '35', ',', '657', '(', '2006', ')']

['Hulaul\xc3\xa1', 'Total', 'speakers', '10', ',', '000']

['Celtiberian', 'Language', 'extinction', '2nd', 'century', 'AD', '?']

['Dazaga', 'Total', 'speakers', '312', ',', '000']

['Khitan', 'Language', 'extinction', 'c', '.', '1243', '(', 'Yel', '\xc3', '\xbc', 'Chucai', ',', 'last', 'person', 'known', 'who', 'could', 'speak', 'and', 'write', 'Khitan', ')']

['Jurchen', 'Language', 'extinction', 'c', '.', '17th', 'century']

['Jurchen', 'Spoken', 'in', 'Southwest', 'Manchuria', '(', 'Northeastern', 'China', ')']

['Miao', 'Writing', 'system', 'Pahawh', 'Hmong', 'script', '(', 'Laos', ')']

['Miao', 'Language', 'codes', 'ISO', '639-3', ':', 'hmn', '\xe2', '\x80', '\x93', ']']

['Miao', 'Language', 'codes', 'ISO', '639-3', 'mmr', '\xe2', '\x80', '\x93', ']']

['Miao', 'Language', 'codes', 'ISO', '639-3', 'muq', '\xe2', '\x80', '\x93', ']']

['Miao', 'Language', 'codes', 'ISO', '639-3', 'hea', '\xe2', '\x80', '\x93', ']']

```

'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hms', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmq', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'cqđ', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'mww', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hrm', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'sfm', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmd', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hml', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmj', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'huj', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmg', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmy', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmi', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmh', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmc', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hme', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmm', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmp', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hmw', '\xe2', '\x80',
'\x93', '']
['Miao', 'Language', 'codes', 'ISO', '639-3', 'hma', '\xe2', '\x80',
'\x93', '']
['Bukusu', 'Total', 'speakers', '800', ',', '000']
['Nobiin,', 'Noban', 'tamen', 'Total', 'speakers', '495', ',', '000', '(',
'SIL', '2005', ')']
['Ibibio', 'Total', 'speakers', '3', '\xc2', '\xbd', 'million', '(',
'1990', '\xe2', '\x80', '\x93', '1998', ')']

```

['Ibibio', 'Language', 'codes', 'ISO', '639-3', 'ibb', '\xe2', '\x80',  
 '\x93', 'Ibibio', 'proper']  
 ['Iraqi', 'Total', 'speakers', '35', ',', '731', ',', '260', '(', '2011',  
 'est', '.', ')']  
 ['Iraqi', 'Regulated', 'by', 'Iraqi', 'Academy', 'of', 'Sciences']  
 ['Saek', 'Total', 'speakers', '25', ',', '000', '(', '1993', 'est', '.',  
 ')']  
 ['Modern', 'Regulated', 'by', 'modelled', 'after', 'the', 'Qur'an', ';',  
 'Academy', 'of', 'the', 'Arabic', 'Language']  
 ['Urum', 'Total', 'speakers', '192', ',', '729']  
 ['Old', 'Irish', 'Spoken', 'in', 'Ireland']  
 ['Old', 'Language', 'extinction', 'Evolved', 'into', 'Middle', 'Irish',  
 'about', 'the', '10th', 'century']  
 ['Primitive', 'Irish', 'Language', 'extinction', 'Evolved', 'into', 'Old',  
 'Irish', 'about', 'the', '6th', 'century', 'AD']  
 ['Niellim', 'Total', 'speakers', '5', ',', '000']  
 ['Nanerige', 'Total', 'speakers', '50', ',', '000', '(', '1985', ')']  
 ['Palaka', 'Total', 'speakers', '8', ',', '000', '(', '1995', ')']  
 ['Gula', 'Total', 'speakers', '3', ',', '500']  
 ['Buryat', 'Spoken', 'in', 'northwestern', 'the', 'People's', 'Republic',  
 'of', 'China']  
 ['Buryat', 'Total', 'speakers', '400', ',', '000']  
 ['Buryat', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-BAA-b']  
 ['Bua', 'Total', 'speakers', '7', ',', '708']  
 ['Hmar', 'Total', 'speakers', 'India', ':', '83', ',', '404', '(',  
 'Census', 'of', 'India', ',', '2001', ')']  
 ['Western', 'Neo-Aramaic', 'Total', 'speakers', '15', ',', '000']  
 ['Western', 'Region', 'Anti-Lebanon', 'mountains', ':', 'Ma'loula', ',',  
 'Bakh'a', 'and', 'Jubb'adin']  
 ['Marquesan', 'Total', 'speakers', '~', '11', ',', '000']  
 ['Marquesan', 'Language', 'codes', 'ISO', '639-3', ':', 'mrq', '\xe2',  
 '\x80', '\x93', 'North', 'Marquesan']  
 ['Marquesan', 'Language', 'codes', 'ISO', '639-3', 'mqm', '\xe2', '\x80',  
 '\x93', 'South', 'Marquesan']  
 ['Rapa', 'Nui', 'language', 'Total', 'speakers', 'up', 'to', '4', ',',  
 '650', '(', 'ethnic', 'Rapa', 'Nui', ',', '2002', ')', ',', 'or', 'as',  
 'low', 'as', '800']  
 ['Rapa', 'Writing', 'system', 'Latin', 'alphabet', ',', 'possibly',  
 'formerly', 'rongorongo']  
 ['Dakota', 'Total', 'speakers', '15', ',', '400']  
 ['Dakota', 'Region', 'Primarily', 'North', 'Dakota', 'and', 'South',  
 'Dakota']  
 ['Sioux', 'Total', 'speakers', '33', ',', '000', 'total', '(', 'including',  
 'Dakota', ',', 'Lakota', 'and', 'Assiniboine', ')', ';', '26', ',', '505',  
 'in', 'the', 'United', 'States', 'and', '6', ',', '495', 'in', 'Canada']  
 ['Pukapukan', 'Spoken', 'in', 'Pukapuka', 'and', 'Nassau', 'islands', ',',

'northern', 'Cook', 'Islands', ';', 'some', 'in', 'Rarotonga', ';', 'also', 'New', 'Zealand', 'and', 'Australia']

['Pukapukan', 'Total', 'speakers', '3000', '+']

['Barzani', 'Jewish', 'Neo-Aramaic', 'Total', 'speakers', '20', 'second-language', 'speakers', ',', 'effectively', 'extinct']

['Gusii', 'Region', 'Western', 'Kenya', ',', 'Gusii', 'district']

['H\u03a9tevin', 'Total', 'speakers', '1', ',', '000']

['H\u03a9tevin', 'Writing', 'system', 'Syriac', 'abjad', '(', 'Madnh', '\u00c4', '\u0081', 'y', '\u00c3', '\u00a2', 'variant', ')']

['Koy', 'Sanjaq', 'Surat', 'Total', 'speakers', '800', 'to', '1', ',', '000']

['Koy', 'Region', 'Koy', 'Sanjaq', 'and', 'Armota', 'in', 'Arbil']

['Luo', 'Total', 'speakers', '4', ',', '410', ',', '000']

['Luo', 'Region', 'East', 'of', 'Lake', 'Victoria', 'in', 'Western', 'Kenya', 'and', 'Northern', 'Tanzania']

['Adangme', 'Total', 'speakers', '800', ',', '000', '(', 'SIL', '2004', ')']

['Adangme', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Adangme', 'variant', ')']

['Ho', 'Total', 'speakers', '3', ',', '803', ',', '126']

['Timbisha', 'Total', 'speakers', '&', 'lt', ';', '20']

['Gronings', 'Spoken', 'in', 'the', 'east', 'of', 'the', 'Frisian', 'municipality', 'Kollumerland']

['Gronings', 'Official', 'language', 'in', 'the', 'Netherlands', '(', 'as', 'part', 'of', 'Low', 'Saxon', ')']

['Koryak', 'Total', 'speakers', '3019', '(', '2002', 'Census']

['Mbum', 'Spoken', 'in', 'Cameroon', ',', 'Central', 'African', 'Republic']

['Mbum', 'Total', 'speakers', '51', ',', '000', '(', '1982', ')']

['Hawaii', 'Language', 'codes', 'ISO', '639-3', 'hps', 'sign', 'language', '\u00e2', '\u0080', '\u0094', 'list', 'of', 'sign', 'languages', '\u00e2', '\u0080', '\u0094', 'legal', 'recognition']

['Mbabaram', 'Language', 'extinction', '1972', 'with', 'the', 'death', 'of', 'Albert', 'Bennett']

['Yanyuwa', 'Total', 'speakers', '5', '\u00e2', '\u0080', '\u0093', '15']

['Miami-Illinois', 'Language', 'extinction', 'mid-20th', 'century', '(', 'currently', 'undergoing', 'revival', ')']

['Tugen', 'Total', 'speakers', 'around', '200', '000']

['Kobon', 'Total', 'speakers', '10', ',', '000']

['Kobon', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Kobon', 'Region', 'and', 'Western', 'Highlands', 'Province', 'on', 'Kaironk', 'River', 'in', 'lower', 'Jimi', 'River', 'area', 'north', 'of', 'Mt', '.', 'Hagen']

['P\u03b6koot', 'Total', 'speakers', '264', ',', '000']

['P\u03b6koot', 'Region', 'West', 'Pokot', 'and', 'Baringo', 'districts', 'in', 'Kenya', '.', 'Karimojong', 'borderland', 'in', 'Uganda']

['Ogiek', 'Total', 'speakers', '~', '500']

['Ogiek', 'Region', 'Sogoo', ':', 'Kenya', ',', 'southern', 'Mau',  
 'forest', 'between', 'the', 'Amala', 'and', 'Ewas', 'Ng'iro', 'rivers']  
 ['Classical', 'Armenian', 'Language', 'extinction', 'developed', 'into',  
 'Middle', 'Armenian']  
 ['Lendu', 'Total', 'speakers', '760', ',', '000', '(', '1996', ')']  
 ['Koyra', 'Chiini', 'Total', 'speakers', '200', ',', '000', '(', 'as',  
 'of', '1999', ')']  
 ['Bactrian', 'Language', 'extinction', 'ca', '.', '9th', 'c', '.', 'AD']  
 ['Walliser', 'Spoken', 'in', 'upper', 'Valais', '&', 'amp', ';', 'Walser',  
 'colonies']  
 ['Walliser', 'Total', 'speakers', 'about', '80', ',', '000']  
 ['Vaeakau-Taumako', 'Total', 'speakers', '1', ',', '142', '(', '1999', ')']  
 ['Wiyot', 'Spoken', 'in', 'California', ',', 'U', '.', 'S']  
 ['Yurok', 'Spoken', 'in', 'Northwestern', 'California', ',', 'U', '.', 'S']  
 ['Yurok', 'Total', 'speakers', '~', '12']  
 ['Xiang', 'Region', 'Central', 'and', 'southwestern', 'Hunan']  
 ['Tunica', 'Language', 'extinction', 'since', 'the', 'death', 'of',  
 'Sesostrie', 'Youchigant']  
 ['Tunica', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact',  
 'distribution', 'of', 'the', 'Tunica', 'language']  
 ['Pinghua', 'Total', 'speakers', '2', ',', '300', ',', '000']  
 ['Palenquero', 'Total', 'speakers', '2', ',', '500']  
 ['Atikamekw', 'Total', 'speakers', '5', ',', '545']  
 ['Tolomako', 'Total', 'speakers', 'fewer', 'than', '500']  
 ['Tolomako', 'Region', 'Big', 'Bay', ',', 'Espiritu', 'Santo', 'Island']  
 ['Kurmanji', 'Spoken', 'in', '&', 'amp', ';', 'neighboring', 'countries']  
 ['Kurmanji', 'Total', 'speakers', 'about', '9', 'million', '(', 'data',  
 '1980', '\xe2', '\x80', '\x93', '2004', ')']  
 ['Macedonian', 'Language', 'extinction', '4th', 'century', 'BCE']  
 ['Ch'orti', 'Total', 'speakers', 'approx', '.', '20', ',', '000']  
 ['Adai', 'Language', 'extinction', '19th', 'century']  
 ['Lusitanian', 'Language', 'extinction', '2nd', 'century', 'AD']  
 ['Lusitanian', 'Region', 'Inland', 'central-west', 'Iberian', 'Peninsula']  
 ['Lisu', 'Total', 'speakers', '723', ',', '000']  
 ['Lisu', 'Writing', 'system', 'Fraser', 'Alphabet', '(', 'Latin', 'Based',  
 ')']  
 ['Lisu', 'Official', 'language', 'in', 'Weixi', 'Lisu', 'Autonomous',  
 'County', ',', 'Nujiang', 'Lisu', 'Autonomous', 'Prefecture', '(', 'PRC',  
 ')']  
 ['Krio', 'Total', 'speakers', '340', ',', '000', 'native', 'speakers', '(',  
 'the', 'Sierra', 'Leone', 'Creole', 'people', ')']  
 ['Krio', 'Total', 'speakers', '6', ',', '250', ',', '000', 'total',  
 'speakers', '-', '(', '97', '%', 'of', 'Sierra', 'Leone's', 'population',  
 ')']  
 ['Krio', 'Language', 'codes', 'Linguasphere', '52-ABB-bb', 'Where', 'it',  
 'is', 'mainly', 'spoken']



['Jud\xc3\xa6o-Tat', 'Spoken', 'in', 'United', 'States', '(', 'New',  
 'York', ')']  
 ['Zarphatic', 'Language', 'extinction', 'fourteenth', 'century']  
 ['Nitinaht', 'Region', 'Southern', 'part', 'of', 'Vancouver', 'Island',  
 ', ', 'British', 'Columbia']  
 ['Goemai', 'Total', 'speakers', '200', ', ', '000']  
 ['Haisla', 'Spoken', 'in', 'Central', 'British', 'Columbia', 'coast',  
 'inlet']  
 ['Heiltsuk', 'Spoken', 'in', 'Northern', 'Central', 'Coast']  
 ['She', 'language', 'Total', 'speakers', '1', ', ', '200', '(', '2002', ')']  
 ['She', 'Region', 'Huidong', 'County', 'and', 'Haifeng', 'County', 'in',  
 'Guangdong']  
 ['Bukhori', 'Total', 'speakers', '~', '110', ', ', '000']  
 ['Krymchak', 'Total', 'speakers', '~', '100']  
 ['Krymchak', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Bhili', 'Total', 'speakers', '5', '.', '6', 'million', '(', '1991',  
 'census', ')']  
 ['Mochica', 'Language', 'extinction', 'ca', '.', '1920']  
 ['Ithkuil', 'Date', 'founded', '1978', '\xe2', '\x80', '\x94', '2004']  
 ['Ithkuil', 'Pronunciation', '\xc9', '\xaa', '\xce', '\xb8', '\xcb',  
 '\x88', 'k', '\xca', '\x8a', '.', 'il', '1']  
 ['Temne', 'Spoken', 'in', 'Central', 'Sierra', 'Leone']  
 ['Temne', 'Total', 'speakers', 'about', '1', ', ', '000', ', ', '000']  
 ['Fula', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Fula', 'Language', 'codes', 'ISO', '639-3', 'fui', '\xe2', '\x80',  
 '\x93', 'Bagirmi', 'Fulfulde']  
 ['Fula', 'Language', 'codes', 'ISO', '639-3', 'ffm', '\xe2', '\x80',  
 '\x93', 'Maasina', 'Fulfulde']  
 ['Fula', 'Language', 'codes', 'ISO', '639-3', 'fuv', '\xe2', '\x80',  
 '\x93', 'Nigerian', 'Fulfulde']  
 ['Fula', 'Language', 'codes', 'ISO', '639-3', 'fuf', '\xe2', '\x80',  
 '\x93', 'Pular', 'language']  
 ['Fula', 'Language', 'codes', 'ISO', '639-3', 'fuh', '\xe2', '\x80',  
 '\x93', 'Western', 'Niger', 'Fulfulde']  
 ['Bullom', 'So', 'Total', 'speakers', '500', '(', '1998', ')', ';', 'near',  
 'extinction']  
 ['Bullom', 'Spoken', 'in', 'Guinea', ', ', 'Sierra', 'Leone']  
 ['Bullom', 'Region', 'coast', 'of', 'Guinea', ', ', 'near', 'the', 'Sierra',  
 'Leone', 'border']  
 ['Krim', 'Total', 'speakers', 'fewer', 'than', '500']  
 ['Krim', 'Region', 'coast', 'of', 'Sierra', 'Leone']  
 ['Sherbro', 'Total', 'speakers', '135', ', ', '000']  
 ['Sherbro', 'Region', 'coast', 'of', 'Sierra', 'Leone', 'and', 'Sherbro',  
 'Island']  
 ['Karon', 'Total', 'speakers', '10', ', ', '420']  
 ['Mlomp', 'Total', 'speakers', '4', ', ', '895']

['Mende', 'Total', 'speakers', '1', ',', '480', ',', '000']  
 ['Mende', 'Region', 'South', 'central', 'Sierra', 'Leone']  
 ['Mende', 'Spoken', 'in', 'Sierra', 'Leone', ',', 'Liberia']  
 ['Mende', 'Writing', 'system', 'Latin', ';;', 'Kisimi', "Kamara's", 'Mende', 'syllabary']  
 ['Tedaga', 'Total', 'speakers', '42', ',', '000']  
 ['Fala', 'Total', 'speakers', '6', ',', '500']  
 ['Kiowa', 'Total', 'speakers', '1', ',', '000', '+', '(', '2000', 'census', ')', ',', '300', '(', 'Mithun', ')']  
 ['Kiowa', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Kiowa', 'language']  
 ['Gagauz', 'Total', 'speakers', '162', ',', '200']  
 ['Gagauz', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Sicani', 'Language', 'codes', 'ISO', '639-3', 'A', 'rough', 'map', 'of', 'Sicilian', 'ethnic', 'locations']  
 ['Shuadit', 'Language', 'extinction', '1977', ',', 'on', 'the', 'death', 'of', 'Armand', 'Lunel']  
 ['Judaéo-Portuguese', 'Spoken', 'in', 'Northern', 'Germany', '(', 'Hamburg', ')']  
 ['Judaéo-Portuguese', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Portuguese', 'variant', ')', 'Hebrew', 'alphabet']  
 ['Kankana-ey', 'Total', 'speakers', '2', '.', '7', 'million', ',', 'est', '.', '1', '.', '3', 'million', 'second', 'language', '=', '4', 'million', 'total']  
 ['Proto-Norse', 'Language', 'extinction', 'evolved', 'into', 'Old', 'Norse', 'from', 'the', '8th', 'century']  
 ['Kumaoni', 'Total', 'speakers', '2', ',', '360', ',', '000', '(', '1998', ')']  
 ['Kumaoni', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Kalenjin', 'Total', 'speakers', '1', '.', '6', 'million']  
 ['Yokutsan', 'Spoken', 'in', 'San', 'Joaquin', 'Valley', ',', 'California', 'Ethnicity']  
 ['Yokutsan', 'Total', 'speakers', '25', '(', 'Northern', 'Foothill', 'Yokuts', ')']  
 ['Logba', 'Total', 'speakers', '7', '500']  
 ['Coahuilteco', 'Language', 'extinction', 'mid-20th', 'century']  
 ['Northern', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Maya', 'Dialects', 'Highland', 'Maya', 'Sign']  
 ['Maya', 'Language', 'codes', 'ISO', '639-3', 'msd', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Dogri', 'Total', 'speakers', '4', '.', '7', 'million']  
 ['Dogri', 'Language', 'codes', 'ISO', '639-3', ':', 'dgo', '\xe2', '\x80',  
 '\x93', 'Dogri', '(', 'individual', ')']  
 ['Dogri', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Anfillo', 'Total', 'speakers', '500', '(', '1990', 'SIL', ')']  
 ['Konso', 'Total', 'speakers', '242', ',', '059', '(', '2007', 'census',  
 ')']  
 ['Konso', 'Region', 'South', 'of', 'Lake', 'Chamo', 'in', 'the', 'bend',  
 'of', 'the', 'Sagan', 'River']  
 ['Bussa', 'Region', 'Dirashe', 'special', 'woreda', ',', 'Southern',  
 'Ethiopia']  
 ['Eblaité', 'Language', 'extinction', 'before', 'the', '2nd', 'millennium',  
 'BCE']  
 ['Quebec', 'Region', 'Concentrated', 'in', 'Quebec', ',', 'but', 'with',  
 'signers', 'in', 'other', 'Canadian', 'urban', 'centres']  
 ['Quebec', 'Language', 'codes', 'ISO', '639-3', 'fcs', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Modern', 'Swedish', 'Language', 'extinction', 'developed', 'into',  
 'contemporary', 'Swedish', 'by', 'the', 'late', '19th', 'century']  
 ['Modern', 'Spoken', 'in', 'Swedish', 'Empire', ',', '(', 'contemporary',  
 'Sweden', ',', 'Finland', 'and', 'Estonia', ')']  
 ['Old', 'Swedish', 'Language', 'extinction', 'Evolved', 'into', 'Modern',  
 'Swedish', 'by', 'the', '16th', 'century']  
 ['Khazar', 'Language', 'extinction', 'sometime', 'between', '1000', 'and',  
 '1300', 'CE']  
 ['Khazar', 'Spoken', 'in', 'Pontic', 'steppes', 'and', 'parts', 'of',  
 'Central', 'Asia']  
 ['Khazar', 'Writing', 'system', 'variant', 'of', 'Old', 'Turkic', 'script',  
 ';', 'Hebrew', 'alphabet', ';', 'possibly', 'others']  
 ['Sawai', 'Total', 'speakers', '12', ',', '000']  
 ['Plains', 'Apache', 'Total', 'speakers', '18']  
 ['Kipsigis', 'Total', 'speakers', '471', ',', '000']  
 ['Bolgo', 'Total', 'speakers', '1', ',', '800']  
 ['Mescalero-Chiricahua', 'Total', 'speakers', '279', '(', '1990', ')']  
 ['Mescalero-Chiricahua', 'Region', 'Oklahoma', ',', 'New', 'Mexico']  
 ['South', 'Jutlandic', 'Total', 'speakers', 'unknown']  
 ['South', 'Region', 'Northern', 'Schleswig', '/', 'Southern', 'Jutland',  
 '(', 'Denmark', ')', 'and', 'in', 'Southern', 'Schleswig', '(', 'Germany',  
 ')']  
 ['Judaéo-Georgian', 'Total', 'speakers', '85', ',', '000']  
 ['Old', 'Turkic', '/', 'Old', 'Uyghur', 'Language', 'extinction',  
 'evolved', 'into', 'Uyghur', 'by', 'the', '13th', 'century']

['Brabantian', 'Language', 'codes', 'Linguasphere', '52-ACB-ak', '(', 'varieties']

['Brabantian', 'Language', 'codes', 'Linguasphere', ':', '52-ACB-aka', 'to-akk', ')']

['Newfoundland', 'Irish', 'Total', 'speakers', 'possibly', 'none', ';', 'moribund', 'or', 'extinct']

['Newfoundland', 'Region', 'Newfoundland', ',', 'mostly', 'on', 'Avalon', 'Peninsula']

['Newfoundland', 'Language', 'codes', 'ISO', '639-3', 'gle', 'The', 'Irish', 'Shore', 'of', 'Newfoundland's', 'Avalon', 'Peninsula']

['Old', 'Total', 'signers', 'Estimated', '200', 'in', 'Paris', 'in', 'the', '1750s']

['Old', 'Dialects', 'French', 'Sign', 'Language']

['Old', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Thompson', 'Spoken', 'in', 'British', 'Columbia', ',', 'Washington']

['Tonkawa', 'Language', 'extinction', 'ca', '.', '1940']

['Tonkawa', 'Region', 'Western', 'Oklahoma', ',', 'South-central', 'Texas', 'and', 'into', 'New', 'Mexico']

['Tonkawa', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Tonkawa', 'language']

['Brigasc', 'Total', 'speakers', 'approximately', '1', ',', '000']

['Gheg', 'Total', 'speakers', '4', ',', '200', ',', '000']

['Gheg', 'Language', 'codes', 'Linguasphere', '55-AAA-aaa', 'to', '55-AAA-aag']

['Tuscan', 'Total', 'speakers', '3', ',', '500', ',', '000']

['Luise\xc3\xb1o', 'Total', 'speakers', '30', '\xe2', '\x80', '\x93', '40']

['Makhuwa', 'Total', 'speakers', '3', '.', '1', 'million', '(', '2006', ')']

['Latgalian', 'Total', 'speakers', '150', ',', '000']

['Chibcha', 'Language', 'extinction', '18th', 'century']

['Chibcha', 'Spoken', 'in', 'Central', 'highlands', 'of', 'Colombia']

['Caddo', 'Region', 'Caddo', 'County', 'in', 'western', 'Oklahoma']

['Classical', 'Nepal', 'Bhasa', 'Spoken', 'in', 'Nepal']

['Classical', 'Language', 'extinction', 'No', 'native', 'speakers', ',', 'used', 'as', 'a', 'literary', 'and', 'liturgical', 'language', 'only']

['Classical', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Yapese', 'Total', 'speakers', '6', ',', '600']

['Argobba', 'Total', 'speakers', '43', ',', '682', '(', '2007', ')']

['Gafat', 'Ethiopic', 'Language', 'extinction', 'Extinct']

['Vunjo', 'Total', 'speakers', '1', '.', '5', 'million']

['Cape', 'Spoken', 'in', 'Cape', 'Verde', ',', 'Cape', 'Verdean',

'diaspora']  
 ['Cape', 'Total', 'speakers', '926', ',', ',', '078']  
 ['Goaria', 'Total', 'speakers', '25', ',', ',', '000']  
 ['German', 'Total', 'signers', '200', ',', ',', '000']  
 ['German', 'Language', 'codes', 'ISO', '639-3', 'gsg', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Northern', 'Ndebele', 'Total', 'speakers', '1', ',', ',', '558', ',', ',', '000']  
 ['Northern', 'Region', 'Matabeleland', 'North', ',', ',', 'Matabeleland',  
 'South', 'in', 'Zimbabwe', ';', 'North-East', 'District', 'in', 'Botswana']  
 ['Akkala', 'Sami', 'Language', 'extinction', 'December', '29', ',', ',', '2003',  
 'with', 'the', 'death', 'of', 'Marja', 'Sergina']  
 ['Oorlans', 'Total', 'speakers', '200', ',', ',', '000']  
 ['Rohingya', '(', 'Total', 'speakers', '~', '1', '.', '5', 'million']  
 ['Fataluku', 'Spoken', 'in', 'eastern', 'East', 'Timor']  
 ['Fataluku', 'Total', 'speakers', '30', ',', ',', '000', '+', 'regional',  
 'usage']  
 ['Garhwali', 'Total', 'speakers', '2', ',', ',', '267', ',', ',', '314', '(',  
 'India', 'Census', '2001', ')']  
 ['Garhwali', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Baga', 'Binari', 'Total', 'speakers', 'About', '3', ',', ',', '000']  
 ['Maasai', 'Total', 'speakers', '900', ',', ',', '000']  
 ['Maasai', 'Region', 'Southern', 'Kenya', 'and', 'Northern', 'Tanzania']  
 ['Rotuman', 'Total', 'speakers', '9', ',', ',', '000']  
 ['Tiv', 'Total', 'speakers', '2', ',', ',', '210', ',', ',', '000']  
 ['Warlpiri', 'Total', 'speakers', '2', ',', ',', '670']  
 ['Cook', 'Islands', 'M', '\xc4', '\x81', 'ori', 'Total', 'speakers', '42',  
 ',', ',', '669', '(', '16', ',', ',', '800', 'in', 'Cook', 'Islands', '(', '1979',  
 'government', 'report', ')', ')']  
 ['Cook', 'Region', 'throughout', 'Cook', 'Islands', '.', '.', '.', ',',  
 'New', 'Zealand']  
 ['Cook', 'Language', 'codes', 'ISO', '639-3', 'pnh', '\xe2', '\x80',  
 '\x93', 'Tongareva', 'dialect', '(', 'Penrhyn', ')']  
 ['Cook', 'Language', 'codes', 'ISO', '639-3', 'rkh', '\xe2', '\x80',  
 '\x93', 'Rakahanga-Manihiki', 'dialect']  
 ['Kamayo', 'Total', 'speakers', '2', '.', '5', 'Million']  
 ['Gela', 'Total', 'speakers', '5', ',', ',', '000', '\xe2', '\x80', '\x93',  
 '10', ',', ',', '000']  
 ['Gela', 'Spoken', 'in', 'Central', 'Solomon', 'Islands']  
 ['Damin', 'Setting', 'and', 'usage', 'Initiation', 'language', 'for',  
 'men', ',', 'used', 'by', 'the', 'Lardil', 'tribe', 'of', 'Mornington',  
 'Island']  
 ['Damin', 'Created', 'by', 'the', 'Lardil', 'tribe']

['Afrikaans', '(', 'Western', 'Cape', 'dialect', ')', 'Total', 'speakers']  
 ['Ndaou', 'Total', 'speakers', '1', ',', '580', ',', '000', 'in',  
 'Mozambique', '(', '2006', ')', ',', 'Population', 'total', 'all',  
 'countries', ':', '2', ',', '380', ',', '000']  
 ['Haya', 'Total', 'speakers', 'about', '1', ',', '200', ',', '000']  
 ['Samburu', 'Total', 'speakers', '128', ',', '000', '(', '147', ',', '000',  
 'including', 'the', 'Camus', ')']  
 ['Samburu', 'Region', 'Samburu', 'district', 'of', 'Rift', 'Valley',  
 'Province']  
 ['Yaaku', 'Total', 'speakers', '50', '(', '1983', ')']  
 ['Khmu', 'Total', 'speakers', '480', ',', '000']  
 ['Chumash', 'Language', 'extinction', 'since', 'the', '1960s']  
 ['Qiong', 'wen', 'Total', 'speakers', '1', '.', '1', 'million']  
 ['Qiong', 'Spoken', 'in', 'China', ',', 'United', 'States', '(', 'New',  
 'York', 'City', ',', 'California', ')', ',', 'Taiwan', ',', 'Malaysia',  
 ',', 'Singapore', ',', 'Hong', 'Kong']  
 ['Arcaicam', 'Esperantom', 'Total', 'speakers', '0']  
 ['Arcaicam', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Arcaicam', 'Category', '(', 'purpose', ')', 'Arcaicam', 'Esperantom']  
 ['Zeelandic', 'Total', 'speakers', '220', ',', '000']  
 ['Zeelandic', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Zeelandic',  
 'variant', ')']  
 ['Zeelandic', 'Language', 'codes', 'Linguasphere', 'Position', 'of',  
 'Zeelandic', '(', 'colour', ':', 'sandy', ')', 'among', 'the', 'other',  
 'minority', 'languages', ',', 'regional', 'languages', 'and', 'dialects',  
 'in', 'the', 'Benelux']  
 ['Ticinese', 'Total', 'speakers', '150', ',', '000']  
 ['oluLuhya', 'Total', 'speakers', '6', ',', '000', ',', '000', 'to', '10',  
 ',', '000', ',', '000']  
 ['Defaka', 'Total', 'speakers', '200', '(', 'Blench', '2003', ')']  
 ['Malvi', 'Total', 'speakers', '20', '.', '0', 'million', '(', '1997', ')']  
 ['Aquitanian', 'Language', 'extinction', 'by', 'the', 'Early', 'Middle',  
 'Ages']  
 ['Aquitanian', 'Language', 'extinction', '(', 'except', 'in', 'the',  
 'Northern', 'Basque', 'Country', ')']  
 ['Aquitanian', 'Region', 'West', 'of', 'the', 'Pyrenees']  
 ['Mecklenburgisch-Vorpommersch', 'Language', 'codes', 'ISO', '639-3',  
 'Low', 'German', 'and', 'Low', 'Franconian', 'dialects']  
 ['Tangut', 'Official', 'language', 'in', 'Western', 'Xia', 'Dynasty', 'of',  
 'China']  
 ['Western', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80',  
 '\x93', 'Lombard', 'group']  
 ['Western', 'Language', 'codes', 'Linguasphere', '51-AAA-odd', '.', '.',  
 '.', '51-AAA-odj']  
 ['Tabasaran', 'Total', 'speakers', '128', ',', '900', '(', '2002', ')']

['Tosk', 'Language', 'codes', 'Linguasphere', '55-AAA-aca', 'to', '55-AAA-ace']  
 ['Highest', 'Total', 'speakers', 'about', '500', ',', '000']  
 ['Highest', 'Language', 'codes', 'ISO', '639-3', ':', 'gsw', '\xe2', '\x80', '\x93', 'Swiss', 'German']  
 ['Highest', 'Language', 'codes', 'ISO', '639-3', 'wae', '\xe2', '\x80', '\x93', 'Walser', 'German']  
 ['Moselle', 'Language', 'codes', 'ISO', '639-3', 'Area', 'where', 'Moselle', 'Franconian', 'and', 'Luxembourgish', 'is', 'spoken']  
 ['Tukang', 'Besi', 'Total', 'speakers', '200', ',', '000']  
 ['Tukang', 'Region', 'Sulawesi', ',', 'Tukang', 'Besi', 'Archipelago']  
 ['Tukang', 'Language', 'codes', 'ISO', '639-3', ':', 'khc', '\xe2', '\x80', '\x93', 'Tukang', 'Besi', 'North']  
 ['Tukang', 'Language', 'codes', 'ISO', '639-3', 'bhq', '\xe2', '\x80', '\x93', 'Tukang', 'Besi', 'South']  
 ['Taba', 'Total', 'speakers', '20', ',', '000']  
 ['Kurukh', 'Total', 'speakers', '2', ',', '053', ',', '000', '(', 'SIL', '1997', ')']  
 ['Tai', 'L', '\xc3', '\xbc', 'Language', 'Total', 'speakers']  
 ['Tai', 'Writing', 'system', 'New', 'Tai', 'L', '\xc3', '\xbc', 'script']  
 ['Tai', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Tai', 'N', '\xc3', '\xbc', 'a', 'Total', 'speakers', '357', ',', '400']  
 ['Lemnian', 'Language', 'extinction', '400', 'BC']  
 ['Lule', 'Total', 'speakers', '~', '2', ',', '000']  
 ['Lule', 'Language', 'codes', 'ISO', '639-3', 'smj', 'Lule', 'Sami', 'is', '4', 'on', 'this', 'map']  
 ['Turkana', 'Total', 'speakers', '340', ',', '000']  
 ['Albay', 'Bikolano', 'Total', 'speakers', '1', ',', '900', ',', '000']  
 ['Albay', 'Language', 'codes', 'ISO', '639-3', ':', 'ubl', '\xe2', '\x80', '\x93', 'Buhi'non', 'Bikol']  
 ['Albay', 'Language', 'codes', 'ISO', '639-3', 'fbl', '\xe2', '\x80', '\x93', 'West', 'Albay', 'Bikol']  
 ['Mussau-Emira', 'Total', 'speakers', '4', ',', '200', 'to', '5', ',', '000']  
 ['Mussau-Emira', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Mussau-Emira', 'Region', 'Islands', 'of', 'Mussau', 'and', 'Emira', '(', 'New', 'Ireland', 'Province', ')']  
 ['Nheengatu', 'Total', 'speakers', '3', ',', '000', '(', 'Brazil', ',', '1998', ')', ',', '3', ',', '000', '(', 'Colombia', ')', ',', '2', ',', '000', '(', 'Venezuela', ')']  
 ['Shasta', 'Language', 'extinction', 'by', 'end', 'of', '20th', 'century']  
 ['Quechan', '/', 'Yuma', 'Total', 'speakers', '150', '(', '1994', ')']  
 ['Utopian', 'Writing', 'system', 'Utopian', 'alphabet', 'Persian', ',',

'with', 'influence', 'of', 'Greek', 'and', 'Latin']  
 ['Savoyard', 'Total', 'speakers', '\xc2', '\xb1', '35', '.', '000',  
 'speakers']  
 ['Savoyard', 'Regulated', 'by', 'Institute', 'Savoyard', 'language']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'okl', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Aluku', 'Total', 'speakers', 'more', 'than', '11', ',', '000']  
 ['Ndyuka', 'Total', 'speakers', '30', ',', '000']  
 ['Ndyuka', 'Language', 'codes', 'Linguasphere', '52-ABB-ay', '(',  
 'varieties']  
 ['Ndyuka', 'Language', 'codes', 'Linguasphere', ':', '52-ABB-aya', 'to', '-  
 aye', ')']  
 ['Chinese', 'Total', 'speakers', 'Possibly', 'evolved', 'into',  
 'Chinglish', ',', 'and', 'into', 'English', 'loanwords', 'in', 'Chinese',  
 'dialects']  
 ['Grenadian', 'Total', 'speakers', '89', ',', '227']  
 ['Antiguan', 'Total', 'speakers', '125', '244']  
 ['Antiguan', 'Language', 'codes', 'Linguasphere', '52-ABB-apl', 'to', '-  
 apm']  
 ['Saint', 'Spoken', 'in', 'Saint', 'Kitts', 'and', 'Nevis']  
 ['Saint', 'Total', 'speakers', '&', 'lt', ';;', '100', ',', '000']  
 ['Turks', 'Spoken', 'in', 'Turks', 'and', 'Caicos', 'Islands']  
 ['Virgin', 'Spoken', 'in', 'British', 'Virgin', 'Islands']  
 ['Virgin', 'Spoken', 'in', 'U', '.', 'S', '.', 'Virgin', 'Islands']  
 ['Virgin', 'Spoken', 'in', 'Virgin', 'Islands', 'diaspora']  
 ['Virgin', 'Total', 'speakers', '&', 'gt', ';;', '100', ',', '000']  
 ['Virgin', 'Language', 'codes', 'Linguasphere', '52-AAB-apa', 'to', '-ape',  
 'and', '52-AAB-apg', 'to', 'api', '(', 'SSS', 'varieties', ')']  
 ['Miskito', 'Total', 'speakers', '&', 'lt', ';;', '30', ',', '000']  
 ['San', 'Spoken', 'in', 'Colombia', '(', 'San', 'Andr', '\xc3', '\xa9',  
 's', 'and', 'Providencia', 'islands', ')']  
 ['Guyanese', 'Total', 'speakers', 'Over', '1', 'million']  
 ['Waima', 'Spoken', 'in', 'Eastern', 'New', 'Guinea']  
 ['Waima', 'Total', 'speakers', '15', ',', '000']  
 ['Tsakonian', 'Region', 'Eastern', 'Peloponnese', 'around', 'Mount',  
 'Parnon']  
 ['Gelao', 'Total', 'speakers', '7', ',', '900', '(', 'Edmondson', '2008',  
 ')']  
 ['Sri', 'Total', 'speakers', '3', ',', '406']  
 ['Icelandic', 'Language', 'codes', 'ISO', '639-3', 'icl', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Fanagalo', 'Total', 'speakers', 'Several', 'hundred', 'thousand', 'in',  
 'South', 'Africa', 'and', 'several', 'hundred', 'thousand', 'in',  
 'Zimbabwe', 'as', 'a', 'second', 'language', 'in', '1975']



['Japanese', 'Total', 'signers', '320', ',', ',', '000', '(', '1986', ')']  
 ['Japanese', 'Regulated', 'by', 'Japanese', 'Federation', 'of', 'the', 'Deaf']  
 ['Japanese', 'Language', 'codes', 'ISO', '639-3', 'jsl', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Khanty', 'Total', 'speakers', '12', ',', ',', '000']  
 ['Mansi', 'Total', 'speakers', '2', ',', ',', '746', '(', '2002', ')']  
 ['Ter', 'Language', 'codes', 'ISO', '639-3', 'sjt', 'Ter', 'Sami', 'is', 'number', '9', 'on', 'the', 'map']  
 ['Ume', 'Total', 'speakers', '~', '10']  
 ['Ume', 'Language', 'codes', 'ISO', '639-3', 'sju', 'Ume', 'Sami', 'is', '2', 'on', 'this', 'map']  
 ['Pite', 'Total', 'speakers', '25', '~', '50']  
 ['Pite', 'Language', 'codes', 'ISO', '639-3', 'sje', 'Pite', 'Sami', 'is', '3', 'on', 'this', 'map']  
 ['Bunun', 'Total', 'speakers', '38', ',', ',', '000']  
 ['Negerhollands', 'Spoken', 'in', 'U', '.', 'S', '.', 'Virgin', 'Islands']  
 ['Angloromani', 'Total', 'speakers', '195', ',', ',', '000']  
 ['Bolaang', 'Mongondow', 'Total', 'speakers', '900', ',', ',', '000', '(', '1989', ')']  
 ['Ayacucho', 'Total', 'speakers', '1', ',', ',', '000', ',', ',', '000']  
 ['Cusco-Collao', 'Total', 'speakers', '4', ',', ',', '000', ',', ',', '000']  
 ['Cusco-Collao', 'Dialects', 'South', 'Bolivian', 'Quechua']  
 ['Cusco-Collao', 'Dialects', 'North', 'Bolivian', 'Quechua']  
 ['Motu', 'Total', 'speakers', '40', ',', ',', '000', '~', 'in', '2008']  
 ['Motu', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Makhuwa-Meetto', 'Total', 'speakers', '800', ',', ',', '000']  
 ['Acholi', 'Total', 'speakers', '800', ',', ',', '000']  
 ['Adjuvilo', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']  
 ['Adjuvilo', 'Setting', 'and', 'usage', 'International', 'auxiliary', 'language']  
 ['Maharashtri', 'Language', 'extinction', 'Dhivehi', 'and', 'Sinhalese']  
 ['Jain', 'Prakrit', 'Language', 'extinction', 'developed', 'into', 'Magadhi']  
 ['Lihir', 'Spoken', 'in', 'Lihir', 'Island', ',', ',', 'off', 'New', 'Ireland']  
 ['Lihir', 'Total', 'speakers', '13', ',', ',', '000']  
 ['Paeonian', 'Language', 'extinction', 'probably', '1st', 'century', 'BC--2nd', 'century', 'AD']  
 ['Mewari', 'Total', 'speakers', '5', '.', '2', 'million', '(', '2006', ')']  
 ['Ligurian', 'Total', 'speakers', '1', ',', ',', '925', ',', ',', '100']  
 ['Ligurian', 'Official', 'language', 'in', 'Officially', 'recognized', 'in', 'Italy', '(', 'Law', '482', '/', '1999', ')', 'and', 'Monaco']  
 ['Ligurian', 'Language', 'codes', 'Linguasphere', '51-AAA-oh', '&', 'amp', ';', '51-AAA-og']

['Carian', 'Language', 'extinction', 'Early', 'CE']  
 ['Carian', 'Spoken', 'in', 'Ancient', 'southwestern', 'Anatolia']  
 ['Arapaho', 'Total', 'speakers', '~', '1', ',', '000']  
 ['Auregnais', 'Language', 'extinction', 'by', 'mid-20th', 'century', '(', 'some', 'rememberers', ')']  
 ['Mundolinco', 'Setting', 'and', 'usage', 'International', 'auxiliary', 'language']  
 ['Mundolinco', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']  
 ['Selkup', 'Total', 'speakers', '1', ',', '570', '(', '1994', 'est', '.', ')']  
 ['Nganasan', 'Total', 'speakers', '505', '(', '2002', 'census', ')']  
 ['Middle', 'Welsh', 'Language', 'extinction', 'Evolved', 'into', 'Modern', 'Welsh', 'about', 'the', '15th', 'century']  
 ['Banyjima', 'Total', 'speakers', '50', '(', 'as', 'of', '1991', ')']  
 ['Banyjima', 'Region', 'Pilbara', 'region', 'of', 'Western', 'Australia']  
 ['Mazatec', 'Language', 'codes', 'ISO', '639-3', ':', 'maa', '\xe2', '\x80', '\x93', 'San', 'Jeronimo', 'Tec', '\xc3', '\xb3', 'atl', 'Mazatec']  
 ['Mazatec', 'Language', 'codes', 'ISO', '639-3', 'maj', '\xe2', '\x80', '\x93', 'Jalapa', 'de', 'D', '\xc3', '\xad', 'az', 'Mazatec']  
 ['Itelmen', 'Total', 'speakers', 'Fewer', 'than', '100']  
 ['Kapingamarangi', 'Region', 'Kapingamarangi', 'and', 'Ponape', 'islands']  
 ['Kangri', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Are', 'Total', 'speakers', '1', ',', '230']  
 ['Are', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Are', 'Region', 'Milne', 'Bay', 'Province', ',', 'tip', 'of', 'Cape', 'Vogel']  
 ['Megleno-Romanian', 'Total', 'speakers', '5', ',', '000']  
 ['Tartessian', 'Language', 'extinction', 'after', '5th', 'century', 'BCE']  
 ['Tartessian', 'Spoken', 'in', 'Southwest', 'Iberian', 'Peninsula']  
 ['Tartessian', 'Language', 'codes', 'ISO', '639-3', 'Tartessian', 'language', 'in', 'the', 'context', 'of', 'paleohispanic', 'languages']  
 ['Min', 'Dong', 'Total', 'speakers', '9', '.', '1', 'million']  
 ['Min', 'Spoken', 'in', 'United', 'States', '(', 'chiefly', 'New', 'York', 'City', ')']  
 ['Min', 'Region', 'eastern', 'Fujian', '(', 'Fuzhou', 'and', 'Ningde', ')']  
 ['Phuthi', 'Total', 'speakers', 'around', '20', ',', '000']  
 ['Phuthi', 'Spoken', 'in', 'Lesotho', ',', 'South', 'Africa']  
 ['South', 'Signed', 'in', 'South', 'Africa']  
 ['South', 'Total', 'signers', '&', 'lt', ';', '12', ',', '000']  
 ['South', 'Language', 'codes', 'ISO', '639-3', 'sfs', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Chilcotin', 'Total', 'speakers', '2', ',', '000']  
 ['Chilcotin', 'Region', 'Chilcotin', 'Country', ',', 'Central', 'Interior',  
 'of', 'British', 'Columbia']  
 ['Miskito', 'Total', 'speakers', '183', ',', '400']  
 ['Miskito', 'Region', 'North', 'Atlantic', 'Autonomous', ',',  
 'neighbouring', 'areas']  
 ['Sumo', 'Total', 'speakers', '7', ',', '000']  
 ['Cacaopera', 'Language', 'extinction', '20th', 'century']  
 ['Matagalpa', 'Language', 'extinction', '19th', 'century']  
 ['Karuk', 'Spoken', 'in', 'Northwestern', 'California', ',', 'USA']  
 ['Karuk', 'Total', 'speakers', '10', '(', '1997', ')', ',', '55', '(',  
 '2000', ')']  
 ['Kusaal', 'Total', 'speakers', '437', ',', '000']  
 ['Sidamo', 'Total', 'speakers', '2', ',', '981', ',', '471', '(', '2007',  
 'census', ')']  
 ['Sidamo', 'Region', 'region', 'of', 'Sidamo', 'Province', ',', 'south',  
 'central']  
 ['Deg', 'Xinag', ',', 'Deg', 'Hit', '\xe2', '\x80', '\x99', 'an', 'Total',  
 'speakers', '15']  
 ['Deg', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan',  
 'variant', ')']  
 ['Nahuatl', 'Total', 'speakers', 'over', '1', '.', '5', 'million']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', ':', 'nci', '\xe2',  
 '\x80', '\x93', 'Classical', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhn', '\xe2', '\x80',  
 '\x93', 'Central', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nch', '\xe2', '\x80',  
 '\x93', 'Central', 'Huasteca', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'ncx', '\xe2', '\x80',  
 '\x93', 'Central', 'Puebla', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhe', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Huasteca', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'azz', '\xe2', '\x80',  
 '\x93', 'Highland', 'Puebla', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhy', '\xe2', '\x80',  
 '\x93', 'Northern', 'Oaxaca', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'ncj', '\xe2', '\x80',  
 '\x93', 'Northern', 'Puebla', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nlv', '\xe2', '\x80',  
 '\x93', 'Orizaba', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'ppl', '\xe2', '\x80',  
 '\x93', 'Pipil', 'language']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhz', '\xe2', '\x80',  
 '\x93', 'Santa', 'Mar', '\xc3', '\xad', 'a', 'la', 'Alta', 'Nahuatl']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhs', '\xe2', '\x80',  
 '\x93', 'Southeastern', 'Puebla', 'Nahuatl']

['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhg', '\xe2', '\x80', '\x93', 'Tetelcingo', 'Nahuatl']

['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nhw', '\xe2', '\x80', '\x93', 'Western', 'Huasteca', 'Nahuatl']

['Pattani', 'Malay', 'Total', 'speakers', '3', ',', '100', ',', '000', '(', '1998', ')']

['Seri', 'Total', 'speakers', '900', 'or', '716']

['Classical', 'Nahuatl', 'Language', 'extinction', 'Divided', 'into', 'various', 'dialects', 'by', 'the', '15th', 'century']

['Monguor', 'Total', 'speakers', '152', ',', '000']

['Yagua', 'Total', 'speakers', '5', ',', '700', '(', 'in', '2000', ')']

['Khowar', 'Total', 'speakers', '400', ',', '000']

['Khowar', 'Writing', 'system', 'Khowar', 'alphabet', '(', 'Nasta', '\xca', '\xbf', 'l', '\xc4', '\xab', 'q', 'script', ')', ',', 'see', 'other', 'less-used', 'writing', 'systems', 'below']

['Khowar', 'Regulated', 'by', '(', 'Pakistan', ')', ';', 'Literary', 'Association', 'for', 'Promotion', 'of', 'Khowar', 'language']

['Fuzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China', '(', 'Fuzhou', 'and', 'its', 'surrounding', 'counties', ')', ';', 'Republic', 'of', 'China', '(', 'Matsu', 'Islands', ')', ';', 'Malaysia', '(', 'Sibu']

['Fuzhou', 'Spoken', 'in', 'Sitiawan', 'and', 'Ayer', 'Tawar', ')', ';', 'Indonesia', ';', 'Singapore', ';', 'and', 'some', 'Chinese', 'communities', 'in', 'Southeast', 'Asia', 'and', 'the', 'west']

['Fuzhou', 'Spoken', 'in', 'particularly', 'in', 'the', 'Chinatowns', 'of', 'New', 'York', 'City']

['Fuzhou', 'Language', 'codes', 'ISO', '639-3', '1', ':', 'Fuzhou', 'City', 'Proper']

['Dameli', 'Total', 'speakers', '5', ',', '000']

['Dameli', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Kalasha', 'Total', 'speakers', '5', ',', '029', '(', '2000', ')', '6', ',', '000', '(', '2010', ')']

['Palula', 'Total', 'speakers', '8', ',', '600', '(', '1990', ')', ',', '10', ',', '000', '(', '2010', ')', 'Source', 'Khowar', 'Academy']

['Palula', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Gawar-Bati', 'Total', 'speakers', '10', ',', '000']

['Gawar-Bati', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Yidgha', 'Total', 'speakers', '6', ',', '145', '(', '2000', ')', ',', '145', '(', '2010', ')']

'8', ',', '000', '(', '2010', ')', '(', 'Source', ':', 'Khowar', 'Academy', ')']

['Panare', 'Total', 'speakers', '3', ',', '000', '\xe2', '\x80', '\x93', '4', ',', '000']

['Brij', 'Bhasa', 'Total', 'speakers', '11', ',', '085', ',', '230']

['Angika', 'Total', 'speakers', '740', ',', '892', '(', 'of', 'whom', 'approx', '.', '725', ',', '000', 'in', 'India', ',', '15', ',', '892', 'in', 'Nepal', ')', '(', 'SIL', ')']

['Angika', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Kambera', 'Total', 'speakers', '10', ',', '000', '(', '1981', ')']

['Algerian', 'Total', 'speakers', '20', ',', '000', ',', '000', '+']

['Fuy\xc3\xbc', 'G', '\xc3', '\xaf', 'rg', '\xc3', '\xaf', 's', 'Total', 'speakers', '&', 'lt', ';', '10']

['Nyoro', 'Total', 'speakers', '700', ',', '000']

['Carrier', 'Total', 'speakers', '1', ',', '000']

['Carrier', 'Region', 'Central', 'Interior', 'of', 'British', 'Columbia']

['Carrier', 'Language', 'codes', 'ISO', '639-3', 'caf', '\xe2', '\x80', '\x93', 'Southern', 'Carrier']

['Kaqchikel', '(', 'Cakchiquel', ')', 'Total', 'speakers', 'approx', '.', '500', ',', '000']

['Atakapa', 'Language', 'extinction', '20th', 'century']

['Atakapa', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Atakapa', 'language']

['Esselen', 'Language', 'extinction', '19th', 'century']

['Esselen', 'Region', 'Big', 'Sur', '(', 'California', ')']

['Karankawa', 'Region', 'Texas', 'coast', ',', 'from', 'Galveston', 'Island', 'to', 'Corpus', 'Christi', 'Ethnicity']

['Karankawa', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Karankawa', 'language']

['Kaw\xc3\xa9sqar', 'Region', 'Wellington', 'Island', 'off', 'south', 'Chilean', 'coast']

['Indo-Pakistani', 'Total', 'signers', '2', '.', '7', 'million', 'in', 'India', ';', 'number', 'in', 'Pakistan', 'and', 'Bangladesh', 'unknown']

['Indo-Pakistani', 'Language', 'codes', 'ISO', '639-3', 'pks', '\xe2', '\x80', '\x93', 'Pakistani', 'Sign', 'Language', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Leonese', 'language', 'Total', 'speakers', 'From', '25', ',', '000', 'to', '50', ',', '000']

['Toda', 'Total', 'speakers', '1', ',', '100']

['Toda', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',

'conjuncts', '.', 'More']  
 ['Malvani', 'Total', 'speakers', '8', ',', '68', ',', '825']  
 ['Malvani', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Neo', 'Regulated', 'by', 'Akademio', 'de', 'Neo', '(', 'now', 'defunct',  
 ')']  
 ['Neo', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Neo', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Martuthunira', 'Language', 'extinction', '6', 'August', '1995', 'with',  
 'the', 'death', 'of', 'Algy', 'Paterson']  
 ['Gutob', 'Total', 'speakers', '8', ',', '000', '(', '2000', ')']  
 ['Paic\xc3\xae', 'Total', 'speakers', '5', ',', '500']  
 ['Paic\xc3\xae', 'Region', 'East', 'coast', 'between', 'Poindimi', '\xc3',  
 '\xa9', 'and', 'Pon', '\xc3', '\xa9', 'rihouen', 'and', 'inland',  
 'valleys']  
 ['Shor', 'Total', 'speakers', '9', ',', '800']  
 ['Afrihili', 'Category', '(', 'sources', ')', 'related', 'to', 'various',  
 'African', 'languages']  
 ['Shina', 'Region', 'Northern', 'Areas', ',', 'Pakistan']  
 ['Bungi', 'Total', 'speakers', '&', 'lt', ';', '\xe2', '\x80', '\x89',  
 '500', 'likely', 'extinct']  
 ['Dahalik', 'Total', 'speakers', '2', ',', '500', '\xe2', '\x80', '\x93',  
 '3', ',', '000']  
 ['Genoese', 'Language', 'codes', 'Linguasphere', '51-AAA-ohd', '.', '.',  
 '.', '-objb']  
 ['Tuscarora', 'Total', 'speakers', '&', 'lt', ';', '50']  
 ['Tuscarora', 'Region', 'Tuscarora', 'Reservation', 'in', 'northwestern',  
 'New', 'York']  
 ['Tuscarora', 'Region', 'and', 'eastern', 'North', 'Carolina']  
 ['Seneca', 'Region', 'Western', 'New', 'York', 'and', 'the', 'Six',  
 'Nations', 'Reserve', ',', 'Ontario']  
 ['Onondaga', 'Region', 'and', 'central', 'New', 'York']  
 ['Canadian', 'Gaelic', 'Total', 'speakers', 'ca', '.', '2000', 'or',  
 'roughly', '1', '.', '3', '%', 'of', 'Cape', 'Breton', 'Island']  
 ['Canadian', 'Region', 'limited', 'to', 'Cape', 'Breton', 'Island']  
 ['Canadian', 'Region', 'Newfoundland', '&', 'amp', ';', 'Labrador', '(',  
 'Codroy', 'Valley', ')']  
 ['Canadian', 'Region', 'Bruce', 'and', 'Grey', 'Counties', ')']  
 ['Canadian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Gaelic',  
 'variant', ')']  
 ['Canadian', 'Language', 'codes', 'ISO', '639-3', 'Distribution',  
 'throughout', 'the', 'Maritimes', 'c', '.', '1850']

['Chenchu', 'Total', 'speakers', '28', ',', '754', '(', '1981', 'census', ')']

['Chenchu', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Khalaj', 'Total', 'speakers', '42', ',', '107', '(', '2000', ')']

['Khalaj', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']

['\xc7\x83Kung', 'Total', 'speakers', '~', '45', ',', '000']

['Middle', 'Irish', 'Spoken', 'in', 'Ireland']

['Middle', 'Language', 'extinction', 'Evolved', 'into', 'Early', 'Modern', 'Irish', 'about', 'the', '12th', 'century']

['Halang', 'Total', 'speakers', '17', ',', '500']

['Sika', 'Total', 'speakers', '180', ',', '000']

['Udi', 'Region', 'Russia', '(', 'North', 'Caucasus', ')']

['Lara\xe2\x80\x99', 'Total', 'speakers', '19', ',', '000']

['Parthian', 'Language', 'extinction', 'Marginalized', 'by', 'Middle', 'Persian', 'from', 'the', '3rd', 'century']

['Swampy', 'Cree', 'Total', 'speakers', '4', ',', '500', '(', '1982', ')']

['Swampy', 'Language', 'codes', 'Linguasphere', '62-ADA-ac', ',', '62-ADA-ad']

['Lithuanian', 'Language', 'codes', 'ISO', '639-3', 'lls', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Sukur', 'Total', 'speakers', '15', ',', '000', '(', '1992', ')']

['Haida', 'Spoken', 'in', 'Canada', '(', 'Haida', 'Gwaii', '/', 'Queen', 'Charlotte', 'Islands', ')', ',', 'Alaska', '(', 'Prince', 'of', 'Wales', 'Island', ')']

['Haida', 'Total', 'speakers', 'First', 'language', ':', '40']

['Haida', 'Total', 'speakers', 'Second', 'language', ':', '275']

['Haida', 'Official', 'language', 'in', 'Council', 'of', 'the', 'Haida', 'Nation']

['Haida', 'Language', 'codes', 'ISO', '639-3', ':', 'hdn', '\xe2', '\x80', '\x93', 'Northern', 'Haida']

['Haida', 'Language', 'codes', 'ISO', '639-3', 'hax', '\xe2', '\x80', '\x93', 'Southern', 'Haida']

['Lari', 'Total', 'speakers', 'ca', '.', '0', '.', '3', 'million']

['Lari', 'Region', 'Southern', 'Iran', '(', 'Mainly', ':', 'Fars', 'and', 'Hormozgan', 'provinces', ')']

['South', '\xc3', '\x89', 'fat', '\xc3', '\xa9', 'Total', 'speakers', '6', ',', '000']

['Yazgulyam', 'Total', 'speakers', '4', ',', '000', '(', '1994', ')']

['Cotoname', 'Language', 'extinction', 'ca', '.', '1900']

['Cotoname', 'Language', 'codes', 'ISO', '639-3', 'Map', 'indicating', 'where', 'Cotoname', 'was', 'spoken']

['Chuukese', 'Total', 'speakers', '45', ',', '000']

['Tsangla', 'Spoken', 'in', 'E', 'Bhutan', ':', '140', ',', '000']  
 ['Tsangla', 'Spoken', 'in', 'Bom', '\xc3', '\xaa', 'County']  
 ['Tsangla', 'Spoken', 'in', 'Tibet', ':', '7']  
 ['Tsangla', 'Total', 'speakers', '175', ',', '200', '(', '2009', 'est',  
 '.', ')']  
 ['Tai', 'Total', 'speakers', 'about', '700', ',', '000']  
 ['Tai', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Basaa', 'Total', 'speakers', '230', ',', '000']  
 ['Bassa', 'Spoken', 'in', 'Liberia', ',', 'Sierra', 'Leone']  
 ['Bassa', 'Total', 'speakers', '409', ',', '000', '(', '2006', ')']  
 ['Unami', 'Spoken', 'in', 'Eastern', 'United', 'States']  
 ['Unami', 'Region', 'Around', 'the', 'lower', 'Delaware', 'and', 'Hudson',  
 'rivers', 'in', 'the', 'United', 'States', ';', 'some', 'Unami', 'groups',  
 'in', 'Oklahoma']  
 ['Munsee', 'Total', 'speakers', '7-8', '(', '1991', ')']  
 ['Munsee', 'Region', 'now', 'in', 'Ontario', ';', 'formerly', 'in', 'New',  
 'York']  
 ['Tanapag', 'Total', 'speakers', '4', ',', '400']  
 ['Tanapag', 'Spoken', 'in', 'Northern', 'Mariana', 'Islands']  
 ['Hupa', 'Total', 'speakers', 'Between', '5', 'and', '20']  
 ['Huastec', 'Total', 'speakers', 'approx', '.', '150', ',', '000']  
 ['Huastec', 'Region', 'San', 'Luis', 'Potos', '\xc3', '\xad', ',',  
 'Veracruz', 'and', 'Tamaulipas']  
 ['Unserdeutsch', 'Spoken', 'in', 'Papua', 'New', 'Guinea', ',',  
 'Australia']  
 ['Unserdeutsch', 'Total', 'speakers', '&', 'lt', ';', '100']  
 ['Plains', 'Cree', 'Total', 'speakers', '34', ',', '000']  
 ['Kharia', 'Spoken', 'in', 'India', '(', 'Jharkhand', ',', 'Chhattisgarh',  
 ',', 'Orissa', ',', 'West', 'Bengal', ',', 'Assam', ',', 'Tripura', ',',  
 'Andaman', 'and', 'Nicobar', 'Islands', ')', ',', 'Nepal']  
 ['Kharia', 'Total', 'speakers', '293', ',', '575']  
 ['Miwok', 'Spoken', 'in', 'California', ',', 'western', 'slopes', 'of',  
 'Sierra', 'Nevadas', 'Ethnicity']  
 ['Dinka', 'Spoken', 'in', 'Southern', 'Sudan', 'and', 'neighboring',  
 'areas']  
 ['Dinka', 'Language', 'codes', 'ISO', '639-3', ':', 'dip', '\xe2', '\x80',  
 '\x93', 'Northeastern', 'Dinka', '(', 'Padang', ')']  
 ['Dinka', 'Language', 'codes', 'ISO', '639-3', 'diw', '\xe2', '\x80',  
 '\x93', 'Northwestern', 'Dinka', '(', 'Ruweng', ')']  
 ['Dinka', 'Language', 'codes', 'ISO', '639-3', 'dib', '\xe2', '\x80',  
 '\x93', 'South', 'Central', 'Dinka', '(', 'Agar', ')']  
 ['Dinka', 'Language', 'codes', 'ISO', '639-3', 'dks', '\xe2', '\x80',  
 '\x93', 'Southeastern', 'Dinka', '(', 'Bor', ',', 'Twic', '/', 'Tuic', ')']



['Dinka', 'Language', 'codes', 'ISO', '639-3', 'dik', '\xe2', '\x80',  
 '\x93', 'Southwestern', 'Dinka', '(', 'Rek', ')']  
 ['Massa', 'Total', 'speakers', '212', ',', ',', '000']  
 ['Skolt', 'Total', 'speakers', '~', '400']  
 ['Skolt', 'Language', 'codes', 'ISO', '639-3', 'Skolt', 'Sami', 'is', '6',  
 'on', 'this', 'regional', 'map', 'of', 'Sami', 'languages']  
 ['Ngunnawal', 'Spoken', 'in', 'New', 'South', 'Wales', '&', 'amp', ';',  
 'ACT', ',', 'Australia']  
 ['Ngunnawal', 'Language', 'codes', 'ISO', '639-3', '08t', '\xe2', '\x80',  
 '\x93']  
 ['Kosraean', 'Total', 'speakers', '8', ',', ',', '000']  
 ['Potawatomi', 'Language', 'codes', 'Linguasphere', '62-ADA-dc', '(',  
 'Potawatomi', ')']  
 ['Car', 'Total', 'speakers', '22', ',', ',', '100']  
 ['Al-Sayyid', 'Language', 'codes', 'ISO', '639-3', 'syy', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Kim', 'Official', 'language', 'in', 'China', '(', 'in', 'Jinxu', 'Yao',  
 'Autonomous', 'County', ')']  
 ['Kwanyama', 'Total', 'speakers', '421', ',', ',', '000', 'in', 'Angola', '&',  
 'amp', ';', 'over', '250', ',', ',', '000', 'in', 'Namibia']  
 ['Kamba', 'Total', 'speakers', '2', ',', ',', '448', ',', ',', '300', '(', '1989',  
 'census', ')']  
 ['Awabakal', 'Language', 'extinction', 'Sometime', 'late', 'in', 'the',  
 '19th', 'century', '.', 'The', 'language', 'is', 'currently', 'in',  
 'early', 'stages', 'of', 'revival']  
 ['Awabakal', 'Spoken', 'in', 'Eastern', 'New', 'South', 'Wales', ',',  
 'Australia']  
 ['Day', 'Total', 'speakers', '49', ',', ',', '916']  
 ['Bathari', 'Total', 'speakers', 'about', '200']  
 ['Isthmus\xe2\x80\x93Mecayapan', 'Nahuatl', 'Total', 'speakers', '~', '20',  
 ',', ',', '000']  
 ['Yidiny', 'Total', 'speakers', '12', '(', '1981', ')']  
 ['Ndonga', 'Total', 'speakers', '690', ',', ',', '000']  
 ['Lan-nang', 'Total', 'speakers', '592', ',', ',', '200', '(', '98', '.', '7',  
 '%', 'of', 'all', 'Chinese', 'in', 'the', 'Philippines', ')']  
 ['Lan-nang', 'Region', 'and', 'other', 'Philippine', 'communities', 'with',  
 'a', 'substantial', 'Chinese', 'minority']  
 ['Providence', 'Language', 'codes', 'ISO', '639-3', 'prz', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Urubu', 'Language', 'codes', 'ISO', '639-3', 'uks', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Mono', 'Total', 'speakers', '65', ',', ',', '000', '(', '1984', ')']  
 ['Mono', 'Region', 'Northwestern', 'corner', 'of', 'Congo', '(', 'DRC',

')]  
 ['Mono', 'Total', 'speakers', '87-91', '(', '1994', ')']  
 ['Nez', 'Perce', 'Total', 'speakers', '200']  
 ['Aji\xc3\xab', 'Spoken', 'in', 'Houailou', ',', 'New', 'Caledonia']  
 ['Aji\xc3\xab', 'Total', 'speakers', '4', ',', '044', '(', '1996',  
 'census', ')']  
 ['Badimaya', 'Total', 'speakers', 'Possibly', '4']  
 ['Badimaya', 'Region', 'Murchison', 'area', 'of', 'Western', 'Australia']  
 ['Comanche', 'Region', 'Oklahoma', '(', 'formerly', ',', 'Texas', ',',  
 'New', 'Mexico', ',', 'Kansas', ',', 'Colorado', ',', 'Oklahoma', ')']  
 ['Mehri', 'Total', 'speakers', '135', ',', '800']  
 ["Ngan'gikurunggurr", 'Total', 'speakers', '275', '(', '1988', 'SIL', ')']  
 ['Vai', 'Spoken', 'in', 'Liberia', ',', 'Sierra', 'Leone']  
 ['Vai', 'Total', 'speakers', '119', ',', '500']  
 ['Shawnee', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'the', 'Shawnee', 'language', 'around', '1650']  
 ['Bamum', 'Total', 'speakers', '215', ',', '000', '(', '1982', ')']  
 ['Bamum', 'Language', 'codes', 'ISO', '639-3', 'Page', 'from', 'a',  
 'manuscript', 'in', 'the', 'Bamum', 'script']  
 ['Ngadha', 'Total', 'speakers', '60', ',', '000', '(', '1995', ')']  
 ['Kriol', 'Total', 'speakers', '10', ',', '000', '(', '1991', 'Borneman',  
 ')']  
 ['Kriol', 'Region', 'Northern', 'Territory', ';', 'Kimberley', 'region',  
 'of', 'Western', 'Australia', ';', 'Gulf', 'Country']  
 ['Kriol', 'Region', 'Lower', 'Cape', 'York', 'Peninsula']  
 ['Kriol', 'Language', 'codes', 'Linguasphere', '52-ABB-ca', '(',  
 'varieties']  
 ['Kriol', 'Language', 'codes', 'Linguasphere', ':', '52-ABB-caa', 'to', '-  
 caf']  
 ['Ngalakan', 'Total', 'speakers', '10', 'Wurm', 'and', 'Hattori', '(',  
 '1981', ')', 'as', 'cited', 'in', 'the', 'Ethnologue']  
 ['Lotha', 'Total', 'speakers', '80', ',', '000', '(', '1997', ')']  
 ['Adhola', 'Total', 'speakers', '~', '250', ',', '000']  
 ['Bobot', 'Total', 'speakers', '4', ',', '500']  
 ['Bobot', 'Region', 'from', 'the', 'village', 'of', 'Atiahu', 'to', 'Kota',  
 'Baru']  
 ['Wajarri', 'Total', 'speakers', 'Possibly', '25']  
 ['Wajarri', 'Region', 'Murchison', 'area', 'of', 'Western', 'Australia']  
 ['Tundra', 'Yukaghir', 'Total', 'speakers', '30', '\xe2', '\x80', '\x93',  
 '150']  
 ['Khams', 'Total', 'speakers', '1', '.', '5', 'million', '(', '1994', ')']  
 ['Flemish', 'Total', 'signers', '6', ',', '000', '(', '2005', ')']  
 ['Flemish', 'Language', 'codes', 'ISO', '639-3', 'vgt', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Igala', 'Total', 'speakers', '800', ',', '000', '(', '1989', ')']

['Nyaw', 'Total', 'speakers', '50', ',', '000']  
 ['Nyaw', 'Region', 'Isan', ',', 'Mekong', 'floodplain']  
 ['Phu', 'Total', 'speakers', '519', ',', '400', '(', '2000', ')']  
 ['Bergamasque', 'Total', 'speakers', '~', '700', ',', '000']  
 ['Bergamasque', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80', '\x93', 'Lombard', 'language']  
 ['Nyangumarta', 'Total', 'speakers', '520', '(', '1991', ')']  
 ['Natchez', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Natchez', 'language']  
 ['Teso', 'Total', 'speakers', '1', ',', '278', ',', '537', '.', '(', '1991', ')']  
 ['Xamtanga', 'Total', 'speakers', '213', ',', '355', '(', '2007', 'census', ')']  
 ['Mongsen', 'Ao', 'Total', 'speakers', '141', ',', '000']  
 ['Western', 'Desert', 'Language', 'Total', 'speakers', 'several', 'thousand']  
 ['Western', 'Region', 'Desert', 'areas', 'of', 'Western', 'Australia', ',', 'South', 'Australia', 'and', 'the', 'Northern', 'Territory']  
 ['Atayal', 'Total', 'speakers', '&', 'lt', ';', '84', ',', '330']  
 ['Turkish', 'Signed', 'in', 'Turkey', ',', 'Northern', 'Cyprus']  
 ['Turkish', 'Language', 'codes', 'ISO', '639-3', 'tsm', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Malgana', 'Language', 'extinction', 'not', 'known']  
 ['Malgana', 'Spoken', 'in', 'Shark', 'Bay', 'area', 'of', 'Western', 'Australia']  
 ['Nhanda', 'Region', 'Geraldton', 'to', 'Shark', 'Bay', 'area', 'of', 'Western', 'Australia']  
 ['Jam', 'Total', 'speakers', '130', ',', '000', '(', '1998', ')']  
 ['East', 'Central', 'German', 'Total', 'speakers', '\xe2', '\x80', '\x93']  
 ['Kambera', 'Total', 'speakers', '234', ',', '574', '(', '2000', ')']  
 ['Palu'e', 'Total', 'speakers', '10', ',', '000']  
 ['Soninke', 'Total', 'speakers', '1', ',', '096', ',', '795']  
 ['Soninke', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']  
 ['Even', 'Total', 'speakers', '7', ',', '168']  
 ['Dhatki', 'Total', 'speakers', '150', ',', '400', '(', '1997', ')']  
 ['T\xc3\xa2y', 'Total', 'speakers', '~', '1', ',', '000']  
 ['Bishnupriya', 'Spoken', 'in', 'Myanmar', 'and', 'several', 'other', 'countries']  
 ['Bishnupriya', 'Total', 'speakers', '450', ',', '000']  
 ['Bishnupriya', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Puma', 'Total', 'speakers', '4', ',', '310']  
 ['Akatek', 'Total', 'speakers', '58', ',', '600', '(', '1998', ')']

['Tonga', 'Total', 'speakers', '1', ',', '517', ',', '000', 'one', 'of', 'the', 'key', 'lingua', 'franca', 'in', 'Zambia', 'and', 'parts', 'of', 'Zimbabwe', 'and', 'Mozambique', ')]']

['Hungarian', 'Language', 'codes', 'ISO', '639-3', 'hsh', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Torlakian', 'Language', 'codes', 'ISO', '639-3', 'Areas', 'where', 'Torlakian', 'dialects', 'are', 'spoken', ',', 'with', 'ethnic', 'divisions', ':', '(', '1', ')', 'Serbian', ',', '(', '2', ')', 'Bulgarian', ',', '(', '3', ')', 'Macedonian', ',', '(', '4', ')', 'Gorani', '(', 'Muslim', ')', ',', '(', '5', ')', 'Krashovani', '(', 'Croatian', ')', ',', '(', '6', ')', 'Torlakian', 'Serb', 'areas', 'of', 'Kosovo', 'prior', 'to', '1999']

['Au', 'Spoken', 'in', 'Sandaun', ',', 'Papua', 'New', 'Guinea']

['Au', 'Total', 'speakers', 'approx', '.', '8', ',', '000']

['Kriol', 'Spoken', 'in', 'Belize', 'approx', '.', '250', ',', '000']

['Kriol', 'Spoken', 'in', 'United', 'States', 'approx', '.', '100', ',', '000']

['Kriol', 'Spoken', 'in', 'Nicaragua', 'approx', '.', '30', ',', '000']

['Kriol', 'Total', 'speakers', 'First', 'language', ':', 'approx', '.', '200', ',', '000']

['Kriol', 'Total', 'speakers', 'Second', 'language', ':', '227', ',', '140']

['Danzhouhua', 'Total', 'speakers', '~', '400', ',', '000']

['Danzhouhua', 'Region', 'Hainan', '(', 'Danzhou', ')]']

['Gonja', 'Total', 'speakers', '230', ',', '000']

['Tillamook', 'Language', 'extinction', 'Last', 'speaker', 'died', 'circa', '1970', ';', 'the', 'ethnic', 'population', 'now', 'uses', 'English', 'as', 'their', 'native', 'language']

['Harari', 'Total', 'speakers', '25', ',', '810', '(', '2007', 'census', ')]']

["Silt'e", 'Total', 'speakers', '827', ',', '764', '(', '1998', 'census', ')]']

["Silt'e", 'Language', 'codes', 'ISO', '639-3', ':', 'stv', '\xe2', '\x80', '\x93', "Silt'e"]

['Bohtan', 'Neo-Aramaic', 'Total', 'speakers', '1', ',', '000']

['Maria', 'Total', 'speakers', '165', ',', '000']

['Kwanga', 'Total', 'speakers', '10', ',', '000']

['Kwanga', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Kwanga', 'Region', 'East', 'Sepik', 'and', 'Sandaun']

['Kam', 'Total', 'speakers', '1', '.', '5', 'million']

['Kam', 'Language', 'codes', 'ISO', '639-3', ':', 'doc', '\xe2', '\x80', '\x93', 'Northern', 'Dong']

['Kam', 'Language', 'codes', 'ISO', '639-3', 'kmc', '\xe2', '\x80', '\x93', 'Southern', 'Dong']

['Vlax', 'Total', 'speakers', '1', '.', '5', 'million']

['Tiwi', 'Total', 'speakers', '1', ',', '500']  
 ['Tiwi', 'Region', 'Bathurst', 'and', 'Melville', 'Islands', ',', 'Northern', 'Territory']  
 ['Tiwi', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of', 'Tiwi', 'lands', 'in', 'red', '.', '(', 'Dark', 'lines', 'are', 'estimated', 'language', 'boundaries', 'before', 'colonization', '.', ')']  
 ["Kwak'wala", 'Spoken', 'in', 'Northern', 'Vancouver', 'Island', ',', 'Canada']  
 ["Kwak'wala", 'Total', 'speakers', '~', '250']  
 ['Cal'xc3xb3', 'Total', 'speakers', '40', ',', '000', '~', '70', ',', '000']  
 ['Tetet'xc3xa9', 'Language', 'extinction', '?', '?', '?']  
 ['Nar', 'Phu', 'Total', 'speakers', '800']  
 ['Amuzgo', 'Total', 'speakers', '&', 'gt', ';', '49', ',', '000']  
 ['Amuzgo', 'Language', 'codes', 'ISO', '639-3', ':', 'amu', '\xe2', '\x80', '\x93', 'Northern', '(', 'Guerrero', ')', 'Amuzgo']  
 ['Amuzgo', 'Language', 'codes', 'ISO', '639-3', 'azg', '\xe2', '\x80', '\x93', 'San', 'Pedro', 'Amuzgos', 'Amuzgo']  
 ['Amuzgo', 'Language', 'codes', 'ISO', '639-3', 'ISO', 'code', 'no', 'ISO', 'code', '\xe2', '\x80', '\x93', 'Southern', 'Amuzgo']  
 ['Loxian', 'Total', 'speakers', 'Two', '(', 'Enya', 'and', 'Roma', 'Ryan', ')']  
 ['Loxian', 'Category', '(', 'sources', ')', 'Welsh', 'and', 'Siberian', 'Yupik']  
 ['Loxian', 'Setting', 'and', 'usage', "'Beings', 'on', 'another', 'planet'"]  
 ['Jersey', 'Total', 'speakers', '17', ',', '000']  
 ['South', 'Total', 'speakers', '3', '.', '7', 'million']  
 ['Mayo', 'Total', 'speakers', '40', ',', '000']  
 ['Assiniboine', 'Total', 'speakers', '200', '\xe2', '\x80', '\x93', '250']  
 ['Assiniboine', 'Region', 'Southern', 'Saskatchewan', 'in', 'Canada', 'and', 'Montana', 'in', 'the', 'United', 'States']  
 ['Sasak', 'Total', 'speakers', '2', ',', '100', ',', '000', '(', '1989', ')']  
 ['Honduras', 'Language', 'codes', 'ISO', '639-3', 'hds', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Ida\xe2\x80\x99an', 'Total', 'speakers', '~', '6', ',', '000']  
 ['Mlabri', 'Total', 'speakers', 'c', '.', '340', '(', '2000', ')']  
 ['Tuamotuan', 'Total', 'speakers', '8700', '~']  
 ['Liburnian', 'Language', 'extinction', 'during', '1st', 'century', 'AD']  
 ['Lak', 'Total', 'speakers', '157', ',', '000', '(', '2002', ')']  
 ['Mbundu', 'Total', 'speakers', '3', ',', '000', ',', '000', '(', '1999', ')']  
 ['Khakas', 'Total', 'speakers', '~', '60', ',', '000']  
 ['Anus', 'Total', 'speakers', '70', '(', '2005', ')']

['Dolgan', 'Total', 'speakers', '~', '5', ',', '000']  
 ['Dolgan', 'Language', 'codes', 'ISO', '639-3', 'Map', 'showing',  
 'locations', 'of', 'Sakha', '(', 'dark', 'blue', ')', 'and', 'Dolgan', '(',  
 'blue', ')']  
 ['Eteocypriot', 'Language', 'extinction', 'effectively', 'extinct', 'from',  
 'about', 'the', 'beginning', 'of', 'the', '4th', 'century', 'BCE', ',', '']  
 ['Irula', 'Total', 'speakers', '4', ',', '500', '(', 'in', '1961', ')']  
 ['Tadaksahak', 'Total', 'speakers', '31', ',', '800']  
 ['Nuu-chah-nulth', 'Total', 'speakers', '150', '\xe2', '\x80', '\x93',  
 '200']  
 ['Nuu-chah-nulth', 'Region', 'West', 'coast', 'of', 'Vancouver', 'Island']  
 ['Babungo', 'Total', 'speakers', '14', ',', '000']  
 ['Babungo', 'Region', 'West', 'Africa', '/', 'Central', 'Africa']  
 ["Q'eqchi'", 'Total', 'speakers', '423', ',', '500']  
 ["Q'eqchi'", 'Total', 'speakers', '12', ',', '300', 'speakers', 'in', 'El',  
 'Salvador']  
 ["Q'eqchi'", 'Region', 'El', 'Quich', '\xc3', '\xa9', ';', 'Toledo']  
 ['Groma', 'Total', 'speakers', '12', ',', '840', 'in', 'China', '(',  
 '1993', ')']  
 ['Gurung', 'Language', 'codes', 'ISO', '639-3', ':', 'ggn', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Gurung']  
 ['Gurung', 'Language', 'codes', 'ISO', '639-3', 'gvr', '\xe2', '\x80',  
 '\x93', 'Western', 'Gurung']  
 ['Lepcha', 'Spoken', 'in', 'Greater', 'Sikkim', '(', 'Sikkim', ',',  
 'Darjeeling', 'and', 'Kalimpong', ',', 'India', ')', ';', 'parts', 'of',  
 'Nepal', 'and', 'Bhutan']  
 ['Lepcha', 'Total', 'speakers', '30', ',', '000', '\xe2', '\x80', '\x93',  
 '52', ',', '830', '(', '2007', ')']  
 ['Limbu', 'Spoken', 'in', 'Nepal', ';', 'Bhutan', ';', 'Sikkim']  
 ['Limbu', 'Spoken', 'in', 'Kashmir', 'and', 'Darjeeling', 'district',  
 'West', 'Bengal']  
 ['Limbu', 'Total', 'speakers', '333', ',', '633', 'in', 'Nepal', '(',  
 '2001', 'census', ')']  
 ['Limbu', 'Total', 'speakers', '~', '28', ',', '000', 'in', 'India', '(',  
 '1997', ')']  
 ['Limbu', 'Total', 'speakers', '~', '8000', 'in', 'Pakistani', 'Kashmir',  
 '(', '1998', ')']  
 ['Magar', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Dhanwar', 'Total', 'speakers', '32', ',', '000']  
 ['Sherpa', 'Total', 'speakers', 'ca', '.', '150', ',', '000', '(', '2000',  
 ')']  
 ['Sunuwar', 'Total', 'speakers', 'one', 'million', 'in', 'Nepal', 'and',  
 '27', ',', '000', 'in', 'Sikkim']

['Thulung', 'Total', 'speakers', '33', ',', '313']  
 ['Yakkha', 'Total', 'speakers', '14', ',', '000']  
 ['Norwegian', 'Total', 'signers', '180', ',', '000', 'in', 'Madagascar',  
 '(', '2006', ')', 'and', '4000', 'in', 'Norway', '(', '1986', ')']  
 ['Norwegian', 'Language', 'codes', 'ISO', '639-3', 'mzc', '\xe2', '\x80',  
 '\x93', 'Malagasy', 'sign', 'language', '\xe2', '\x80', '\x94', 'list',  
 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Iu', 'Mien', 'Total', 'speakers', '2', ',', '172', ',', '000']  
 ['Iu', 'Official', 'language', 'in', 'China', '(', 'in', 'Jinxiu', 'Yao',  
 'Autonomous', 'County', ')']  
 ['Naukan', 'Yupik', 'Total', 'speakers', 'approximately', '70']  
 ['Pecheneg', 'Language', 'extinction', '12th', 'century']  
 ['M\xc6\xb0\xe1\xbb\x9dng', 'Total', 'speakers', '1', '.', '2', 'million']  
 ['Tongva', 'Language', 'extinction', '1970s', '?']  
 ['Balkan', 'Total', 'speakers', '331', ',', '000']  
 ['Anal', 'Total', 'speakers', '13', ',', '853']  
 ['Libyan', 'Total', 'speakers', '4', '.', '3', 'million', '(', 'SIL',  
 'Ethnologue', ')']  
 ['Libyan', 'Language', 'codes', 'ISO', '639-3', 'Extension', 'of',  
 'Libyan', 'Arabic', '(', 'es', ':', 'Usuario', ':', 'Fobos92', '2010', ')']  
 ['Gamilaraay', 'Spoken', 'in', 'Central', 'northern', 'New', 'South',  
 'Wales']  
 ['Gamilaraay', 'Total', 'speakers', '3', '(', '1997', ')']  
 ['Wiradjuri', 'Total', 'speakers', '3', '(', '1981', ')']  
 ['Old', 'Welsh', 'Language', 'extinction', 'Evolved', 'into', 'Middle',  
 'Welsh', 'about', 'the', '12th', 'century']  
 ['Hill', 'Mari', 'Total', 'speakers', '36', ',', '822']  
 ['Hill', 'Region', 'Mari', 'El', '(', 'Gornomariysky', ',', 'Yurinsky',  
 ',', 'Kilemarsky', 'districts', ')']  
 ['Gao-Yang', 'Language', 'codes', 'ISO', '639-3', 'Gao-Yang', 'is', 'red',  
 ',', 'in', 'the', 'middle', 'of', 'the', 'Yue', 'area']  
 ['Kele', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Munichi', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of',  
 'Munichi', 'language']  
 ['Garifuna', 'Total', 'speakers', 'approx', '.', '300', ',', '000']  
 ['Garifuna', 'Spoken', 'in', 'Nicaragua's', 'Mosquito', 'Coast']  
 ['Garifuna', 'Region', 'Historically', 'the', 'Northern', 'Caribbean',  
 'coast', 'of', 'Central', 'America', 'from', 'Belize', 'to', 'Nicaragua']  
 ['Tweants', 'Total', 'speakers', '338', ',', '000', '(', '2003', ')']  
 ['Tweants', 'Official', 'language', 'in', 'Netherlands', '(', 'as', 'part',  
 'of', 'Low', 'Saxon', ')']  
 ['Sikkimese', 'Total', 'speakers', '70', ',', '300', '(', '2001', ')']  
 ['Eastern', 'Pomo', 'Total', 'speakers', '1']  
 ['Mexican', 'Total', 'signers', 'estimated', '87', ',', '000', '\xe2',  
 '\x80', '\x93', '100', ',', '000', '(', '1986', ')']  
 ['Mexican', 'Language', 'codes', 'ISO', '639-3', 'mfs', 'sign', 'language',

'\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Berbice', 'Total', 'speakers', '4', 'or', '5', '(', 'As', 'of', '1993', ')']

['Kenyan', 'Total', 'signers', '340', ',', '000']

['Kenyan', 'Language', 'codes', 'ISO', '639-3', 'xki', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['I\xe2\x80\x99saka', 'Spoken', 'in', 'Sandaun', 'Province', ',', 'Papua', 'New', 'Guinea']

['Mazandarani', 'Total', 'speakers', '3-4', '.', '5', 'million', '(', 'native', 'speakers', ')']

['Mazandarani', 'Region', 'South', 'coast', 'of', 'the', 'Caspian', 'Sea']

['Mazandarani', 'Language', 'codes', 'ISO', '639-3', 'Areas', 'where', 'Mazandarani', 'is', 'spoken', 'as', 'mother', 'tongue']

['Wutung', 'Spoken', 'in', 'Sandaun', 'Province', ',', 'Papua', 'New', 'Guinea']

['R\xc3\xa9union', 'Total', 'speakers', '600', ',', '000']

['Zay', 'Total', 'speakers', '4', ',', '880']

['Shipibo-Conibo', 'Total', 'speakers', '26', ',', '000']

['Huaorani', 'Total', 'speakers', '1', ',', '600-2', ',', '000']

['Huaorani', 'Region', 'Oriente', 'or', 'Ecuadorian', 'Amazon']

['Godoberi', 'Total', 'speakers', '3', ',', '000', '(', '1996', ')']

['Enga', 'Spoken', 'in', 'Enga', 'Province', ',', 'Papua', 'New', 'Guinea']

['Enga', 'Total', 'speakers', '165', ',', '000', '(', '1981', ')']

['Enga', 'Official', 'language', 'in', 'Papua', 'New', 'Guinea']

['Tauya', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Libras', 'Language', 'codes', 'ISO', '639-3', 'bzs', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Pumi', 'language', 'Total', 'speakers', 'more', 'than', '50', ',', '000']

['Pumi', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']

['Pumi', 'Language', 'codes', 'ISO', '639-3', ':', 'pmi', '\xe2', '\x80', '\x93', 'Northern', 'Pumi']

['Pumi', 'Language', 'codes', 'ISO', '639-3', 'pmj', '\xe2', '\x80', '\x93', 'Southern', 'Pumi']

['Northern', 'Total', 'speakers', '130', ',', '000']

['Southern', 'Total', 'speakers', '81', ',', '300']

['Mararit', 'Total', 'speakers', 'Approximately', '43', ',', '000', '(', '1993', ')']

['Mararit', 'Region', 'western', 'sudan', 'and', 'Eastern', 'Chad']

['Shatt', 'Total', 'speakers', '15', ',', '000', '(', '1984', ')']

['Northeastern', 'Maidu', 'Total', 'speakers', '1', 'or', '2']

['Tucano', 'Total', 'speakers', '4', ',', '632']

['Tariana', 'Total', 'speakers', '~', '100']

['Tariana', 'Region', 'Upper', 'and', 'Middle', 'Vaup', '\xc3', '\xa9',



's', 'River', 'in', 'Amazonas']  
 ['Northern', 'Pomo', 'Language', 'extinction', '1994', 'with', 'the',  
 'death', 'of', 'Edna', 'Guerrero']  
 ['Pintupi', 'Spoken', 'in', 'Western', 'Australia', 'Northern',  
 'Territory']  
 ['Fox', 'Language', 'codes', 'ISO', '639-3', ':', 'sac', '\xe2', '\x80',  
 '\x93', 'Fox', 'and', 'Sauk']  
 ['Tooro', 'Total', 'speakers', '500', ',', '000']  
 ['Drehu', 'Spoken', 'in', 'Lifou', ',', 'New', 'Caledonia']  
 ['Drehu', 'Total', 'speakers', '11', ',', '338', '(', '1996', 'census',  
 ')']  
 ['American', 'Language', 'codes', 'ISO', '639-3', 'ase', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Armenian', 'Total', 'speakers', '6', '.', '7', 'million']  
 ['Armenian', 'Language', 'codes', 'ISO', '639-3', ':', 'hye', '\xe2',  
 '\x80', '\x93', 'Modern', 'Armenian']  
 ['Armenian', 'Language', 'codes', 'ISO', '639-3', 'xcl', '\xe2', '\x80',  
 '\x93', 'Classical', 'Armenian']  
 ['Armenian', 'Language', 'codes', 'ISO', '639-3', 'axm', '\xe2', '\x80',  
 '\x93', 'Middle', 'Armenian']  
 ['Armenian', 'Language', 'codes', 'Linguasphere', '57-AAA-a', '(', '31',  
 'varieties', ')']  
 ['Aramaic', 'Total', 'speakers', '~', '500', ',', '000']  
 ['Aramaic', 'Region', 'Throughout', 'the', 'Middle', 'East', ',', 'Europe',  
 'and', 'America']  
 ['Aramaic', 'Writing', 'system', 'Arabic', 'Abjad', '(', 'vernacular', '}',  
 'with', 'a', 'handful', 'of', 'inscriptions', 'found', 'in', 'Demotic',  
 'and', 'Chinese']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', ':', 'arc', '\xe2',  
 '\x80', '\x93', 'Imperial', 'and', 'Official', 'Aramaic', '(', '700-300',  
 'BCE', ')']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'oar', '\xe2', '\x80',  
 '\x93', 'Old', 'Aramaic', '(', 'before', '700', 'BCE', ')']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'aii', '\xe2', '\x80',  
 '\x93', 'Assyrian', 'Neo-Aramaic']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'aij', '\xe2', '\x80',  
 '\x93', 'Lishanid', 'Noshan']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'amw', '\xe2', '\x80',  
 '\x93', 'Western', 'Neo-Aramaic']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'bhn', '\xe2', '\x80',  
 '\x93', 'Bohtan', 'Neo-Aramaic']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'bjf', '\xe2', '\x80',  
 '\x93', 'Barzani', 'Jewish', 'Neo-Aramaic']  
 ['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'cld', '\xe2', '\x80',  
 '\x93', 'Chaldean', 'Neo-Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'jpa', '\xe2', '\x80', '\x93', 'Jewish', 'Palestinian', 'Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'kqd', '\xe2', '\x80', '\x93', 'Koy', 'Sanjaq', 'Surat']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'lsd', '\xe2', '\x80', '\x93', 'Lishana', 'Deni']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'mid', '\xe2', '\x80', '\x93', 'Modern', 'Mandaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'myz', '\xe2', '\x80', '\x93', 'Classical', 'Mandaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'sam', '\xe2', '\x80', '\x93', 'Samaritan', 'Aramaic']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'syc', '\xe2', '\x80', '\x93', 'Syriac', '(', 'classical', ')']

['Aramaic', 'Language', 'codes', 'ISO', '639-3', 'tmr', '\xe2', '\x80', '\x93', 'Jewish', 'Babylonian', 'Aramaic']

['Bulgarian', 'Regulated', 'by', 'Institute', 'for', 'the', 'Bulgarian', 'language', 'at', 'the', 'Bulgarian', 'Academy', 'of', 'Sciences', '(', '\xd0', '\x98', '\xd0', '\xbd', '\xd1', '\x81', '\xd1', '\x82', '\xd0', '\xb8', '\xd1', '\x82', '\xd1', '\x83', '\xd1', '\x82', '\xd0', '\xb7', '\xd0', '\xb0', '\xd0', '\xb1', '\xd1', '\x8a', '\xd0', '\xbb', '\xd0', '\xb3', '\xd0', '\xb0', '\xd1', '\x80', '\xd1', '\x81', '\xd0', '\xba', '\xd0', '\xb8', '\xd0', '\xb5', '\xd0', '\xb7', '\xd0', '\xb8', '\xd0', '\xba', '\xd0', '\xba', '\xd1', '\x8a', '\xd0', '\xbc', '\xd0', '\x91', '\xd1', '\x8a', '\xd0', '\xbb', '\xd0', '\xb3', '\xd0', '\xb0', '\xd1', '\x80', '\xd1', '\x81', '\xd0', '\xba', '\xd0', '\xb0', '\xd1', '\x82', '\xd0', '\xb0', '\xd0', '\xb0', '\xd0', '\xba', '\xd0', '\xb0', '\xd0', '\xb4', '\xd0', '\xb5', '\xd0', '\xbc', '\xd0', '\xb8', '\xd1', '\x8f', '\xd0', '\xbd', '\xd0', '\xb0', '\xd0', '\xbd', '\xd0', '\xb0', '\xd1', '\x83', '\xd0', '\xba', '\xd0', '\xb8', '\xd1', '\x82', '\xd0', '\xb5', '(', '\xd0', '\x91', '\xd0', '\x90', '\xd0', '\x9d', ')', ')']

['Bulgarian', 'Language', 'codes', 'Linguasphere', '53-AAA-hb', '&', 'lt', ';', '53-AAA-h']

['Bambara', 'Total', 'speakers', '4', ',', '000', ',', '000', ',', 'second', 'language', 'speakers', '8-10', 'millions', '(', 'estimation', ',', '2010', ';', 'the', 'figures', 'by', 'the', 'Ethnologue', 'of', '2', ',', '700', ',', '000', 'are', 'out', 'of', 'date', ')']

['Bambara', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]

['Breton', 'Total', 'speakers', 'ca', '.', '200', ',', '000']

['Breton', 'Language', 'codes', 'Linguasphere', '50-ABB-b', '(', 'varieties']

['Breton', 'Language', 'codes', 'Linguasphere', ':', '50-ABB-ba', 'to', '-be', ')']

['Blissymbols', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']

['Khmer', 'Total', 'speakers', '15', '.', '7', 'to', '21', '.', '6',

'million', '(' , '2004', ')']  
 ['Khmer', 'Total', 'speakers', 'Native', 'speakers', ':', '14', '.', '7',  
 'to', '20', '.', '6', 'million']  
 ['Khmer', 'Total', 'speakers', 'Cambodia', ':', '12', '.', '1', 'million']  
 ['Khmer', 'Total', 'speakers', 'Vietnam', ':', '1', ',', '055', ',', '174']  
 ['Khmer', 'Total', 'speakers', 'Thailand', ':', '1', '.', '2', 'million']  
 ['Khmer', 'Total', 'speakers', 'USA', ':', '190', ',', '000']  
 ['Khmer', 'Total', 'speakers', 'France', ':', 'ca', '.', '50', ',', '000']  
 ['Khmer', 'Total', 'speakers', 'Australia', ':', '22', ',', '000']  
 ['Khmer', 'Total', 'speakers', 'Canada', ':', '16', ',', '500', '2nd',  
 'language', 'speakers', ':', '1', 'million', 'in', 'Cambodia']  
 ['Khmer', 'Language', 'codes', 'ISO', '639-3', ':', 'khm', '\xe2', '\x80',  
 '\x93', 'Central', 'Khmer']  
 ['Khmer', 'Language', 'codes', 'ISO', '639-3', 'kxm', '\xe2', '\x80',  
 '\x93', 'Northern', 'Khmer']  
 ['Khmer', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Cornish', 'Regulated', 'by', 'Cornish', 'Language', 'Partnership']  
 ['Czech', 'Regulated', 'by', 'Czech', 'Language', 'Institute']  
 ['Czech', 'Language', 'codes', 'Linguasphere', '53-AAA-da', '&', 'lt', ';',  
 '53-AAA-b', '.', '.', '.', '-d']  
 ['Czech', 'Language', 'codes', 'Linguasphere', '(' , 'varieties', ':', '53-  
 AAA-daa', 'to', '53-AAA-dam', ')']  
 ['Danish', 'Regulated', 'by', 'Dansk', 'Sprogn', '\xc3', '\xa6', '\vn', '(' ,  
 '\xe2', '\x80', '\x9c', 'Danish', 'Language', 'Committee', '\xe2', '\x80',  
 '\x9d', ')']  
 ['Danish', 'Language', 'codes', 'Linguasphere', '52-AAA-bf', '&', 'amp',  
 ';', '-ca', 'to', '-cj']  
 ['Dioula', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Etruscan', 'Language', 'extinction', '1st', 'century', 'AD']  
 ['Estonian', 'Total', 'speakers', '1', ',', '048', ',', '660']  
 ['Estonian', 'Spoken', 'in', 'Estonian', 'immigrant', 'communities', '(' ,  
 'see', 'Estonians', ')']  
 ['Estonian', 'Regulated', 'by', 'Institute', 'of', 'the', 'Estonian',  
 'Language', '/', 'Eesti', 'Keele', 'Instituut', ',', 'Emakeele', 'Selts',  
 '(' , 'semi-official', ')']  
 ['Estonian', 'Language', 'codes', 'ISO', '639-3', ':', 'ekk', '\xe2',  
 '\x80', '\x93', 'Standard', 'Estonian']  
 ['West', 'Flemish', 'Total', 'speakers', '~', '1', '.', '16', 'million',  
 'speakers']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', ':', 'grc', '\xe2', '\x80',  
 '\x93', 'Ancient', 'Greek']  
 ['Greek', 'Language', 'codes', 'ISO', '639-3', 'ell', '\xe2', '\x80',  
 '\x93', 'Modern', 'Greek']

['Greek', 'Language', 'codes', 'ISO', '639-3', 'pnt', '\xe2', '\x80', '\x93', 'Pontic', 'Greek']

['Greek', 'Language', 'codes', 'ISO', '639-3', 'gmy', '\xe2', '\x80', '\x93', 'Mycenaean', 'Greek']

['Greek', 'Language', 'codes', 'ISO', '639-3', 'gkm', '\xe2', '\x80', '\x93', 'Medieval', 'Greek']

['Greek', 'Language', 'codes', 'ISO', '639-3', 'cpg', '\xe2', '\x80', '\x93', 'Cappadocian', 'Greek']

['Greek', 'Language', 'codes', 'ISO', '639-3', 'tsd', '\xe2', '\x80', '\x93', 'Tsakonian', 'Greek']

['Greek', 'Language', 'codes', 'Linguasphere', '56-AAA-a', '(', 'varieties']

['Greek', 'Language', 'codes', 'Linguasphere', ':', '56-AAA-aa', 'to', '-am', ')']

['Gallurese', 'Total', 'speakers', '100', ',', '000', '(', '1993', ')']

['Hebrew', 'Total', 'speakers', 'Total', 'Speakers', '&', 'lt', ';', '9', ',', '000', ',', '000']

['Hebrew', 'Total', 'speakers', 'First', 'Language', '5', ',', '300', ',', '000', '(', '2009', ')', ';']

['Hebrew', 'Total', 'speakers', 'Second', 'Language', '2', ',', '500', ',', '000', '(', '2009', ')']

['Hebrew', 'Total', 'speakers', 'Home', 'Language', '200', ',', '000', '(', 'approx', '.', ')', 'in', 'the', 'United', 'States', 'speak', 'Hebrew', 'at', 'home1']

['Hebrew', 'Total', 'speakers', '1United', 'States', 'Census', '2000', 'PHC-T-37', '.', 'Ability', 'to', 'Speak', 'English', 'by', 'Language', 'Spoken', 'at', 'Home', ':', '2000', '.', 'Table', '1a', '.', 'PDF', '(', '11', '.', '8', 'KB', ')']

['Hebrew', 'Total', 'speakers', 'Palestinian', 'territories', 'Second', 'Language', '500', ',', '000', '-', '1', ',', '000', ',', '000']

['Hebrew', 'Total', 'speakers', 'Extinct', 'as', 'a', 'regularly', 'spoken', 'language', 'by', 'the', '4th', 'century', 'CE', ',', 'but', 'survived', 'as', 'a', 'liturgical', 'and', 'literary', 'language', ';', 'revived', 'in', 'the', '1880s']

['Hebrew', 'Spoken', 'in', 'Global', '(', 'as', 'a', 'liturgical', 'language', 'for', 'Judaism', ')', ',', 'in', 'West', 'Bank', ',', 'and', 'Gaza']

['Hebrew', 'Pronunciation', 'standard', 'Israeli', ':', '(', '\xca', '\x94', ')', 'iv', '\xcb', '\x88', '\xca', '\x81', 'it', '-', '(', '\xca', '\x94', ')', 'iv', '\xcb', '\x88', '\xc9', '\xbe', 'it', ',']

['Hebrew', 'Pronunciation', 'standard', 'Israeli', '(', 'Sephardi', ')', ':', '\xca', '\x95', 'iv', '\xcb', '\x88', '\xc9', '\xbe', 'it', ',']

['Hebrew', 'Pronunciation', 'Iraqi', ':', '\xca', '\x95', 'ib', '\xcb', '\x88', 'ri', '\xcb', '\x90', '\xce', '\xb8', ',']

['Hebrew', 'Regulated', 'by', 'Academy', 'of', 'the', 'Hebrew', 'Language']

['Hebrew', 'Regulated', 'by', '\xd7', '\x94', '\xd7', '\x90', '\xd7',

'\xa7', '\xd7', '\x93', '\xd7', '\x9e', '\xd7', '\x99', '\xd7', '\x94',  
 '\xd7', '\x9c', '\xd7', '\x9c', '\xd7', '\xa9', '\xd7', '\x95', '\xd7',  
 '\x9f', '\xd7', '\x94', '\xd7', '\xa2', '\xd7', '\x91', '\xd7', '\xa8',  
 '\xd7', '\x99', '\xd7', '\xaa', '(', 'HaAkademia', 'LaLashon', 'Ha',  
 '\xe2', '\x80', '\x98', 'Ivrit', ')']  
 ['Hebrew', 'Language', 'codes', 'ISO', '639-3', ':', 'heb', '\xe2', '\x80',  
 '\x93', 'Modern', 'Hebrew']  
 ['Hebrew', 'Language', 'codes', 'ISO', '639-3', 'hbo', '\xe2', '\x80',  
 '\x93', 'Ancient', 'Hebrew']  
 ['Hindi', 'Total', 'speakers', '550', 'million', 'in', '2011']  
 ['Hindi', 'Regulated', 'by', 'Central', 'Hindi', 'Directorate', '(',  
 'India', ')']  
 ['Hindi', 'Language', 'codes', 'Linguasphere', '59-AAF-q', '(', 'with',  
 'Urdu', ',']  
 ['Hindi', 'Language', 'codes', 'Linguasphere', 'including', '58',  
 'varieties', ':', '59-AAF-qaa', 'to', '59-AAF-qil', ')']  
 ['Hindi', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Hawaiian', 'Spoken', 'in', 'Hawai', '\xca', '\xbb', 'i', ':',  
 'concentrated', 'on', 'Ni', '\xca', '\xbb', 'ihau', 'and', 'Hawai', '\xca',  
 '\xbb', 'i', ', ', 'but', 'speakers', 'throughout', 'the', 'Hawaiian',  
 'Islands', 'and', 'the', 'U', '.', 'S', '.', 'mainland']  
 ['Hawaiian', 'Total', 'speakers', '~', '2000', 'native']  
 ['Hawaiian', 'Total', 'speakers', '~', '27', ', ', '160']  
 ['Hawaiian', 'Writing', 'system', 'Latin', 'plus', 'Hawaiian', 'symbols']  
 ['Hawaiian', 'Official', 'language', 'in', 'Hawai', '\xca', '\xbb', 'i',  
 '(', 'with', 'English', ')']  
 ['Hawaiian', 'Official', 'language', 'in', ':', 'U', '.', 'S', '.',  
 'Mainland']  
 ['Italian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Italian',  
 'variant', ')']  
 ['Ido', 'Total', 'speakers', '100', '\xe2', '\x80', '\x93', '200']  
 ['Ido', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Ido', 'Created', 'by', 'A', 'group', 'of', 'reformist', 'Esperanto',  
 'speakers']  
 ['Ido', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Interlingua', 'Created', 'by', 'International', 'Auxiliary', 'Language',  
 'Association']  
 ['Japanese', 'Writing', 'system', 'Japanese', 'logographs', 'and',  
 'syllabaries']  
 ['Japanese', 'Regulated', 'by', 'Japanese', 'government', 'plays', 'major',  
 'role']

['Korean', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Korean', 'Total', 'speakers', '78', 'million', '(', '1999', 'est', '.', '(',')']  
 ['Korean', 'Official', 'language', 'in', 'Yanbian', ',', 'People's', 'Republic', 'of', 'China']  
 ['Korean', 'Regulated', 'by', ':', 'The', 'National', 'Institute', 'of', 'the', 'Korean', 'Language']  
 ['Korean', 'Regulated', 'by', 'The', 'Language', 'Research', 'Institute', 'of', 'Social', 'Science']  
 ['Klingon', 'Regulated', 'by', 'Klingon', 'Language', 'Institute']  
 ['Klingon', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Klingon', 'writing', 'systems', 'constructed', 'languages']  
 ['Klingon', 'Setting', 'and', 'usage', 'Star', 'Trek', 'films', 'and', 'television', 'series', '(', 'TNG', ',', 'DS9', ',', 'Voyager', ',', 'Enterprise', ')', 'and', 'the', 'opera', '"u"']  
 ['Kashubian', 'Total', 'speakers', '50', ',', '000']  
 ['Kashubian', 'Regulated', 'by', 'Kashubian', 'Language', 'Council']  
 ['Loglan', 'Category', '(', 'sources', ')', 'eight', 'of', 'the', 'most', 'common', 'languages', ':', 'English', ',', 'Chinese', '(', 'Beijing', 'dialect', ')', ',', 'Hindi', ',', 'Russian', ',', 'Spanish', ',', 'French', ',', 'Japanese', ',', 'German']  
 ['Maltese', 'Total', 'speakers', '371', ',', '900', '(', '1975', ')']  
 ['Maltese', 'Regulated', 'by', 'National', 'Council', 'for', 'the', 'Maltese', 'Language']  
 ['Mandarin', 'Total', 'speakers', 'Overall', ':', '1', ',', '365', ',', '053', ',', '177']  
 ['Mandarin', 'Dialects', '(', 'Jin', ',', 'disputed', ')']  
 ['Mandarin', 'Region', 'Most', 'of', 'northern', 'and', 'southwestern', 'China', ',', 'Taiwan', 'and', 'Singapore', '(', 'and', 'other', 'overseas', 'Chinese', 'communities', ')']  
 ['Mandarin', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Mandarin', 'Spoken', 'in', 'United', 'States', '(', 'chiefly', 'New', 'York', 'City', ')']  
 ['Manx', 'Total', 'speakers', 'and', '1', ',', '689', 'people', '(', '2', '.', '2', '%', 'total', 'population', ')', 'in', 'the', 'Isle', 'of', 'Man', 'professing', 'some', 'knowledge', 'of', 'the', 'language', '(', '2001', ')']  
 ['Manx', 'Regulated', 'by', 'Coonseil', 'ny', 'Gaelgey', '(', 'Manx', 'Gaelic', 'Council', ')']  
 ['Marathi', 'Region', 'Dadra', '&', 'amp', ';', 'Nagar', 'Haveli']  
 ['Marathi', 'Region', 'and', 'Daman', '&', 'amp', ';', 'Diu']  
 ['Marathi', 'Spoken', 'in', 'Australia', '&', 'amp', ';', 'New', 'Zealand']  
 ['Marathi', 'Regulated', 'by', 'Maharashtra', 'Sahitya', 'Parishad', '&', 'amp', ';', 'various', 'other', 'institutions']  
 ['Marathi', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Marathi', 'speakers', 'in', 'India', 'This', 'page', 'contains',

'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']

['Norwegian', 'Spoken', 'in', 'Norway', '(', '4', '.', '8', 'million', ')',  
 ',']

['Norwegian', 'Spoken', 'in', 'United', 'States', '(', '55', ',', '311',  
 ')']

['Norwegian', 'Spoken', 'in', 'Denmark', '(', '15', ',', '000', ')']

['Norwegian', 'Spoken', 'in', 'Canada', '(', '7', ',', '710', ')']

['Norwegian', 'Standard', 'forms', 'Nynorsk', '(', 'official', ')']

['Norwegian', 'Standard', 'forms', 'Bokm', '\xc3', '\xa5', 'l', '(',  
 'official', ')', '/', 'Riksm', '\xc3', '\xa5', 'l', '(', 'unofficial', ')']

['Norwegian', 'Regulated', 'by', 'Norwegian', 'Language', 'Council', '(',  
 'Bokm', '\xc3', '\xa5', 'l', 'and', 'Nynorsk', ')']

['Norwegian', 'Regulated', 'by', 'Norwegian', 'Academy', '(', 'Riksm',  
 '\xc3', '\xa5', 'l', ')']

['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-ba', 'to', '-  
 be', '&', 'amp', ';']

['Norwegian', 'Language', 'codes', 'Linguasphere', '52-AAA-cf', 'to', '-  
 cg']

['Northern', 'Sotho', 'Total', 'speakers', '4', ',', '208', ',', '940']

['Northern', 'Regulated', 'by', 'Pan', 'South', 'African', 'Language',  
 'Board']

['Novial', 'Category', '(', 'sources', ')', 'Romance', 'and', 'Germanic',  
 'languages', ';', 'also', 'Occidental', 'and', 'Ido']

['Norn', 'Language', 'extinction', 'by', 'the', '18th', 'century', '(',  
 '19th', 'century', 'at', 'the', 'latest', ')', ';', 'much', 'earlier',  
 'in', 'Caithness']

['Prussian', 'Language', 'extinction', 'Late', '17th', '/', 'Early',  
 '18th', 'century']

['Old', 'Norse', 'Language', 'extinction', 'developed', 'into', 'the',  
 'various', 'North', 'Germanic', 'languages', 'by', 'the', '14th',  
 'century']

['Old', 'Writing', 'system', 'Runic', ',', 'later', 'Latin', 'alphabet',  
 '(', 'Old', 'Norse', 'variant', ')']

['Old', 'English', 'Language', 'extinction', 'mostly', 'developed', 'into',  
 'Middle', 'English', 'by', 'the', '13th', 'century']

['Old', 'Writing', 'system', 'Runic', ',', 'later', 'Latin', 'alphabet',  
 '(', 'Old', 'English', 'variant', ')']

['Occidental', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']

['Occidental', 'Created', 'by', 'Edgar', 'de', 'Wahl', '(', '1922', ')']

['Polish', 'Official', 'language', 'in', ':', 'Czech', 'Republic']

['Polish', 'Regulated', 'by', 'Polish', 'Language', 'Council']

['Polish', 'Language', 'codes', 'Linguasphere', '53-AAA-cc', '&', 'lt',  
 ';', '53-AAA-b', '.', '.', '.', '-d']

['Polish', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-AAA-cca', 'to', '53-AAA-ccu', ')']

['Portuguese', 'Total', 'speakers', 'Native', ':', '236', '.', '1', 'million']

['Portuguese', 'Total', 'speakers', 'Total', ':', '272', '.', '9', 'million']

['Portuguese', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Portuguese', 'variant', ')']

['Portuguese', 'Pronunciation', 'po', '\xc9', '\xbe', 'tu', '\xcb', '\x88', '\xc9', '\xa1', 'es', '(', 'BP-ga', '\xc3', '\xba', 'cho', ')', '1']

['Portuguese', 'Regulated', 'by', 'International', 'Portuguese', 'Language', 'Institute', ';', 'CPLP', ';', 'Academia', 'Brasileira', 'de', 'Letras', '(', 'Brazil', ')', ';', 'Academia', 'das', 'Ci', '\xc3', '\xaa', 'ncias', 'de', 'Lisboa', ',', 'Classe', 'de', 'Letras', '(', 'Portugal', ')']

['Standard', 'Spoken', 'in', "People's", 'Republic', 'of', 'China']

['Standard', 'Official', 'language', 'in', "People's", 'Republic', 'of', 'China']

['Standard', 'Official', 'language', 'in', 'Singapore', '(', 'official', 'but', 'not', 'national', 'language', ')']

['Standard', 'Regulated', 'by', 'In', 'the', 'PRC', ':', 'National', 'Language', 'Regulating', 'Committee']

['Standard', 'Regulated', 'by', 'In', 'Singapore', ':', 'Promote', 'Mandarin', 'Council', '/', 'Speak', 'Mandarin', 'Campaign']

['Punjabi', 'Language', 'Total', 'speakers', '88', ',', '000', ',', '000', '(', 'Ethnologue', '2005', 'estimate', ')']

['Punjabi', 'Writing', 'system', 'Devanagari', 'in', 'Haryana', ',', 'Himachal', 'Pradesh', ',', 'Delhi', 'and', 'Punjabi', 'Hindu', 'diaspora', ')']

['Punjabi', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Punjabi', 'speakers', 'in', 'India', 'and', 'Pakistan', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Romansh', 'Total', 'speakers', '35', ',', '095', '(', 'Swiss', 'federal', 'census', '2000', ')']

['Romanian', 'Region', 'Southeastern', ',', 'Central', 'and', 'Eastern', 'Europe']

['Romanian', 'Spoken', 'in', ':', 'North', 'and', 'South', 'America']

['Romanian', 'Spoken', 'in', 'Western', 'and', 'Southern', 'Europe']

['Romanian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Romanian', 'variant', ')']

['Romanian', 'Language', 'codes', 'Linguasphere', '51-AAD-c', '(', 'varieties']

['Romanian', 'Language', 'codes', 'Linguasphere', ':', '51-AAD-ca', 'to', '-ck', ')', 'Map', 'of', 'the', 'Romanian-speaking', 'territories']



['Swedish', 'Total', 'speakers', '~', '10', 'million']  
 ['Swedish', 'Region', 'Northern', 'Europe', ',', 'parts', 'of', 'USA',  
 'and', 'other', 'countries']  
 ['Swedish', 'Spoken', 'in', 'Sweden', '(', '9', '.', '4', 'million', ')']  
 ['Swedish', 'Spoken', 'in', 'Finland', '(', '290', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'USA', '(', '70', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'Spain', '(', '40', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'Great', 'Britain', '(', '30', ',', '000', ')']  
 ['Swedish', 'Spoken', 'in', 'Canada', '(', '20', ',', '000', ')']  
 ['Swedish', 'Regulated', 'by', 'Swedish', 'Language', 'Council', '(', 'in',  
 'Sweden', ')']  
 ['Swedish', 'Language', 'codes', 'Linguasphere', '52-AAA-ck', 'to', '-cw']  
 ['Spanish', 'Official', 'language', 'in', 'Organization', 'of',  
 'American', 'States']  
 ['Spanish', 'Official', 'language', 'in', 'Union', 'of', 'South',  
 'American', 'Nations']  
 ['Spanish', 'Official', 'language', 'in', 'Central', 'American',  
 'Integration', 'System']  
 ['Spanish', 'Official', 'language', 'in', 'North', 'American', 'Free',  
 'Trade', 'Agreement']  
 ['Spanish', 'Regulated', 'by', 'Association', 'of', 'Spanish', 'Language',  
 'Academies', '(', 'Real', 'Academia', 'Espa', '\xc3', '\xb1', 'ola', 'and',  
 '21', 'other', 'national', 'Spanish', 'language', 'academies', ')']  
 ['Logudorese', 'Total', 'speakers', '500', ',', '000-700', ',', '000']  
 ['Sardinian', 'Total', 'speakers', '1', '.', '85', 'million']  
 ['Sardinian', 'Regulated', 'by', 'Limba', 'Sarda', 'Comuna', '(', 'Common',  
 'Sardinian', 'Language', ')', 'code']  
 ['Sardinian', 'Language', 'codes', 'Linguasphere', '51-AAA-s', '+', '(',  
 'Corso-Sardinian', ')']  
 ['Sardinian', 'Language', 'codes', 'Linguasphere', '51-AAA-pd', '&', 'amp',  
 ';', '-pe']  
 ['Swiss', 'German', 'Total', 'speakers', '4', ',', '500', ',', '000']  
 ['Swiss', 'Language', 'codes', 'Linguasphere', '52-ACB-f', '(', '45',  
 'varieties', ':', '52-ACB-faa', 'to', '-fkb', ')']  
 ['Sanskrit', 'Total', 'speakers', '14', ',', '135', 'native', 'speakers',  
 'in', 'India', '(', '2001', ')']  
 ['Sanskrit', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Serbo-Croatian', 'Total', 'speakers', '16', '.', '3', 'million']  
 ['Serbo-Croatian', 'Region', 'the', 'Western', 'Balkans', 'Ethnicity']  
 ['Serbo-Croatian', 'Standard', 'forms', 'Standard', 'Serbo-Croatian', '(',  
 'defunct', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Bosnia', 'and',  
 'Herzegovina', '(', 'as', 'Bosnian', ',', 'Croatian', ',', 'Serbian', ')']

['Serbo-Croatian', 'Official', 'language', 'in', 'Croatia', '(', 'as',  
 'Croatian', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Kosovo', '(', 'as',  
 'Serbian', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Montenegro', '(', 'as',  
 'Montenegrin', ')']  
 ['Serbo-Croatian', 'Official', 'language', 'in', 'Serbia', '(', 'as',  
 'Serbian', ')']  
 ['Serbo-Croatian', 'Regulated', 'by', 'Council', 'for', 'Standard',  
 'Croatian', 'Language', 'Norm']  
 ['Serbo-Croatian', 'Regulated', 'by', 'Board', 'for', 'Standardization',  
 'of', 'the', 'Serbian', 'Language']  
 ['Slovene/Slovenian', 'Total', 'speakers', '2', '.', '5', 'million']  
 ['Slovene/Slovenian', 'Region', 'Central', 'Southern', 'and',  
 'Southeastern', 'Europe']  
 ['Slovene/Slovenian', 'Writing', 'system', 'Latin', 'alphabet', '(',  
 'Slovene', 'variant', ')']  
 ['Slovene/Slovenian', 'Official', 'language', 'in', 'Regional', 'or',  
 'local', 'official', 'language', 'in']  
 ['Slovene/Slovenian', 'Language', 'codes', 'Linguasphere', '53-AAA-f', '(',  
 '51', 'varieties', ')']  
 ['Slovak', 'Spoken', 'in', 'Czech', 'Republic', ',']  
 ['Slovak', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Slovak',  
 'variant', ')']  
 ['Slovak', 'Regulated', 'by', 'Ministry', 'of', 'Culture', 'of', 'the',  
 'Slovak', 'Republic']  
 ['Slovak', 'Language', 'codes', 'Linguasphere', '53-AAA-db', '&', 'lt',  
 ';', '53-AAA-b', '.', '.', '.', '-d']  
 ['Slovak', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-  
 AAA-dba', 'to', '53-AAA-dbs', ')']  
 ['Scottish', 'Total', 'speakers', '92', ',', '400', 'people', 'aged',  
 'three', 'and', 'over', 'in', 'Scotland', 'had', 'some', 'Gaelic',  
 'language', 'ability', 'in', '2001', 'with', 'an', 'additional', '2', ',',  
 '000', 'in', 'Nova', 'Scotia', '.', '1', ',', '610', 'speakers', 'in',  
 'the', 'United', 'States', 'in', '2000', '.', '822', 'in', 'Australia',  
 'in', '2001', '.', '669', 'in', 'New', 'Zealand', 'in', '2006']  
 ['Swahili', 'Spoken', 'in', 'Comoros', '(', 'as', 'Comorian', ')']  
 ['Swahili', 'Spoken', 'in', 'Seychelles', '(', 'as', 'Shimaore', ')']  
 ['Swahili', 'Total', 'speakers', 'First', 'and', 'second', 'language', ':',  
 '50', '+', 'million']  
 ['Tamil', 'Total', 'speakers', '65', ',', '675', ',', '200']  
 ['Tamil', 'Official', 'language', 'in', 'India', '(', 'Tamil', 'Nadu', ',',  
 'Puducherry', ')', ',']  
 ['Tamil', 'Official', 'language', 'in', 'Sri', 'Lanka', ',', 'and']  
 ['Tamil', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'native', 'Tamil', 'speakers', 'in', 'India', 'and', 'Sri', 'Lanka',

'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Setswana', 'Total', 'speakers', '4', ',', '521', ',', '700']

['Tocharian', 'languages', 'Language', 'extinction', '9th', 'century', 'AD']

['Tocharian', 'Spoken', 'in', 'Tarim', 'Basin', 'in', 'Central', 'Asia']

['Tocharian', 'Language', 'codes', 'ISO', '639-3', ':', 'xto', '\xe2', '\x80', '\x93', 'Tocharian', 'A']

['Tocharian', 'Language', 'codes', 'ISO', '639-3', 'txb', '\xe2', '\x80', '\x93', 'Tocharian', 'B']

['Vietnamese', 'Total', 'speakers', '70-73', 'million', 'native', '(', 'includes', '3', 'million', 'overseas', ')']

['Vietnamese', 'Total', 'speakers', '80', '+', 'million', 'total']

['Vietnamese', 'Pronunciation', 'ti', '\xc9', '\x99', '\xcc', '\x8c', '\xc5', '\x8b', 'vi', '\xc9', '\x99', '\xcc', '\x80', '\xcb', '\x80', 't', '(', 'Northern', ')']

['Vietnamese', 'Pronunciation', 'ti', '\xc9', '\x99', '\xcc', '\x8c', '\xc5', '\x8b', 'ji', '\xc9', '\x99', '\xcc', '\x80', 'k', '(', 'Southern', ')']

['Volap\xc3\xbck', 'Date', 'founded', '1879', '\xe2', '\x80', '\x93', '1880']

['Volap\xc3\xbck', 'Writing', 'system', 'Latin', 'vocabulary', 'from', 'English']

['Volap\xc3\xbck', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']

['Volap\xc3\xbck', 'Setting', 'and', 'usage', 'International', ':', 'mostly', 'in', 'Europe']

['Yiddish', 'Total', 'speakers', '1', ',', '762', ',', '320']

['Yiddish', 'Official', 'language', 'in', 'Official', 'minority', 'language', 'in']

['Yiddish', 'Official', 'language', 'in', 'Jewish', 'Autonomous', 'Oblast']

['Yiddish', 'Language', 'codes', 'ISO', '639-3', ':', 'ydd', '\xe2', '\x80', '\x93', 'Eastern', 'Yiddish']

['Yiddish', 'Language', 'codes', 'ISO', '639-3', 'yih', '\xe2', '\x80', '\x93', 'Western', 'Yiddish']

['Yiddish', 'Language', 'codes', 'Linguasphere', '52-ACB-g', '=', '52-ACB-ga', '(', 'West', ')', '+', '52-ACB-gb', '(', 'East', ')', ';', 'totalling', '11', 'varieties']

['Indonesian', 'Spoken', 'in', 'East', 'Timor', '(', 'as', 'a', '"', 'working', 'language', '"', ')']

['Malayalam', 'Total', 'speakers', '35', ',', '893', ',', '990']

['Malayalam', 'Total', 'speakers', '33', ',', '015', ',', '420', 'in', 'India', '(', '2001', ')', ',', '']

['Malayalam', 'Total', 'speakers', '1', ',', '847', ',', '902', 'in', 'other', 'countries', '(', '2007', ')']

['Malayalam', 'Total', 'speakers', '\xe2', '\x80', '\xa2', '447', ',', '440', 'in', 'Saudi', 'Arabia']

['Malayalam', 'Language', 'codes', 'ISO', '639-3', 'Number', 'of', 'Malayalam', 'speakers', 'by', 'Countries', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Language', 'Total', 'speakers', '38', 'million', 'native', '(', '2001', ',', 'only', 'India', ')', ',', '13', 'million', 'as', 'a', 'second', 'language', '(', '2001', ')', '=', '51', 'million']

['Language', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Kannada', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Telugu', 'Total', 'speakers', '130', 'million', 'native', 'speakers', 'as', 'of', '2011']

['Telugu', 'Region', 'Andhra', 'Pradesh', ',', 'Tamil', 'Nadu']

['Telugu', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Telugu', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Dari', '(', 'Afghan', 'Persian', ')', 'Total', 'speakers', '12-15', 'million', ',', 'as', 'first', 'language', '(', 'Tajiks', '11-13', 'm', ';', 'Hazaras', '2', '.', '5-3', 'm', ';', 'Aym', '\xc4', '\x81', 'qs', '3-3', '.', '5', 'm', ';', 'Pashtuns', '(', 'unknown', ')', ')]

['Dari', 'Language', 'codes', 'ISO', '639-3', ':', 'prs', '\xe2', '\x80', '\x93', 'Dari', ',', 'Afghan', 'Persian']

['Dari', 'Language', 'codes', 'Linguasphere', '58-AAC-ce', '(', 'Dari', ')', '+', '58-AAC-cdo', '&', 'amp', ';', 'cdp', '(', 'Hazaragi', ')', '+', '58-AAC-ck', '(', 'Aimaq', ')]

['Kurdish', 'Writing', 'system', 'Kurdish', 'alphabet', '(', 'modified', 'Perso-Arabic', 'alphabet', 'in', 'Iraq', 'and', 'Iran', ',', 'Latin', 'alphabet', 'in', 'Turkey', ',', 'Syria', 'and', 'Armenia', ',', 'modified', 'Cyrillic', 'in', 'the', 'former', 'USSR', ')]

['Kurdish', 'Official', 'language', 'in', 'Iraq', ':', 'status', 'as', 'official', 'language', 'alongside', 'Arabic']

['Kurdish', 'Language', 'codes', 'ISO', '639-3', 'sdh', '\xe2', '\x80', '\x93', 'Southern', 'Kurdish']

['Kurdish', 'Language', 'codes', 'Linguasphere', '58-AAA-a', '(', 'North', 'Kurdish', 'incl', '.', 'Kurmanji', '&', 'amp', ';', 'Kurmanjiki', ')', '+', '58-AAA-b', '(', 'Central', 'Kurdish', 'incl', '.', 'Dimli', '/', 'Zaza', '&', 'amp', ';', 'Gurani', ')', '+', '58-AAA-c', '(', 'South', 'Kurdish', 'incl', '.', 'Kurdi', ')]

['Sudovian', 'Language', 'extinction', '17th', 'c', '.', '?']

['Old', 'Church', 'Slavonic', 'Spoken', 'in', 'formerly', 'in', 'Slavic', 'areas', ',', 'under', 'the', 'influence', 'of', 'Byzantium', '(', 'both', 'Catholic', 'and', 'Orthodox', ')']

['Old', 'Language', 'extinction', 'in', 'use', 'only', 'as', 'a', 'liturgical', 'language']

['Latvian', 'Total', 'speakers', 'native', '1', '.', '39', 'million', '(', 'Latvia', ')']

['Latvian', 'Total', 'speakers', '1', '.', '5', 'million', '(', 'Worldwide', ')']

['Latvian', 'Total', 'speakers', 'second', 'language', ':', '400', ',', '000']

['Latvian', 'Regulated', 'by', 'Latvian', 'State', 'Language', 'Center']

['Latvian', 'Language', 'codes', 'ISO', '639-3', 'lvs', '\xe2', '\x80', '\x93', 'Standard', 'Latvian']

['Mon', 'Total', 'speakers', 'Myanmar', ':', '742', ',', '900', ',', 'Total', ':', '850', ',', '530']

['Mon', 'Spoken', 'in', 'Burma', '(', 'Myanmar', ')', ',', 'Thailand']

['Mon', 'Writing', 'system', 'Burmese', 'alphabet', '(', 'itself', 'derived', 'from', 'the', 'Old', 'Mon', 'Indic-based', 'script', ')']

['Michif', 'Region', 'Turtle', 'Mountain', 'Indian', 'Reservation', 'in', 'North', 'Dakota']

['British', 'Language', 'codes', 'ISO', '639-3', 'bfi', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Thai', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Duala', 'Total', 'speakers', '90', ',', '000', '(', '1982', ')']

['Akkadian', 'Language', 'extinction', '100', 'AD']

['Akkadian', 'Official', 'language', 'in', 'initially', 'Akkad', '(', 'central', 'Mesopotamia', ')', ';', 'lingua', 'franca', 'of', 'the', 'Middle', 'East', 'and', 'Egypt', 'in', 'the', 'late', 'Bronze', 'and', 'early', 'Iron', 'Ages']

['Middle', 'English', 'Language', 'extinction', 'developed', 'into', 'Early', 'Modern', 'English', ',', 'Scots', 'and', 'Yola', 'in', 'Wexford', 'by', 'the', '16th', 'century']

['Middle', 'Spoken', 'in', 'south', 'east', 'Scotland', 'and', 'in', 'Scottish', 'burghs']

['M\u0481ori', 'Total', 'speakers', '157', ',', '110', 'New', 'Zealand', 'residents', 'claimed', 'they', 'could', 'converse', 'in', 'M', '\u0481', 'ori', 'about', 'everyday', 'things', ',', 'in', 'the', '2006', 'Census']

['M\u0481ori', 'Regulated', 'by', 'M', '\u0481', 'ori', 'Language', 'Commission']

['Cree', 'Total', 'speakers', '117', ',', '410', '(', 'including',

'Montagnais-Naskapi', 'and', 'Atikamekw', ')']  
 ['Cree', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Cree',  
 'syllabics', '(', 'variation', 'of', 'Canadian', 'Aboriginal', 'syllabics',  
 ')']  
 ['Cree', 'Language', 'codes', 'ISO', '639-3', 'crm', '\xe2', '\x80',  
 '\x93', 'Moose', 'Cree']  
 ['Cree', 'Language', 'codes', 'ISO', '639-3', 'crl', '\xe2', '\x80',  
 '\x93', 'Northern', 'East', 'Cree']  
 ['Cree', 'Language', 'codes', 'ISO', '639-3', 'crj', '\xe2', '\x80',  
 '\x93', 'Southern', 'East', 'Cree']  
 ['Cree', 'Language', 'codes', 'ISO', '639-3', 'crw', '\xe2', '\x80',  
 '\x93', 'Swampy', 'Cree']  
 ['Cree', 'Language', 'codes', 'ISO', '639-3', 'cwd', '\xe2', '\x80',  
 '\x93', 'Woods', 'Cree']  
 ['Cree', 'Language', 'codes', 'ISO', '639-3', 'crk', '\xe2', '\x80',  
 '\x93', 'Plains', 'Cree']  
 ['Tetum', 'Total', 'speakers', '800', ',', '000']  
 ['Hakka', 'Region', 'in', 'China', ':', 'Eastern', 'Guangdong', 'province',  
 ';', 'adjoining', 'regions', 'of', 'Fujian', 'and', 'Jiangxi', 'provinces']  
 ['Hakka', 'Spoken', 'in', 'Japan', '(', 'due', 'to', 'presence', 'of',  
 'Taiwanese', 'community', 'in', 'Tokyo-Yokohama', 'Metropolitan', 'Area',  
 ')']  
 ['Hakka', 'Spoken', 'in', 'India', 'and', 'other', 'countries', 'where',  
 'Hakka', 'Chinese', 'migrants', 'have', 'settled']  
 ['Hakka', 'Official', 'language', 'in', 'none', '(', 'legislative',  
 'bills', 'have', 'been', 'proposed', 'for', 'it', 'to', 'be', 'one', 'of',  
 'the', '"national", "languages"', 'in', 'the', 'Republic', 'of', 'China',  
 ')', ';', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public',  
 'transport', 'announcements', 'in', 'the', 'ROC', ';', 'ROC', 'government',  
 'sponsors', 'Hakka', 'language', 'television', 'station', 'to', 'preserve',  
 'language']  
 ['Hakka', 'Regulated', 'by', 'The', 'Guangdong', 'Provincial', 'Education',  
 'Department', 'created', 'an', 'official', 'romanisation', 'of', 'Meixian',  
 'Hakka', 'dialect', 'in', '1960', ',', 'one', 'of', 'four', 'languages',  
 'receiving', 'this', 'status', 'in', 'Guangdong', '.', 'It', 'is',  
 'called', 'Kejiahua', 'Pinyin', "Fang'an"]  
 ['Hakka', 'Language', 'codes', 'ISO', '639-1', 'zh', '(', 'Chinese', ')']  
 ['Nyabwa', 'Total', 'speakers', '43', ',', '000']  
 ['Quebec', 'Total', 'speakers', '(', 'mother', 'tongue', ')', '6',  
 'million', 'in', 'Quebec', ',', '700', ',', '000', 'elsewhere', 'in',  
 'Canada']  
 ['Quebec', 'Official', 'language', 'in', 'Quebec', '(', 'as', 'French',  
 ')']  
 ['Quebec', 'Language', 'codes', 'Linguasphere', '51-AAA-hq', '&', 'amp',  
 ';', '51-AAA-ii']  
 ['Dida', 'Total', 'speakers', '200', ',', '000', '(', '1993', ')']

['Dida', 'Language', 'codes', 'ISO', '639-3', 'dic', '\xe2', '\x80', '\x93', 'Lakota', 'Dida']

['Taiwanese', 'Hokkien', 'Total', 'speakers', 'About', '15', 'million', 'in', 'Taiwan', ';', '49', 'million', '(', 'Min', 'Nan', 'as', 'a', 'group', ')']

['Taiwanese', 'Writing', 'system', 'Latin', '(', 'pe', '\xcc', '\x8d', 'h-', '\xc5', '\x8d', 'e-j', '\xc4', '\xab', ')', ',', 'Hanzi']

['Syriac', 'Language', 'extinction', 'marginalized', 'by', 'Arabic', 'from', 'the', '8th', 'century', ',', 'evolved', 'into', 'Neo-Aramaic', 'vernaculars', 'around', 'AD', '1200', '(', 'conventional', 'date', ')']

['Syriac', 'Language', 'codes', 'ISO', '639-3', 'syc', '\xe2', '\x80', '\x93', 'Syriac', '(', 'classical', ')']

['Balochi', 'Total', 'speakers', '7', '\xe2', '\x80', '\x93', '8', 'million', '(', '1998', ',', 'Ethnologue', ')']

['Balochi', 'Official', 'language', 'in', 'Pakistan', '(', 'Balochistan', ')', '(', 'Regional', 'Language', ')']

['Balochi', 'Regulated', 'by', 'Balochi', 'Academy', '(', 'Pakistan', ')']

['Balochi', 'Language', 'codes', 'ISO', '639-3', ':', 'bgp', '\xe2', '\x80', '\x93', 'Eastern', 'Balochi']

['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bgn', '\xe2', '\x80', '\x93', 'Western', 'Balochi']

['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bcc', '\xe2', '\x80', '\x93', 'Southern', 'Balochi']

['Balochi', 'Language', 'codes', 'Linguasphere', '58-AAB-a', '&', 'gt', ';', '58-AAB-aa', '(', 'East', 'Balochi', ')', '+', '58-AAB-ab', '(', 'West', 'Balochi', ')', '+', '58-AAB-ac', '(', 'South', 'Balochi', ')', '+', '58-AAB-ad', '(', 'Bashkardi', ')']

['Balochi', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Zulu', 'Total', 'speakers', 'First', 'language', '-', '10', 'million']

['Zulu', 'Total', 'speakers', 'Second', 'language', '-', '16', 'million']

['Zulu', 'Regulated', 'by', 'Pan', 'South', 'African', 'Language', 'Board']

['Nepali', 'Official', 'language', 'in', 'West', 'Bengal', '(', 'India', ')']

['Nepali', 'Regulated', 'by', 'Language', 'Academy', 'of', 'Nepal']

['Nepali', 'Language', 'codes', 'ISO', '639-3', 'World', 'map', 'with', 'significant', 'Nepali', 'language', 'speakers', 'Dark', 'Blue', ':', 'Main', 'official', 'language']

['Nepali', 'Language', 'codes', 'ISO', '639-3', 'Light', 'blue', ':', 'One', 'of', 'the', 'official', 'languages']

['Nepali', 'Language', 'codes', 'ISO', '639-3', 'Red', ':', 'Places', 'with', 'significant', 'population', 'or', 'greater', 'than', '20', '%', 'but', 'without', 'official', 'recognition', '.', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',

'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']

['Scots', 'Total', 'speakers', 'est', '.', '200', ',', '000', '(',  
 'ethnologue', ')', 'to', 'over', '1', '.', '5', 'million', '(', 'General',  
 'Register', 'Office', 'for', 'Scotland', ',', '1996', ')']

['Scots', 'Total', 'speakers', 'A', 'Scottish', 'Government', 'study',  
 'found', 'that', '85', '%', 'of', 'respondents', '(', 'being', 'a',  
 'representative', 'sample', 'of', 'Scotland's', 'adult', 'population', ')',  
 'claim', 'to', 'speak', 'Scots', 'to', 'varying', 'degrees']

['Scots', 'Region', 'Scotland', ':', 'Scottish', 'Lowlands']

['Scots', 'Official', 'language', 'in', '\xe2', '\x80', '\x94',  
 'Classified', 'as', 'a', '"', 'traditional', 'language', '"', 'by', 'the',  
 'Scottish', 'Government']

['Scots', 'Official', 'language', 'in', '\xe2', '\x80', '\x94',  
 'Classified', 'as', 'a', '"', 'regional', 'or', 'minority', 'language',  
 '"', 'under', 'the', 'European', 'Charter', 'for', 'Regional', 'or',  
 'Minority', 'Languages', ',', 'ratified', 'by', 'the', 'United', 'Kingdom',  
 'in', '2001']

['Scots', 'Official', 'language', 'in', '\xe2', '\x80', '\x94',  
 'Classified', 'as', 'a', '"', 'traditional', 'language', '"', 'by', 'The',  
 'North', '/', 'South', 'Language', 'Body']

['Scots', 'Regulated', 'by', '\xe2', '\x80', '\x94', 'Scotland', ':',  
 'None', ',', 'although', 'the', 'Dictionary', 'of', 'the', 'Scots',  
 'Language', 'carries', 'great', 'authority', '(', 'the', 'Scottish',  
 'Government's', 'Partnership', 'for', 'a', 'Better', 'Scotland',  
 'coalition', 'agreement', '(', '2003', ')', 'promises', '"', 'support',  
 '"', ')']

['Scots', 'Language', 'codes', 'Linguasphere', '52-ABA-aa', '(',  
 'varieties']

['Scots', 'Language', 'codes', 'Linguasphere', ':', '52-ABA-aaa', 'to', '-  
 aav', ')']

['Sumerian', 'Language', 'extinction', 'Effectively', 'extinct', 'from',  
 'about', '1800', 'BC', ';', 'used', 'as', 'classical', 'language', 'until',  
 'about', 'AD', '100']

['Hindi-Urdu', 'Total', 'speakers', 'Native', ':', '570', 'million', '(',  
 '2011', ')']

['Hindi-Urdu', 'Total', 'speakers', 'Second', 'language', ':', '490',  
 'million', '(', '2011', ')']

['Hindi-Urdu', 'Total', 'speakers', 'Total', ':', '970', 'million', '(',  
 '2011', ')']

['Hindi-Urdu', 'Spoken', 'in', 'India', 'and', 'Pakistan', '.', 'Also',  
 'various', 'other', 'countries', 'due', 'to', 'South', 'Asian',  
 'immigration']

['Hindi-Urdu', 'Standard', 'forms', 'Standard', 'Hindi']

['Hindi-Urdu', 'Standard', 'forms', 'Standard', 'Urdu']

['Hindi-Urdu', 'Official', 'language', 'in', 'India', '(', 'as', 'Hindi',



'and', 'Urdu', ')']  
 ['Hindi-Urdu', 'Official', 'language', 'in', 'Pakistan', '(', 'as', 'Urdu', ')']  
 ['Hindi-Urdu', 'Regulated', 'by', 'Central', 'Hindi', 'Directorate', '(', 'Hindi', ',', 'India', ')', ',', '']  
 ['Hindi-Urdu', 'Regulated', 'by', 'National', 'Language', 'Authority', ',', '(', 'Urdu', ',', 'Pakistan', ')', ';']  
 ['Hindi-Urdu', 'Regulated', 'by', 'National', 'Council', 'for', 'Promotion', 'of', 'Urdu', 'language', '(', 'Urdu', ',', 'India', ')']  
 ["Ge'ez", 'Language', 'extinction', 'Extinct', '.', 'Ceased', 'to', 'be', 'a', 'spoken', 'tongue', '(', 'in', '4th', 'century', 'CE', 'according', 'to', ')', '(', 'sometime', 'before', 'the', '10th', 'century', 'CE', 'according', 'to', ')', ',', 'remains', 'in', 'use', 'as', 'a', 'liturgical', 'language']  
 ['Ainu', 'Total', 'speakers', '~', '100s']  
 ['Kazakh', 'Total', 'speakers', '8-8', '.', '3', 'million']  
 ['Uzbek', 'Total', 'speakers', '23', '.', '5', 'million']  
 ['Uzbek', 'Language', 'codes', 'ISO', '639-3', ':', 'uzn', '\xe2', '\x80', '\x93', 'Northern', 'Uzbek']  
 ['Uzbek', 'Language', 'codes', 'ISO', '639-3', 'uzs', '\xe2', '\x80', '\x93', 'Southern', 'Uzbek']  
 ['Tok', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Tok', 'Official', 'language', 'in', 'Papua', 'New', 'Guinea']  
 ['Serbian', 'Regulated', 'by', 'Board', 'for', 'Standardization', 'of', 'the', 'Serbian', 'Language']  
 ['Serbian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-g']  
 ['(Paraguayan)', 'Guaran', '\xc3', '\xad', 'Total', 'speakers', '4', ',', '650', ',', '000', 'in', 'Paraguay', '(', 'SIL', '1995', ')', ',', '200', ',', '000', 'in', 'Argentina']  
 ['(Paraguayan)', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Guaran', '\xc3', '\xad', 'variant', ')']  
 ['Lithuanian', 'Total', 'speakers', '3', '.', '5', 'million', '(', 'Lithuania', ')']  
 ['Lithuanian', 'Regulated', 'by', 'Commission', 'of', 'the', 'Lithuanian', 'Language']  
 ['Alsatian', 'Total', 'speakers', 'more', 'than', '700', ',', '000']  
 ['Kyrgyz', 'Total', 'speakers', 'Approx', '.', '4', '.', '5', 'million']  
 ['Kyrgyz', 'Writing', 'system', 'Cyrillic', 'alphabet', '(', 'Kyrgyz', 'variant', ')', ';', 'Arabic', 'alphabet', '(', 'Kyrgyz', 'variant', ')']  
 ['Ganda', 'Total', 'speakers', 'First', 'language', ':', '6', 'million', '(', '2008', ')']  
 ['Ganda', 'Total', 'speakers', 'Second', 'language', ':', '4', 'million', '(', '2008', ')']  
 ['Tigrinya', 'Total', 'speakers', '6', '.', '7', 'million']  
 ['Wolof', 'Total', 'speakers', '5', '.', '5', 'million', '(', 'mother', 'tongue', ')']

['Wolof', 'Total', 'speakers', '12', '.', '3', 'million', '(', 'second',  
 'language', ')']  
 ['Xhosa', 'Total', 'speakers', '7', '.', '9', 'million']  
 ['Xhosa', 'Region', 'Eastern', 'Cape', ',', 'Western', 'Cape']  
 ['Classical', 'Latin', 'Spoken', 'in', 'Roman', 'republic', ',', 'Roman',  
 'empire']  
 ['Classical', 'Language', 'extinction', 'developed', 'into', 'Late',  
 'Latin']  
 ['Classical', 'Language', 'codes', 'ISO', '639-3', 'The', 'range', 'of',  
 'Latin', ',', '60', 'AD']  
 ['Pashto', 'Language', 'codes', 'ISO', '639-3', ':', 'pst', '\xe2', '\x80',  
 '\x93', 'Central', 'Pashto']  
 ['Pashto', 'Language', 'codes', 'ISO', '639-3', 'pbu', '\xe2', '\x80',  
 '\x93', 'Northern', 'Pashto']  
 ['Pashto', 'Language', 'codes', 'ISO', '639-3', 'pbt', '\xe2', '\x80',  
 '\x93', 'Southern', 'Pashto']  
 ['Malay', 'Spoken', 'in', 'Malaysia', '(', 'as', 'Malaysian', ')']  
 ['Malay', 'Spoken', 'in', 'Indonesia', '(', 'as', 'Indonesian', ')']  
 ['Malay', 'Spoken', 'in', 'East', 'Timor', '(', 'as', 'Indonesian', ')']  
 ['Malay', 'Total', 'speakers', '40', 'million', 'native', '(', 'incl', '.',  
 '23', 'million', 'Indonesian', ')', ';', 'approx', '.', '180', 'million',  
 'total', '(', '90', '%', 'Indonesian', ')']  
 ['Malay', 'Writing', 'system', 'Rumi', '(', 'Latin', 'alphabet', ')', ',',  
 'Malay', 'Variant', '(', 'official', 'in', 'Malaysia', ',', 'Singapore',  
 'and', 'Indonesia', ';', 'co-official', 'in', 'Brunei', ')']  
 ['Malay', 'Writing', 'system', 'Jawi', '(', 'Arabic', 'script', ')', '(',  
 'co-official', 'in', 'Brunei', 'and', 'Malaysia', ')']  
 ['Malay', 'Writing', 'system', 'Historically', 'written', 'in', 'Pallava',  
 ',', 'Kawi', 'and', 'Rencong']  
 ['Malay', 'Regulated', 'by', 'Dewan', 'Bahasa', 'dan', 'Pustaka', '(',  
 'Institute', 'of', 'Language', 'and', 'Literature', ';']  
 ['Malay', 'Regulated', 'by', 'Majlis', 'Bahasa', 'Brunei', '-',  
 'Indonesia', '-', 'Malaysia', '(', 'Brunei', '-', 'Indonesia', '-',  
 'Malaysia', 'Language', 'Council', '\xe2', '\x80', '\x94', 'MABBIM', ')',  
 '(', 'a', 'trilateral', 'joint-venture', ')']  
 ['Malay', 'Recognised', 'minority', 'language', 'in', 'Indonesia', '(',  
 'Malay', 'enjoys', 'a', 'status', 'as', 'a', 'regional', 'language', 'in',  
 'Indonesia', ',', 'separated', 'from', 'Indonesian', 'which', 'receives',  
 'a', 'national', 'language', 'status', ')']  
 ['Malay', 'Recognised', 'minority', 'language', 'in', 'Thailand', '(',  
 'Malay', 'is', 'the', 'language', 'of', 'the', 'Muslim', 'community', 'in',  
 'Southern', 'Thailand', ')']  
 ['Malay', 'Language', 'codes', 'ISO', '639-3', 'mhp', '\xe2', '\x80',  
 '\x93', 'Balinese', 'Malay']  
 ['Malay', 'Language', 'codes', 'ISO', '639-3', 'lrt', '\xe2', '\x80',  
 '\x93', 'Larantuka', 'Malay']

['Malay', 'Language', 'codes', 'ISO', '639-3', 'mfp', '\xe2', '\x80', '\x93', 'Makassar', 'Malay']

['Malay', 'Language', 'codes', 'ISO', '639-3', 'max', '\xe2', '\x80', '\x93', 'North', 'Moluccan', 'Malay']

['Valencian', 'Total', 'speakers', '2', '.', '4', 'million', '(', '2004', ')']

['Valencian', 'Total', 'speakers', '(', '11', '.', '5', 'milion', 'in', 'the', 'whole', 'dominion', ')']

['Valencian', 'Official', 'language', 'in', ':', 'Valencian', 'Community']

['Tutchone', 'Total', 'speakers', '400', 'in', '1995']

['Tutchone', 'Region', '1400', 'Southern', 'Tutchone', 'people', ',', '1000', 'Northern', 'Tutchone', 'people']

['Lao', 'Total', 'speakers', '5', ',', '225', ',', '552', '(', '2006', ')', ',', 'roughly', '20', 'million', 'if', 'Isan', 'speakers', 'are', 'included']

['Lao', 'Spoken', 'in', 'U', '.', 'S']

['Gujarati', 'Total', 'speakers', '46', '.', '1', 'million']

['Gujarati', 'Spoken', 'in', 'U', '.', 'S']

['Gujarati', 'Writing', 'system', 'former', 'use', 'of', 'Devanagari', 'before', 'invention', 'of', 'Gujarati', 'Script']

['Gujarati', 'Writing', 'system', 'also', 'use', 'of', 'Arabic', 'script', 'by', 'the', 'Ismaili', 'community', 'and', 'other', 'Gujarati', 'communities']

['Gujarati', 'Official', 'language', 'in', 'Daman', 'and', 'Diu', '(', 'India', ')']

['Gujarati', 'Language', 'codes', 'ISO', '639-3', 'Countries', 'where', 'Gujarati', 'is', 'spoken', '.', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Southern', 'Ndebele', 'Total', 'speakers', '586', ',', '961', '(', '1996', 'census', ')']

['Dalmatian', 'Language', 'extinction', '10', 'June', '1898', ',', 'when', 'Tuone', 'Udaina', 'was', 'killed']

['Vandalic', 'Language', 'extinction', '6th', 'century', 'AD']

['Vandalic', 'Spoken', 'in', 'Spain', ',', 'North', 'Africa']

['Romani', 'Spoken', 'in', 'Central', 'and', 'Eastern', 'Europe', ',', 'worldwide']

['Romani', 'Total', 'speakers', '2', '.', '5', 'million', '(', 'SIL', 'estimate', ')']

['Romani', 'Official', 'language', 'in', '(', 'not', 'official', 'language', ')', 'recognised', 'as', 'minority', 'language', 'only', 'in', 'parts', 'of']

['Romani', 'Language', 'codes', 'ISO', '639-3', ':', 'rmn', '\xe2', '\x80', '\x93', 'Balkan', 'Romani']

['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmf', '\xe2', '\x80', '\x93', 'Balkan', 'Romani']

'\x93', 'Finnish', 'Kalo']  
 ['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmy', '\xe2', '\x80', '\x93', 'Vlax', 'Romani']  
 ['Romani', 'Language', 'codes', 'ISO', '639-3', 'rmw', '\xe2', '\x80', '\x93', 'Welsh', 'Romani']  
 ['Yue', 'Total', 'speakers', '56', 'million', 'in', '1984']  
 ['Yue', 'Region', 'in', 'China', ':', 'the', 'Pearl', 'River', 'Delta', '(', 'central', 'Guangdong', ';', 'Hong', 'Kong', ',', 'Macau', ')', ';', 'the', 'eastern', 'and', 'southern', 'Guangxi', ';', 'parts', 'of', 'Hainan', ';']  
 ['Yue', 'Spoken', 'in', 'China', ',', 'Malaysia', '(', 'Kuala', 'Lumpur', ',', 'Ipoh', ',', 'Kota', 'Kinabalu', ')', ',', 'Vietnam', ',', 'Indonesia', '(', 'Medan', ')', ',', 'United', 'States', ',', 'Canada', ',', 'United', 'Kingdom', 'and', 'other', 'countries', 'where', 'Cantonese', 'migrants', 'have', 'settled']  
 ['Yue', 'Official', 'language', 'in', 'Hong', 'Kong', 'and', 'Macau', '(', 'de', 'facto', ',', 'though', 'officially', 'referred', 'to', 'as', '"', 'Chinese', '"', ';', 'Cantonese', 'and', 'occasionally', 'Mandarin', 'are', 'used', 'in', 'government', ')', '.', 'Recognised', 'regional', 'language', 'in', 'Suriname']  
 ['Siwi', 'Total', 'speakers', 'about', '15', ',', '000']  
 ['Zenaga', 'Total', 'speakers', 'about', '200']  
 ['Tuareg', 'language', '(', 's', ')', 'Total', 'speakers', '1', '.', '2', 'million', '(', 'Ethnologue', ')']  
 ['Ancient', 'Greek', 'Language', 'extinction', 'developed', 'into', 'Koin', '\xc3', '\xa9', 'Greek', 'by', 'the', '4th', 'century', 'BC']  
 ['Black', 'Category', '(', 'purpose', ')', 'Black', 'Speech']  
 ['Black', 'Date', 'founded', 'c', '.', '1945', '\xe2', '\x80', '\x93', '1973']  
 ['Polabian', 'Language', 'extinction', '18th', 'century']  
 ['Polabian', 'Language', 'codes', 'Linguasphere', 'Former', 'settlement', 'area', 'of', 'the', 'Polabian', 'Slavs']  
 ['Slovincian', 'Language', 'extinction', '20th', 'century']  
 ['Rusyn', 'Total', 'speakers', 'Estimated', ':', 'At', 'least', '600', ',', '000']  
 ['Rusyn', 'Total', 'speakers', 'Census', 'population', ':', '50', ',', '000', '.', 'These', 'are', 'numbers', 'from', 'national', 'official', 'bureaus', 'for', 'statistics']  
 ['Rusyn', 'Total', 'speakers', ':', 'Slovakia', '-', '24', ',', '201']  
 ['Rusyn', 'Total', 'speakers', 'Serbia', '-', '15', ',', '626']  
 ['Rusyn', 'Total', 'speakers', 'Ukraine', '-', '6', ',', '725']  
 ['Rusyn', 'Total', 'speakers', 'Croatia', '-', '2', ',', '337']  
 ['Rusyn', 'Total', 'speakers', 'Poland', '-', '5', ',', '800']  
 ['Rusyn', 'Total', 'speakers', 'Hungary', '-', '1', ',', '098']  
 ['Rusyn', 'Total', 'speakers', 'Czech', 'Republic', '-', '1', ',', '106']  
 ['Rusyn', 'Language', 'codes', 'Linguasphere', '53-AAA-ec', '&', 'lt', ';',

'53-AAA-e']  
 ['Rusyn', 'Language', 'codes', 'Linguasphere', '(', 'varieties', ':', '53-AAA-eca', 'to', '53-AAA-ecc', ')']  
 ['Romano-Serbian', 'Total', 'speakers', '172', ',', '000', '(', 'in', '2006', ')']  
 ['Manchu', 'Total', 'speakers', '60', '(', '1999', 'Zhao', 'Aping', ')']  
 ['Cimbrian', 'Total', 'speakers', '2', ',', '230', '&', 'gt', ';']  
 ['Austro-Bavarian', 'Spoken', 'in', 'Italy', ',', 'South', 'Tyrol']  
 ['Galindan', 'Galindian', 'Language', 'extinction', '14th', 'century']  
 ['Galindan', 'Spoken', 'in', 'Northern', 'Poland', 'and', 'Russia']  
 ['Mongolian', 'Total', 'speakers', '5', '.', '2', 'million']  
 ['Mongolian', 'Spoken', 'in', 'Mongolia', ',', 'People's', 'Republic', 'of', 'China']  
 ['Mongolian', 'Writing', 'system', 'Mongolian', 'script', ',', 'Mongolian', 'Cyrillic', 'alphabet']  
 ['Mongolian', 'Official', 'language', 'in', 'People's', 'Republic', 'of', 'China', '(', 'Inner', 'Mongolia', ')']  
 ['Mongolian', 'Regulated', 'by', 'State', 'Language', 'Council', '(', 'Mongolia', ')', ',', 'Council', 'for', 'Language', 'and', 'Literature', 'Work', '(', 'Inner', 'Mongolia', ')']  
 ['Mongolian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-BAA-b']  
 ['Rotokas', 'Total', 'speakers', '4', ',', '320']  
 ['Rotokas', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Luba', 'Total', 'speakers', '9', ',', '500', ',', '000', '(', '2009', ')']  
 ['Burushaski', 'Total', 'speakers', '87', ',', '000', '(', '2000', ')']  
 ['Acadian', 'Total', 'speakers', '371', ',', '614', '(', '1996', ')']  
 ['Acadian', 'Total', 'speakers', '371', ',', '590', '(', '2006', ')', ',', 'Canadian', 'census', ',', 'ethnic', 'data']  
 ['Norwegian', 'Spoken', 'in', 'Norway', '(', '4', '.', '8', 'million', ')', ',', '']  
 ['Norwegian', 'Spoken', 'in', 'United', 'States', '(', '55', ',', '311', ')']  
 ['Norwegian', 'Spoken', 'in', 'Canada', '(', '7', ',', '710', ')']  
 ['Norwegian', 'Standard', 'forms', 'Nynorsk', '(', 'official', ')']  
 ['Norwegian', 'Standard', 'forms', 'Bokm', '\xc3', '\xa5', 'l', '(', 'official', ')', '/', 'Riksm', '\xc3', '\xa5', 'l', '(', 'unofficial', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Language', 'Council', '(', 'Bokm', '\xc3', '\xa5', 'l', 'and', 'Nynorsk', ')']  
 ['Norwegian', 'Regulated', 'by', 'Norwegian', 'Academy', '(', 'Riksm', '\xc3', '\xa5', 'l', ')']  
 ['Dungan', 'Total', 'speakers', '41', ',', '400', '(', '2001', ')']  
 ['Dungan', 'Region', 'Fergana', 'Valley', ',', 'Chu', 'Valley']  
 ['Norwegian', 'Spoken', 'in', 'Norway', '(', '4', '.', '8', 'million', ')', ',', '']

['Norwegian', 'Spoken', 'in', 'United', 'States', '(', '55', ',', '311', ')']

['Norwegian', 'Spoken', 'in', 'Canada', '(', '7', ',', '710', ')']

['Norwegian', 'Standard', 'forms', 'Nynorsk', '(', 'official', ')']

['Norwegian', 'Standard', 'forms', 'Bokm', '\xc3', '\xa5', 'l', '(', 'official', ')', '/', 'Riksm', '\xc3', '\xa5', 'l', '(', 'unofficial', ')']

['Norwegian', 'Regulated', 'by', 'Norwegian', 'Language', 'Council', '(', 'Bokm', '\xc3', '\xa5', 'l', 'and', 'Nynorsk', ')']

['Norwegian', 'Regulated', 'by', 'Norwegian', 'Academy', '(', 'Riksm', '\xc3', '\xa5', 'l', ')']

['Brandenburgisch', 'Language', 'codes', 'ISO', '639-3', 'Low', 'Saxon', 'and', 'Low', 'Franconian', 'dialects']

['Haz\xc4\x81rag\xc4\xab', 'Total', 'speakers', '1', '.', '8', 'to', 'over', '2', '.', '2', 'million']

['Navajo', 'Total', 'speakers', '170', ',', '717', '(', '2007', ')']

['Bosnian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-g']

['Okinawan', 'Total', 'speakers', '984', ',', '000', '(', '2000', ')']

['Okinawan', 'Writing', 'system', 'R', '\xc5', '\x8d', 'maji']

['Texas', 'German', 'Total', 'speakers', '~', '3', ',', '000', ',', 'declining']

['Hutterite', 'Spoken', 'in', 'North', 'and', 'South', 'Dakota']

['Hutterite', 'Total', 'speakers', '34', ',', '000', '(', '2003', 'estimate', ')']

['Plautdietsch', 'Total', 'speakers', '260', ',', '710', '\xe2', '\x80', '\x93', '318', ',', '500']

['Mohawk', 'Region', 'Ontario', ',', 'Quebec', 'and', 'northern', 'New', 'York']

['Chinook', 'Jargon', 'Total', 'speakers', 'Unknown', '.', 'Definitely', 'more', 'than', '10', '-', 'Also', 'a', 'small', ',', 'yet', 'present', 'revival', 'effort', 'in', 'place']

['Chinook', 'Region', 'Pacific', 'Northwest', '(', 'Interior', 'and', 'Coast', ')']

['Chinook', 'Writing', 'system', 'de', 'facto', 'Latin', 'Alphabet']

['Chinook', 'Official', 'language', 'in', 'De', 'facto', 'in', 'Pacific', 'Northwest', 'until', 'about', '1900']

['Chechen', 'Total', 'speakers', 'circa', '1', ',', '500', ',', '000', 'self-reported', 'speakers', 'worldwide', '(', '2009', ')']

['East', 'Frisian', 'Low', 'Saxon', 'Total', 'speakers', '230', ',', '000', 'in', 'Germany']

['Taiwanese', 'Language', 'codes', 'ISO', '639-3', 'tss', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Gullah', 'Total', 'speakers', '250', ',', '000']

['Gullah', 'Region', 'Coastal', 'low', 'country', 'region', 'of', 'South', 'Carolina', 'and', 'Georgia', 'including', 'the', 'Sea', 'Islands']

['Arvanitika', 'Total', 'speakers', '30', ',', '000', '-', '150', ',', '000']

'000']

['Arvanitika', 'Region', 'Saronic', 'Islands', ';', 'Western', 'Thrace',  
;', 'Peloponnese', ';', 'some', 'villages', 'in', 'NW', 'of', 'Greece',  
;', 'N', 'of', 'island', 'of', 'Andros', ';', 'more', 'than', '500',  
'villages', 'in', 'total']

['Arvanitika', 'Writing', 'system', 'Greek', 'alphabet', '(', 'Arvanitic',  
'variant', ')']

['Arvanitika', 'Language', 'codes', 'Linguasphere', '55-AAA-aea', 'to',  
'55-AAA-aed']

['Modern', 'English', 'Total', 'speakers', 'First', 'language', ':', '309',  
'\xe2', '\x80', '\x93', '380', 'million']

['Modern', 'Total', 'speakers', 'Overall', ':', '\xe2', '\x89', '\x88',  
'1', '.', '8', 'billion']

['Modern', 'Pronunciation', '/', '\xcb', '\x88', '\xc9', '\xaa', '\xc5',  
'\x8b', '\xc9', '\xa1', 'l', '\xc9', '\xaa', '\xca', '\x83', '/', '1']

['Modern', 'Language', 'codes', 'ISO', '639-3', 'Countries', 'where',  
'English', 'is', 'a', 'primary', 'language', 'are', 'dark', 'blue', ';',  
'countries', 'where', 'it', 'is', 'an', 'official', 'but', 'not', 'a',  
'primary', 'language', 'are', 'light', 'blue', '.', 'English', 'is',  
'also', 'one', 'of', 'the', 'official', 'languages', 'of', 'the',  
'European', 'Union']

['Assamese', 'Total', 'speakers', '13', ',', '079', ',', '696', '(', 'in',  
'1991', ')']

['Assamese', 'Spoken', 'in', 'India', ',', 'Bhutan', '&', 'amp', ';',  
'USA', '(', 'DE', ',', 'NJ', '&', 'amp', ';', 'NY', ')']

['Assamese', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
'of', 'conjuncts', '.', 'More']

['Spanish', 'Language', 'codes', 'ISO', '639-3', 'Spanish', 'language',  
'spread', 'in', 'the', 'United', 'States']

['Lydian', 'Language', 'extinction', '1st', 'c', '.', 'BC']

['Lycian', 'Language', 'extinction', '1st', 'century', 'BC']

['Lycian', 'Writing', 'system', 'Lycian', 'script', 'Names', 'and',  
'inscriptions', ',', 'presumed', 'spoken']

['Tatar', 'Total', 'speakers', '6', ',', '496', ',', '600']

['Tatar', 'Writing', 'system', 'Cyrillic', 'alphabet', ',', 'Latin',  
'alphabet', 'and', 'Arabic', 'alphabet']

['Tatar', 'Regulated', 'by', 'Institute', 'of', 'Language', ',',  
'Literature', 'and', 'Arts', 'of', 'the', 'Academy', 'of', 'Sciences',  
'of', 'the', 'Republic', 'of', 'Tatarstan']

['Azerbaijani', 'Writing', 'system', 'Latin', 'alphabet', 'for', 'North',  
'Azeri', 'in', 'Azerbaijan', ',', 'Perso-Arabic', 'script', 'for', 'South',  
'Azeri', 'in', 'Iran']

['Azerbaijani', 'Official', 'language', 'in', 'Azerbaijan', '(', 'North',  
'Azeri', ')']

['Azerbaijani', 'Official', 'language', 'in', 'Iran', '(', 'South', 'Azeri', ')', '-', 'constitutional', 'status', 'as', 'a', 'regional', 'language']

['Azerbaijani', 'Official', 'language', 'in', 'Iraq', '(', 'South', 'Azeri', ')', '-', 'constitutional', 'status', 'as', 'a', 'regional', 'language', '(', 'referred', 'to', 'as', 'Turkmen', 'in', 'the', 'constitution', ')']

['Azerbaijani', 'Official', 'language', 'in', 'Russia', '-', 'One', 'of', 'the', 'official', 'languages', 'of', 'Dagestan', 'Republic']

['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', ':', 'azj', '\xe2', '\x80', '\x93', 'North', 'Azerbaijani']

['Azerbaijani', 'Language', 'codes', 'ISO', '639-3', 'azb', '\xe2', '\x80', '\x93', 'South', 'Azerbaijani']

['Azerbaijani', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']

['Azerbaijani', 'Language', 'codes', 'Linguasphere', 'Map', 'showing', 'locations', 'of', 'Azerbaijani', 'speakers', '(', 'until', '1991', ')']

['Shelta', 'Total', 'speakers', '6', ',', '000', 'in', 'Ireland', ';;', '86', ',', '000', 'world-wide', 'Irish', 'Travelers']

['Shelta', 'Region', 'Used', 'by', 'some', 'Irish', 'Travellers']

['Southern', 'Region', 'Southern', 'Fujian', 'province', ';;', 'the', 'Chaozhou-Shantou', '(', 'Chaoshan', ')', 'area', 'and', 'Leizhou', 'Peninsula', 'in', 'Guangdong', 'province', ';;', 'extreme', 'south', 'of', 'Zhejiang', 'province', ';;', 'much', 'of', 'Hainan', 'province', '(', 'if', 'Hainanese', 'or', 'Qiong', 'Wen', 'is', 'included', ')', ';;', 'and', 'most', 'of', 'Taiwan', ';;']

['Southern', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']

['Southern', 'Spoken', 'in', 'United', 'States', '(', 'New', 'York', 'City', ')']

['Southern', 'Spoken', 'in', 'and', 'other', 'areas', 'of', 'Southern', 'Min', 'and', 'Hoklo', 'settlement']

['Southern', 'Official', 'language', 'in', 'None', '(', 'Legislative', 'bills', 'have', 'been', 'proposed', 'for', 'Taiwanese', '(', 'Amoy', 'Southern', 'Min', ')', 'to', 'be', 'one', 'of', 'the', '"national", "languages"', 'in', 'Taiwan', ')', ';;', 'one', 'of', 'the', 'statutory', 'languages', 'for', 'public', 'transport', 'announcements', 'in', 'the', 'ROC']

['Southern', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'Southern', 'Min']

['Ubykh', 'Language', 'extinction', 'October', '1992', 'when', 'Tevfik', 'Esen', '\xc3', '\xa7', 'died']

['Chagatai', 'Spoken', 'in', 'Central', 'Asia', ',', 'Khorasan']

['Proto-Germanic', 'Language', 'extinction', 'evolved', 'into', 'Proto-Norse']

['Proto-Germanic', 'Language', 'extinction', 'Frankish', 'and', 'Ingvaeonic', 'by', 'the', '4th', 'century']



['Abkhaz', 'Total', 'speakers', '125', '000']  
 ['Abkhaz', 'Writing', 'system', 'Cyrillic', 'alphabet', '(', 'Abkhaz',  
 'variant', ')']  
 ['Northern', 'Thai', 'Language', 'Total', 'speakers', '6', 'million']  
 ['Northern', 'Writing', 'system', 'Tai', 'Tham', 'script', ',', 'Thai',  
 'script']  
 ['Northern', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Dyirbal', 'Total', 'speakers', '~', '5']  
 ['Cheyenne', 'Total', 'speakers', '1', ',', '721']  
 ['Tanacross', 'Total', 'speakers', '30', '(', '2000', ')']  
 ['Tanacross', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan',  
 'variant', ')']  
 ['Nux\xc3\xa1lk', '(', 'Bella', 'Coola', ')', 'Total', 'speakers', '20']  
 ['Filipino', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Filipino', 'Regulated', 'by', '(', 'Komisyon', 'sa', 'Wikang',  
 'Filipino', ')']  
 ['Middle', 'Chinese', 'Spoken', 'in', 'China']  
 ['Middle', 'Language', 'extinction', 'Evolved', 'into', 'Proto-Mandarin',  
 'and', 'other', 'Chinese', 'dialects', 'apart', 'from', 'Min']  
 ['Middle', 'Official', 'language', 'in', 'Southern', 'and', 'Northern',  
 'Dynasties']  
 ['Sona', 'Category', '(', 'sources', ')', 'The', 'language', 'has', '375',  
 'radicals', 'or', 'root', 'words', '--', 'based', 'on', 'the', 'terms',  
 'in', 'Roget's', 'original', 'thesaurus', '.', 'Ideas', 'and', 'sentences',  
 'are', 'formed', 'by', 'juxtaposing', 'the', 'radicals']  
 ['Old', 'Curonian', 'language', 'Language', 'extinction', '14-17th',  
 'century']  
 ['Nivkh', 'Total', 'speakers', '1', ',', '089', '(', '1989', 'census', ')']  
 ['Ga', 'Total', 'speakers', '600', ',', '000', '(', 'SIL', '2004', ')']  
 ['Ga', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Ga', 'variant', ')']  
 ['Tulu', 'Total', 'speakers', '1', '.', '95', 'million', '(', '1997', ')']  
 ['Tulu', 'Region', 'Tulu', 'Nadu', ':', 'of', 'Karnataka', 'and', 'Kerala',  
 'States']  
 ['Tulu', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'native', 'Tulu', 'speakers', 'in', 'India', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Chukchi', 'Total', 'speakers', '7742', '(', '2002', 'Census', ')']  
 ['Karelian', 'Total', 'speakers', '118', ',', '000']  
 ['Karelian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Karelian',  
 'variant', ')']

['Veps', 'language', 'Total', 'speakers', '6', ',', '355', '(', '1989', 'census', ')']

['Veps', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Vepsian', 'variant', ')']

['Turkmen', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']

['Proven\u00e7al', 'Total', 'speakers', '362', ',', '000']

['Kodava', 'Total', 'speakers', '122', ',', '000', '(', '1997', ')']

['Kodava', 'Writing', 'system', 'Kannada', 'script', '&', 'amp', ';', 'Coorgi-Cox', 'alphabet']

['Kodava', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Mirandese', 'Total', 'speakers', '10', ',', '000-15', ',', '000']

['Mirandese', 'Region', 'Northeastern', 'Portugal', '(', 'municipalities', 'of', 'Miranda', 'do', 'Douro', ',', 'Vimioso', 'and', 'Mogadouro', ')']

['Modern', 'Language', 'codes', 'ISO', '639-3', 'ell', '\u00e2', '\u0080', '\u0093', 'Modern', 'Greek']

['Modern', 'Language', 'codes', 'Linguasphere', 'part', 'of', '56-AAA-a']

['Upper', 'Sorbian', 'Total', 'speakers', '40', ',', '000']

['Upper', 'Language', 'codes', 'Linguasphere', '53-AAA-bb', '&', 'lt', ';', '53-AAA-b', '&', 'lt', ';', '53-AAA-b', '.', '.', '.', '-d', '(', 'varieties', ':', '53-AAA-bba', 'to', '53-AAA-bbf', ')']

['Lower', 'Sorbian', 'Total', 'speakers', '14', ',', '000']

['Lower', 'Language', 'codes', 'Linguasphere', '53-AAA-ba', '&', 'lt', ';', '53-AAA-b', '&', 'lt', ';', '53-AAA-b', '.', '.', '.', '-d', '(', 'varieties', ':', '53-AAA-baa', 'to', '53-AAA-bah', ')']

['Abaza', 'Total', 'speakers', '48', ',', '280']

['Abaza', 'Official', 'language', 'in', 'Russian', 'autonomous', 'republic', 'of', 'Karachay-Cherkess']

['Sundanese', 'Region', 'parts', 'of', 'western', 'Central', 'Java']

['Walloon', 'Total', 'speakers', 'est', '.', '600', ',', '000']

['Orokolo', 'Total', 'speakers', '13', ',', '000']

['Orokolo', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Tiwa', 'Total', 'speakers', '&', 'lt', ';', '3000']

['Tiwa', 'Region', 'New', 'Mexico', ',', 'Arizona', 'Ethnicity']

['Tiwa', 'Language', 'codes', 'ISO', '639-3', ':', 'tix', '\u00e2', '\u0080', '\u0093', 'Southern', 'Tiwa']

['Tiwa', 'Language', 'codes', 'ISO', '639-3', 'twf', '\u00e2', '\u0080', '\u0093', 'Northern', 'Tiwa']

['Cumbric', 'Language', 'extinction', '11th', '-', '12th', 'century']

['Cumbric', 'Spoken', 'in', 'Lancashire', 'and', 'possibly', 'North', 'Yorkshire']

['Camuno', '(', '?', ')', 'Language', 'extinction', 'first', 'millennium', 'AD']

['Camuno', 'Spoken', 'in', 'the', 'south', 'side', 'of', 'the', 'Central',

'Alps', '(', 'Valcamonica', ',', 'Valtellina', ')', ',', 'Italy']  
 ['Western', 'Total', 'speakers', '879', ',', '612', '-', 'Out', 'of', '6',  
 '.', '7', 'million', 'total', 'speakers', 'of', 'Armenian', '(', 'both',  
 'Western', 'and', 'Eastern', 'Armenian', ')']  
 ['Western', 'Language', 'codes', 'ISO', '639-3', ':', 'hye', '\xe2',  
 '\x80', '\x93', 'Modern', 'Armenian']  
 ['Western', 'Language', 'codes', 'ISO', '639-3', 'xcl', '\xe2', '\x80',  
 '\x93', 'Classical', 'Armenian']  
 ['Western', 'Language', 'codes', 'ISO', '639-3', 'axm', '\xe2', '\x80',  
 '\x93', 'Middle', 'Armenian']  
 ['Western', 'Language', 'codes', 'Linguasphere', '57-AAA-aca', 'to', '57-  
 AAA-act']  
 ['Friulan', 'Total', 'speakers', '794', ',', '000']  
 ['Swabian', 'Total', 'speakers', '819', ',', '000', '(', '2006', ')']  
 ['Old', 'French', 'Language', 'extinction', 'evolved', 'into', 'Middle',  
 'French', 'by', 'the', '14th', 'century']  
 ['Sotho', 'Regulated', 'by', 'Pan', 'South', 'African', 'Language',  
 'Board']  
 ['Sindhi', 'Dialects', 'of', 'Sindhi', 'Spoken', 'in']  
 ['Sindhi', 'Dialects', 'South', 'Asia', 'Total', 'speakers']  
 ['Sindhi', 'Dialects', '35', 'million', 'Language', 'family']  
 ['Sindhi', 'Dialects', 'Sri', 'Lanka', 'Region']  
 ['Sindhi', 'Regulated', 'by', 'Sindhi', 'Language', 'Authority', '(',  
 'Pakistan', ')', ',', '']  
 ['Sindhi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Tup\xc3\xad', 'Language', 'extinction', '(', 'survives', 'as',  
 'Nheengatu', ')']  
 ['Tup\xc3\xad', 'Dialects', 'L', '\xc3', '\xad', 'ngua', 'Geral', 'of',  
 'S', '\xc3', '\xa3', 'o', 'Paulo', '(', 'Tup', '\xc3', '\xad', 'Austral',  
 ')']  
 ['Tup\xc3\xad', 'Language', 'codes', 'ISO', '639-3', ':', 'tpw', '\xe2',  
 '\x80', '\x93', 'Tup', '\xc3', '\xad', '(', 'Old', 'Tup', '\xc3', '\xad',  
 ')']  
 ['Tajik', 'Total', 'speakers', 'approximately', '24', ',', '380', ',', '000',  
 '(', '1991', ')']  
 ['Uyghur', 'Official', 'language', 'in', 'Xinjiang', 'Uyghur',  
 'Autonomous', 'Region']  
 ['Uyghur', 'Regulated', 'by', 'Working', 'Committee', 'of', 'Ethnic',  
 'Language', 'and', 'Writing', 'of', 'Xinjiang', 'Uyghur', 'Autonomous',  
 'Region']  
 ['Oromo', 'Total', 'speakers', 'From', 'more', 'than', 'the', '70', ',', '000',  
 ',', '000', 'Ethiopians', ',', 'the', '2007', 'census', 'lists',  
 '25', ',', '488', ',', '344', 'ethnic', 'Oromos', ',', 'with', 'a',

'number', 'of', 'others', 'also', 'speaking', 'Oromo', '.', 'The',  
 'Ethnologue', 'lists', 'about', '263', ',', '000', 'speakers', 'of',  
 'Oromo', 'languages', 'in', 'neighbouring', 'Somalia', 'and', 'Kenya']  
 ['Oromo', 'Writing', 'system', 'Latin', 'Qubee', ',', 'formerly', 'also',  
 'Ethiopic', 'syllabary']  
 ['Oromo', 'Language', 'codes', 'ISO', '639-3', 'hae', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Oromo']  
 ['Oromo', 'Language', 'codes', 'ISO', '639-3', 'gaz', '\xe2', '\x80',  
 '\x93', 'West', 'Central', 'Oromo']  
 ['Oriya', 'Total', 'speakers', '31', 'million', '(', '1996', ')']  
 ['Oriya', 'Language', 'codes', 'Linguasphere', 'Distribution', 'of',  
 'Oriya', 'speakers', 'in', 'World', 'This', 'page', 'contains', 'Indic',  
 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see',  
 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts',  
 '.', 'More']  
 ['Shona', 'Total', 'speakers', '11', ',', '000', ',', '000']  
 ['Aleut', 'Total', 'speakers', 'ca', '.', '150', 'in', '2007']  
 ['Hindko', 'Total', 'speakers', 'approx', '.', '2', '.', '2', '\xe2',  
 '\x80', '\x93', '4', 'million']  
 ['Hindko', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Lower', 'Silesian', 'language', 'Total', 'speakers', '22', ',', '900']  
 ['Lower', 'Region', 'Lower', 'Silesia', '.', 'Also', 'spoken', 'in',  
 'Czech', 'Republic', ',', 'eastern', 'Germany']  
 ['Mari', 'Standard', 'forms', 'Meadow', 'Mari']  
 ['Mari', 'Standard', 'forms', 'Hill', 'Mari']  
 ['Mari', 'Spoken', 'in', 'Russian', 'Federation', ':', 'autonomous',  
 'republics', 'Mari', 'El']  
 ['Mari', 'Total', 'speakers', 'more', 'than', '600', ',', '000']  
 ['Mari', 'Language', 'codes', 'ISO', '639-3', ':', 'mhr', '\xe2', '\x80',  
 '\x93', 'Meadow', 'Mari', '(', 'Eastern', 'Mari', ')']  
 ['Mari', 'Language', 'codes', 'ISO', '639-3', 'mrj', '\xe2', '\x80',  
 '\x93', 'Hill', 'Mari', '(', 'Western', 'Mari', ')']  
 ['Sassarese', 'Total', 'speakers', '\xe2', '\x89', '\x88', '125', '.',  
 '000']  
 ['Sassarese', 'Writing', 'system', 'Latin', 'Alphabet', '(', 'Italian',  
 'variant', ')']  
 ['Hittite', 'Language', 'extinction', 'records', 'cease', 'after', '1200',  
 'BC', ',', 'extinction', 'likely', 'by', '1100', 'BC']  
 ['Istriot', 'Total', 'speakers', '1', ',', '000-2', ',', '000']  
 ['Cebuano', 'Region', '(', 'Central', 'Visayas', 'and', 'Mindanao', ')']  
 ['Huilliche', 'Total', 'speakers', '2', ',', '000']  
 ['Mam', 'Total', 'speakers', '477', ',', '717', '(', '2002', ')']  
 ['Lingua', 'Franca', 'Nova', 'Total', 'speakers', '&', 'gt', ';', '100']

['Lingua', 'Category', '(', 'purpose', ')', 'International', 'auxiliary', 'language']

['Lingua', 'Category', '(', 'purpose', ')', 'Lingua', 'Franca', 'Nova']

['Lingua', 'Setting', 'and', 'usage', 'International', 'auxiliary', 'language']

['Silesian', 'Total', 'speakers', '(', 'about', '60', ',', '000', 'people', 'declared', 'Silesian', 'as', 'their', 'native', 'language', 'in', 'Polish', 'census', ')']

['Silesian', 'Spoken', 'in', 'Silesian', 'Voivodeship', 'and', 'Opole', 'Voivodeship', 'in', 'Poland', ',', 'Moravian-Silesian']

['Silesian', 'Region', 'and', 'Olomouc', '(', 'only', 'Jesen', '\xc3', '\xad', 'k', 'District', ')', 'in', 'Czech', 'Republic', ',', 'Germany', 'and', 'USA']

['Silesian', 'Region', 'Upper', 'Silesia', '/', 'Silesia']

['Cherokee', 'Total', 'speakers', '12', ',', '000', 'to', '22', ',', '000']

['Cherokee', 'Region', 'Oklahoma', 'and', 'the', 'Qualla', 'Boundary', ',', 'North', 'Carolina']

['Yor\xc3\xb9b\xc3\xa1', 'Total', 'speakers', '19', 'million', '(', 'Johnstone', '1993', 'as', 'cited', 'in', 'Ethnologue', ')']

['Baoule', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]

['Baoule', 'Total', 'speakers', '2', ',', '130', ',', '000']

['J\xc3\xa8rriais', 'Total', 'speakers', '2', ',', '874', '(', '113', 'as', 'main', 'language', ';;', '2', ',', '761', 'as', 'secondary', 'language', ')', '(', '2001', 'Census', ')']

['Papiamentu', 'Total', 'speakers', '329', ',', '002']

['Taa', 'Total', 'speakers', '4', ',', '200']

['Taa', 'Region', 'western', 'Southern', 'and', 'western', 'Kweneng', 'districts', 'in', 'Botswana', ';;', 'southern', 'Omaheke', 'and', 'northeastern', 'Hardap', 'regions', 'in', 'Namibia']

['Vedic', 'Sanskrit', 'Language', 'extinction', 'evolved', 'into', 'Classical', 'Sanskrit', 'by', 'the', 'sixth', 'century', 'BC']

['Balinese', 'Total', 'speakers', '3', '.', '9', 'million', '(', 'as', 'of', '2001', ')']

['Buginese', 'Spoken', 'in', 'South', 'Sulawesi', ',', '']

['Buginese', 'Total', 'speakers', '3', '.', '5', 'to', '4', 'million']

['Gayo', 'Total', 'speakers', '550', ',', '000']

['Iban', 'Total', 'speakers', '700', ',', '000']

['Javanese', 'Language', 'codes', 'ISO', '639-3', 'jas', '\xe2', '\x80', '\x93', 'New', 'Caledonian', 'Javanese']

['Javanese', 'Language', 'codes', 'ISO', '639-3', 'osi', '\xe2', '\x80', '\x93', 'Osing', 'language']

['Javanese', 'Language', 'codes', 'ISO', '639-3', 'kaw', '\xe2', '\x80', '\x93', 'Old', 'Javanese']

['Madurese', 'Total', 'speakers', '14', 'million', '(', '1995', ')']

['Makassar', 'Language', 'Total', 'speakers', '1', ',', '600', ',', '000', '(', '1989', ')']

['Makassar', 'Region', 'southern', 'tip', 'of', 'South', 'Sulawesi', ',', 'Sulawesi', '(', 'Celebes', ')']  
 ['Minangkabau', 'Total', 'speakers', '6', ',', '500', ',', '000']  
 ['Minangkabau', 'Region', 'Negeri', 'Sembilan', '(', 'Malaysia', ')']  
 ['Nicaraguan', 'Total', 'signers', '3', ',', '000', '(', 'estimated', 'in', '1997', ';', 'no', 'official', 'census', 'data', 'available', ')']  
 ['Nicaraguan', 'Language', 'codes', 'ISO', '639-3', 'ncs', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Fijian', 'Region', 'Ogea', 'and', 'Vatooa', ';', 'spoken', 'as', 'second', 'language', 'in', 'the', 'rest', 'of', 'Fiji']  
 ['Kashmiri', 'Total', 'speakers', '4', '.', '6', 'million']  
 ['Kashmiri', 'Region', 'Northwestern', 'region', 'of', 'the', 'Indian', 'subcontinent']  
 ['Rundi', 'Total', 'speakers', '9', 'million', '(', '2011', ')']  
 ['Latino', 'Category', '(', 'purpose', ')', 'Latino', 'sine', 'flexione']  
 ['Latino', 'Category', '(', 'sources', ')', 'Completely', 'based', 'on', 'Latin', ',', 'but', 'influenced', 'by', 'ideas', 'in', 'other', 'auxiliary', 'languages']  
 ['Ligurian', '(', 'ancient', ')', 'Language', 'extinction', 'after', '210', 'BC']  
 ['Ligurian', 'Region', 'Northern', 'Mediterranean', 'Coast', 'straddling', 'South-east', 'French', 'and', 'North-east', 'Italian', 'coasts']  
 ['Sango', 'Total', 'speakers', 'an', 'estimated', '404', ',', '000', 'native', 'speakers', ',', '1', '.', '6', 'million', 'second-language', 'speakers', 'as', 'of', '1988']  
 ['Sango', 'Official', 'language', 'in', 'Central', 'African', 'Republic']  
 ['Tongan', 'Total', 'speakers', '200', ',', '000', '(', 'as', 'of', '2010', ')']  
 ['Tsonga', 'Total', 'speakers', '3', ',', '275', ',', '105']  
 ['Zhuang', 'Language', 'codes', 'ISO', '639-3', ':', 'zch', '\xe2', '\x80', '\x93', 'Central', 'Hongshuihe', 'Zhuang']  
 ['Zhuang', 'Language', 'codes', 'ISO', '639-3', 'zeh', '\xe2', '\x80', '\x93', 'Eastern', 'Hongshuihe', 'Zhuang']  
 ['Auslan', 'Region', 'Northern', 'dialect', '(', 'NSW', ',', 'QLD', ')', ',', 'Victorian', 'dialect', 'and', 'South-Eastern', 'dialect', '(', 'all', 'other', 'states', ')']  
 ['Auslan', 'Total', 'signers', '14', ',', '000']  
 ['Auslan', 'Language', 'codes', 'ISO', '639-3', 'asf', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Algonquin', 'Total', 'speakers', '2', ',', '680']  
 ['Algonquin', 'Region', 'Quebec', 'and', 'into', 'Ontario']  
 ['French', 'Total', 'signers', '50', ',', '000', '\xe2', '\x80', '\x93', '100', ',', '000', '(', '1986', ')']  
 ['French', 'Language', 'codes', 'ISO', '639-3', 'fsl', 'sign', 'language',

'\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Dacian', 'Language', 'extinction', 'probably', 'by', 'the', 'sixth', 'century', 'AD']

['Dacian', 'Spoken', 'in', 'and', 'Northern', 'Bulgaria']

['Wenedyk', 'Category', '(', 'sources', ')', '(', 'Romance', 'language', 'based', 'on', 'Polish', ')']

['Wenedyk', 'Setting', 'and', 'usage', 'if', 'Latin', 'had', 'replaced', 'Polish's', 'ancestor']

['Sranan', 'Total', 'speakers', '300', ',', '000']

["Martha's", 'Vineyard', 'Sign', 'Language', 'Language', 'extinction', 'with', 'the', 'death', 'of', 'Katie', 'West', '(', '1952', ')']

["Martha's", 'Region', "Martha's", 'Vineyard', ',', 'Massachusetts']

["Martha's", 'Language', 'codes', 'ISO', '639-3', 'mre', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Avar', 'Total', 'speakers', '788', ',', '960']

['Ruthenian', 'Language', 'extinction', 'developed', 'into', 'Belarusian', ',', 'Ukrainian', 'and', 'Rusyn']

['Ruthenian', 'Spoken', 'in', 'Polish-Lithuanian', 'Commonwealth', ',', '(', 'language', 'of', 'administration', 'of', 'Grand', 'Duchy', 'of', 'Lithuania', 'until', '1699', ')']

['Old', 'Persian', 'Language', 'extinction', 'Ancestor', 'of', 'Middle', 'Persian']

['Middle', 'Dutch', 'Language', 'extinction', 'developed', 'into', 'modern', 'Dutch', 'by', 'the', 'middle', 'of', 'the', '16th', 'century']

['Middle', 'Spoken', 'in', 'the', 'Low', 'Countries']

['Middle', 'High', 'German', 'Language', 'extinction', 'developed', 'into', 'Early', 'New', 'High', 'German', 'from', 'the', '14th', 'century']

['Early', 'New', 'High', 'German', 'Language', 'extinction', 'developed', 'into', 'Modern', 'German', 'from', 'the', '1650s']

["Dena'ina", 'Region', 'Alaska', '(', 'Cook', 'Inlet', 'region', ',', 'Lake', 'Clark', ',', 'Lake', 'Iliamna', ')']

['Ilocano', 'Total', 'speakers', '7', '.', '7', 'million', ',', '2', '.', '3', 'million', '2nd', 'language', '=', '10', 'million', 'total', ';', '3rd', 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']

['Ilocano', 'Writing', 'system', 'Latin', '(', 'Ilocano', 'or', 'Filipino', 'variant', ')', ';']

['Ilocano', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']

['Phoenician', 'Language', 'extinction', 'continued', 'in', 'its', 'Punic', 'form', 'perhaps', 'as', 'late', 'as', '7th', 'century', 'AD']

['Phoenician', 'Spoken', 'in', 'Southern', 'Mediterranean', 'Iberia']

['Zauruan', 'Total', 'speakers', '7', ',', '000']

['Brithenig', 'Setting', 'and', 'usage', 'if', 'Latin', 'had', 'replaced', 'Celtic']

['Zimakani', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Old', 'High', 'German', 'Language', 'extinction', 'developed', 'into',  
 'Middle', 'High', 'German', 'from', 'the', '11th', 'century']  
 ['Old', 'Spoken', 'in', 'Sporadic', 'communities', 'in', 'Eastern', 'Gaul']  
 ['Xavante', 'Total', 'speakers', '\xe2', '\x89', '\x88', '10', ',', '000']  
 ['Marwari', 'Language', 'codes', 'ISO', '639-3', 'rwr', '\xe2', '\x80',  
 '\x93', 'Marwari', '(', 'India', ')']  
 ['Marwari', 'Language', 'codes', 'ISO', '639-3', 'mve', '\xe2', '\x80',  
 '\x93', 'Marwari', '(', 'Pakistan', ')']  
 ['Mozarabic', 'Language', 'extinction', 'by', 'the', 'Late', 'Middle',  
 'Ages']  
 ['Hattic', 'Language', 'extinction', '1100', 'BC', '?']  
 ['Somali', 'Spoken', 'in', 'Somali', 'communities', 'in', 'the', 'Middle',  
 'East', ',', 'Europe', 'and', 'North', 'America']  
 ['Gutnish', 'Total', 'speakers', '~', '500', '?']  
 ['Venda', 'Total', 'speakers', '875', '000']  
 ['Bemba', 'Total', 'speakers', '4', ',', '208', ',', '940']  
 ['Hopi', 'Total', 'speakers', '\xe2', '\x89', '\x88', '5000']  
 ['Kaonde', 'Total', 'speakers', '240', ',', '000']  
 ['Lozi', 'Total', 'speakers', 'around', '500', ',', '000']  
 ['Lozi', 'Region', 'Western', 'Zambia', ',', 'Caprivi']  
 ['Ekajuk', 'Total', 'speakers', '30', ',', '000', '(', '1986', ')']  
 ['Ekajuk', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Nsibidi',  
 'ideographs', ')']  
 ['Western', 'Panjabi', 'Total', 'speakers', '60', ',', '600', ',', '000',  
 'in', 'Pakistan', '(', '2000', ')', ',', '1', ',', '910', ',', '000', 'in',  
 'India', '(', '2007', ')', '., 'Population', 'total', 'all', 'countries',  
 ':', '62', ',', '648', ',', '000']  
 ['Western', 'Language', 'codes', 'ISO', '639-3', 'Punjabi', 'Dialects']  
 ['Lunda', 'Total', 'speakers', '\xce', '\xa3', '400', ',', '000', '(',  
 '2001', ')', '(', '222', ',', '000', 'in', 'Zambia', ',', '178', ',',  
 '000', 'in', 'Angola', ')']  
 ['Masaba', 'Total', 'speakers', '1', ',', '500', ',', '000']  
 ['Yorkshire', 'Dialect', 'Total', 'speakers', 'unknown']  
 ['Mauritian', 'Total', 'speakers', '1', '., '2', 'million']  
 ['Mauritian', 'Language', 'codes', 'Linguasphere', '(', 'to', '51-AAC-cee',  
 ')']  
 ['Tahitian', 'Total', 'speakers', '120', ',', '000']  
 ['Palauan', 'Spoken', 'in', 'Northern', 'Mariana', 'Islands']  
 ['Palauan', 'Total', 'speakers', 'about', '15', ',', '000']  
 ['Franco-Proven\xe7al', 'Arpitan', 'Total', 'speakers', '(', 'est',  
 '., ')', '113', ',', '400']  
 ['Franco-Proven\xe7al', 'Language', 'codes', 'Linguasphere', 'Map',  
 'of', 'the', 'Franco-Proven', '\xc3', '\xa7', 'al', 'Language', 'Area']  
 ['Sicilian', 'Spoken', 'in', 'Calabria', '(', 'Central', 'and', 'Southern',  
 ')']



['Sicilian', 'Spoken', 'in', 'United', 'States', '(', 'New', 'York', 'City', ',', 'Boston', ',', 'Baltimore', ')']

['Sicilian', 'Total', 'speakers', 'est', '.', '8', 'million', '(', 'Gordon', ',', '2005', ')']

['Sicilian', 'Language', 'codes', 'Linguasphere', '51-AAA-re', '&', 'amp', ';', '-rf']

['Sicilian', 'Language', 'codes', 'Linguasphere', '(', 'mainland', '51-AAA-rc', '&', 'amp', ';', '-rd', ')']

['Comorian', 'Total', 'speakers', '350', ',', '702', '(', '2004', ')']

['Maithili', 'Total', 'speakers', '34', ',', '700', ',', '000']

['Maithili', 'Region', 'parts', 'of', 'West', 'Bengal', 'in', 'India']

['Maithili', 'Official', 'language', 'in', 'In', '8th', 'schedule', 'of', 'Constitution', 'of', 'India', ',', 'Bihar', 'state', 'in', 'India', 'and', 'Nepal']

['Maithili', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Haitian', 'Total', 'speakers', '12', ',', '000', ',', '000']

['Haitian', 'Regulated', 'by', 'Minist', '\xc3', '\xa8', 're', 'de', 'l"', '\xc3', '\xa9', 'ducation', 'nationale', 'et', 'de', 'la', 'formation', 'professionnelle']

['Old', 'East', 'Slavic', 'Language', 'extinction', 'developed', 'into', 'the', 'various', 'East', 'Slavic', 'languages']

['Bhojpuri', 'Total', 'speakers', '38', ',', '546', ',', '000']

['Bhojpuri', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Cajun', 'Total', 'speakers', 'over', '250', ',', '000']

['Cajun', 'Language', 'codes', 'Linguasphere', 'French', 'spread', 'in', 'Louisiana', '.', 'Parishes', 'marked', 'in', 'yellow', 'are', 'those', 'where', '4', '\xe2', '\x80', '\x93', '10', '%', 'of', 'the', 'population', 'speak', 'French', 'or', 'Cajun', 'French', 'at', 'home']

['Cajun', 'Language', 'codes', 'Linguasphere', 'orange', '10', '\xe2', '\x80', '\x93', '15', '%']

['Cajun', 'Language', 'codes', 'Linguasphere', 'red', '15', '\xe2', '\x80', '\x93', '20', '%']

['Cajun', 'Language', 'codes', 'Linguasphere', 'brown', '20', '\xe2', '\x80', '\x93', '30', '%']

['Picard', 'Total', 'speakers', '~', '700', '000']

['Picard', 'Official', 'language', 'in', '(', 'official', 'recognition', 'as', 'regional', 'language', 'by', 'Belgium', ')']

['Crimean', 'Gothic', 'Language', 'extinction', 'by', 'the', '18th', 'century', '(', '?', ')']

['Selonian', 'Selian', 'Language', 'extinction', '16th', 'century']

['Thracian', 'Language', 'extinction', 'Fifth', 'century']  
 ['Thracian', 'Spoken', 'in', 'Eastern', 'and', 'Central', 'Balkans', ';']  
 ['Chewa', 'Total', 'speakers', '9', '.', '3', 'million']  
 ['Slavey', 'Total', 'speakers', 'Total', ':', '3', ',', '545']  
 ['Slavey', 'Total', 'speakers', 'North', 'Slavey', ':', '1', ',', '235']  
 ['Slavey', 'Total', 'speakers', 'South', 'Slavey', ':', '2', ',', '310']  
 ['Slavey', 'Language', 'codes', 'ISO', '639-3', ':', 'scs', '\xe2', '\x80', '\x93', 'North', 'Slavey']  
 ['Slavey', 'Language', 'codes', 'ISO', '639-3', 'xsl', '\xe2', '\x80', '\x93', 'South', 'Slavey']  
 ['Wu', 'Total', 'speakers', '~', '90', 'million']  
 ['Wu', 'Region', 'Fujian', 'province', ';', 'North', 'Point']  
 ['Wu', 'Spoken', 'in', 'China', ';', 'and', 'overseas', 'Chinese', 'in', 'urban', 'emigrant', 'communities', 'originating', 'from', 'Wu', 'speaking', 'areas', 'of', 'China', '-', 'particularly', 'United', 'States', '(', 'New', 'York', 'City', ')']  
 ['Tuvan', 'Total', 'speakers', '264', ',', '400']  
 ['Auvergnat', 'Total', 'speakers', '1', '.', '3', 'million']  
 ['Soqotri', 'Total', 'speakers', '64', ',', '000']  
 ['L\x3\xa1adan', 'Category', '(', 'sources', ')', 'a', 'priori', 'language', ',', 'with', 'influences', 'from', 'Navajo', 'and', 'English']  
 ['L\x3\xa1adan', 'Setting', 'and', 'usage', 'experiment', 'in', 'feminist', 'linguistics', ',', 'and', 'featured', 'in', 'Elgin's', 'novel', 'Native', 'Tongue']  
 ['Piedmontese', 'Total', 'speakers', '~', '2', ',', '000', ',', '000']  
 ['Suzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Suzhou', 'Region', 'Suzhou', ',', 'Jiangsu', 'province']  
 ['K\x3\xb6lsch', 'Total', 'speakers', '250', ',', '000']  
 ['Low', 'Language', 'codes', 'ISO', '639-3', 'nds', '\xe2', '\x80', '\x93', 'Low', 'German', '(', 'generic', ')']  
 ['Wyandot', 'Language', 'extinction', 'Spoken', 'until', 'recently', 'near', 'Sandwich']  
 ['Wyandot', 'Language', 'extinction', 'Ontario', 'and', 'Wyandotte']  
 ['Wyandot', 'Language', 'extinction', 'Oklahoma', '.', 'There', 'were', 'two', 'older', 'adult', 'speakers', 'still', 'alive', 'in', '1961']  
 ['Elamite', 'Language', 'extinction', 'by', 'the', 'end', 'of', 'the', '4th', 'century', 'BC']  
 ['Elamite', 'Spoken', 'in', 'Elamite', 'Empire', '(', 'extinct', ')']  
 ['Chiac', 'Region', 'Acadian', 'communities', 'throughout', 'the', 'Maritime', 'provinces']  
 ['Norfolk', 'Total', 'speakers', '580', 'on', 'Norfolk', '(', '1989', ')']  
 ['Norfolk', 'Total', 'speakers', '36', 'on', 'Pitcairn', '(', '2002', ')', ',', 'Australia', ',', 'New', 'Zealand']  
 ['Bundeli', 'Total', 'speakers', '.', '5-1', 'Million', '(', '2001', ')']  
 ['Bagheli', 'Total', 'speakers', '.', '5-1', 'Million', '(', '2001', ')']  
 ['Bagheli', 'Official', 'language', 'in', 'India', 'Madhya', 'Pradesh',

')]  
 ['Khwarezmian', 'Language', 'extinction', 'ca', '.', '13th', 'century']  
 ['Hawaii', 'Total', 'speakers', '600', '000']  
 ['Torres', 'Language', 'codes', 'Linguasphere', '52-ABB-cb', '(',  
 'varieties']  
 ['Torres', 'Language', 'codes', 'Linguasphere', ':', '52-ABB-cba', 'to', '-  
 cbd', ')']  
 ['Zarma', 'Total', 'speakers', '2', '.', '4', 'million']  
 ['Zyrian', 'Total', 'speakers', '293', ',', '000', '(', 'census', '2002',  
 '.', ')']  
 ['Saraiki', 'Total', 'speakers', '13', '.', '9', 'million', 'in',  
 'Pakistan', '(', '1998', 'Population', 'and', 'Housing', 'Census', ',',  
 'Pakistan', ')', ',', ]  
 ['Saraiki', 'Total', 'speakers', '68', ',', '000', 'in', 'India', '(',  
 'Census', 'of', 'India', ',', '2001', ')', '(', 'combined', 'figure',  
 'for', 'persons', 'claiming', 'either', 'the', 'Multani', 'dialect', 'or',  
 'the', 'Bahawalpuri', 'dialect', ')']  
 ['Saraiki', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Shanghainese', 'Region', 'Shanghai', ';', 'also', 'spoken', 'in',  
 'overseas', 'emigrant', 'communities', ',', 'including', 'in', 'United',  
 'States', '(', 'New', 'York', 'City', ')']  
 ['Shanghainese', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Lakota', 'Total', 'speakers', '6', ',', '000']  
 ['Lakota', 'Region', 'Primarily', 'North', 'Dakota', 'and', 'South',  
 'Dakota']  
 ['Komi-Permyak', 'Total', 'speakers', '116', ',', '000']  
 ['Erzya', 'Total', 'speakers', ':', 'in', 'Russia', '440', ',', '000']  
 ['Erzya', 'Total', 'speakers', 'worldwide', '517', ',', '575']  
 ['Evenki', 'Total', 'speakers', '29', ',', '000']  
 ['Shoshoni', 'Total', 'speakers', 'USA', ':', '2905', '(', '2000', ')']  
 ['Balti', 'Total', 'speakers', '337', ',', '000', ',', '(', '450', ',',  
 '000', ')', 'Source', ':', 'Khowar', 'Academy']  
 ['Balti', 'Region', 'Pakistan', '(', 'Kashmir', ',', 'Gilgit-Baltistan',  
 ')', ',', 'Kargil', ',', 'India', 'and', 'small', 'pockets', 'in',  
 'Karachi', ',', 'Rawalpindi', ',', 'Islamabad', '&', 'amp', ';', 'Lahore']  
 ['Balti', 'Writing', 'system', 'Urdu', 'script', 'and', 'modified',  
 'Tibetan', 'script']  
 ['Kalmyk', 'Total', 'speakers', '174', ',', '000']  
 ['Kalmyk', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-BAA-b']  
 ['Beja', 'Total', 'speakers', '1', ',', '178', ',', '000']  
 ['Kabardian', 'Total', 'speakers', '1', ',', '632', ',', '500']  
 ['Blackfoot', 'Total', 'speakers', '5', ',', '000', 'to', '&', 'lt', ';',  
 '8', ',', '000']

['Kongo', 'Language', 'codes', 'ISO', '639-3', 'kwy', '\xe2', '\x80',  
 '\x93', 'San', 'Salvador', 'Kongo']  
 ['Lombard', 'Total', 'speakers', '3', '.', '5', 'million']  
 ['Lombard', 'Language', 'codes', 'Linguasphere', '51-AAA-oc', '&', 'amp',  
 ';', '51-AAA-od']  
 ['Moabite', 'Language', 'extinction', '5th', 'century', 'BC']  
 ['O'odham", 'Total', 'speakers', 'USA', ':', '9595', '(', '2000', ')', ';',  
 'Mexico', ':', '153', '(', '2000', ')']  
 ['Phrygian', 'Language', 'extinction', '5th', 'century']  
 ['Phrygian', 'Spoken', 'in', 'Central', 'Asia', 'Minor']  
 ['Adyghe', 'Total', 'speakers', '500', ',', '000']  
 ['H\xc3\xa4n', 'Writing', 'system', 'Latin', '(', 'Den', '\xc3', '\xa9',  
 'variant', ')']  
 ['Ojibwe', 'language', 'Total', 'speakers', '56', ',', '531', '(', '47',  
 ',', '740', 'in', 'Canada', ';', '8', ',', '791', 'in', 'the', 'United',  
 'States', ')']  
 ['Ojibwe', 'Region', 'British', 'Columbia', ';', 'United', 'States', ':',  
 'Michigan']  
 ['Ojibwe', 'Region', 'groups', 'in', 'North', 'Dakota']  
 ['Ojibwe', 'Writing', 'system', 'Latin', 'alphabet', ',', 'various',  
 'orthographies', 'in', 'Canada', 'and', 'the', 'United', 'States', ';',  
 'Ojibwe', 'syllabics', 'in', 'Canada', ';', 'Great', 'Lakes', 'Algonquian',  
 'syllabary', 'in', 'the', 'United', 'States']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', ':', 'ojs', '\xe2', '\x80',  
 '\x93', 'Severn', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ojg', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ojc', '\xe2', '\x80',  
 '\x93', 'Central', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'obj', '\xe2', '\x80',  
 '\x93', 'Northwestern', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'ISO', '639-3', 'ojw', '\xe2', '\x80',  
 '\x93', 'Western', 'Ojibwa']  
 ['Ojibwe', 'Language', 'codes', 'Linguasphere', '62-ADA-d', '(', 'Ojibwa',  
 '+', 'Anissinapek', ')']  
 ['Ojibwe', 'Language', 'codes', 'Linguasphere', 'Location', 'of', 'all',  
 'Ojibwe', 'Reservations', '/', 'Reserves', 'and', 'cities', 'with', 'an',  
 'Ojibwe', 'population', 'in', 'North', 'America', ',', 'with', 'diffusion',  
 'rings', 'about', 'communities', 'speaking', 'the', 'Ojibwe', 'language']  
 ['Gen', 'Total', 'speakers', '358', ',', '900']  
 ['Stadsfries', 'Total', 'speakers', '45', ',', '000']  
 ['Stadsfries', 'Language', 'codes', 'ISO', '639-3', 'Map', 'of', 'the',  
 'areas', 'in', 'which', 'Stadsfries', 'is', 'spoken']  
 ['West', 'Frisian', 'Total', 'speakers', '360', ',', '000', '\xe2', '\x80',  
 '\x93', '700', ',', '000']  
 ['Saterland', 'Frisian', 'Total', 'speakers', '2', ',', '000']

['North', 'Frisian', 'Total', 'speakers', '10', ',', '000']  
 ['North', 'Language', 'codes', 'Linguasphere', '52-ACA-e', '(',  
 'varieties']  
 ['North', 'Language', 'codes', 'Linguasphere', ':', '52-ACA-eaa', 'to', '-  
 eak', '&', 'amp', ';']  
 ['North', 'Language', 'codes', 'Linguasphere', 'extinct', '-eba', '&',  
 'amp', ';', '-ebb', ')']  
 ['North', 'Language', 'codes', 'Linguasphere', 'North', 'Frisian',  
 'dialects']  
 ['Sireniki', 'Eskimo', 'language', 'Spoken', 'in', 'Russian', 'Federation']  
 ['Sireniki', 'Language', 'extinction', 'Last', 'speaker', 'died', 'in',  
 '1997']  
 ['Sireniki', 'Region', 'Bering', 'Strait', 'region', ',', 'mixed',  
 'populations', 'in', 'settlements', 'Sireniki', 'and', 'Imtuk']  
 ['Soga', 'Total', 'speakers', 'First', 'language', '(', 'native',  
 'speakers', ')', ':', '+', '3', 'million', '(', '2007', 'census', ')']  
 ['Soga', 'Total', 'speakers', 'Second', 'language', ':', '+', '100', '000',  
 '(', '2007', 'census', ')']  
 ['Soga', 'Official', 'language', 'in', 'Major', 'language', 'of', 'Uganda',  
 ',', 'official', 'status', 'unclear', '/', 'pending']  
 ['Soga', 'Regulated', 'by', 'Lusoga', 'Language', 'Authority', '(', 'LULA',  
 ')', '(', 'Uganda', ')']  
 ['Pangasinense', 'Total', 'speakers', '1', '.', '5', 'million', ';', '9th',  
 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']  
 ['Pangasinense', 'Region', 'Ilocos', 'and', 'Central', 'Luzon']  
 ['Pangasinense', 'Writing', 'system', 'Latin', ',', 'Alibata', ',',  
 'Pangasinense', 'or', 'Filipino', 'variant', ')', ';']  
 ['Pangasinense', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Ottoman', 'Turkish', 'Language', 'extinction', 'Reformed', 'into',  
 'Modern', 'Turkish', 'in', '1928']  
 ['Finnish', 'Total', 'signers', '5', ',', '000', '(', '1986', ')']  
 ['Finnish', 'Language', 'codes', 'ISO', '639-3', 'fse', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Bulgar', 'Language', 'extinction', 'between', 'the', '9th', 'century',  
 'on', 'the', 'Danube', 'and', 'the', '14th', 'century', 'on', 'the',  
 'Volga']  
 ['Bulgar', 'Spoken', 'in', 'from', 'Central', 'Asia', 'to', 'the',  
 'steppes', 'North', 'of', 'the', 'Caucasus', ',', 'the', 'Volga', ',',  
 'and', 'the', 'Danube', ',', 'and', 'Southern', 'Italy', '(', 'Molise',  
 ',', 'Campania', ')']  
 ['Limousin', 'Language', 'codes', 'Linguasphere', 'Approximate',  
 'distribution', 'of', 'Limousin', 'within', 'the', 'Occitan', 'area']  
 ['Meitei', 'Total', 'speakers', '1', ',', '500', ',', '000', '+']  
 ['Meitei', 'Total', 'speakers', '1', ',', '466', ',', '705', 'in', 'India',

'(' , '2001' , 'census' , ')']  
 ['Sylheti' , 'Total' , 'speakers' , '10' , ',' , '300' , ',' , '000']  
 ['Kemi' , 'Sami' , 'Language' , 'extinction' , 'ca' , '.' , '1900']  
 ['Papi\xc3\xa1' , 'Total' , 'speakers' , '5' , ',' , '000']  
 ['Niuean' , 'Total' , 'speakers' , '8' , ',' , '000']  
 ['Mi'kmaq' , 'Total' , 'speakers' , '10' , ',' , '850' , '(' , '8' , ',' , '750' ,  
 'in' , 'Canada' , ',' , '2' , ',' , '100' , 'in' , 'the' , 'United' , 'States' , ')']  
 ['Mi'kmaq' , 'Region' , 'Gaspé' , 'Peninsula' , '(' , 'Quebec' , ')']  
 ['Mandaic' , 'Region' , 'Iraq' , '-' , 'Baghdad' , ',' , 'Basra']  
 ['Mandaic' , 'Region' , 'Iran' , '-' , 'Kh' , '\xc5' , '\xab' , 'zest' , '\xc4' ,  
 '\x81' , 'n']  
 ['Mandaic' , 'Language' , 'codes' , 'ISO' , '639-3' , ':' , 'mid' , '\xe2' ,  
 '\x80' , '\x93' , 'Modern' , 'Mandaic']  
 ['Mandaic' , 'Language' , 'codes' , 'ISO' , '639-3' , 'myz' , '\xe2' , '\x80' ,  
 '\x93' , 'Classical' , 'Mandaic']  
 ['Ammonite' , 'Language' , 'extinction' , '5th' , 'century' , 'BC']  
 ['Meroitic' , 'Language' , 'extinction' , '~' , '400' , 'CE']  
 ['Santhali' , 'Total' , 'speakers' , '6' , ',' , '050' , ',' , '000']  
 ['Galician\xe2\x80\x93Portuguese' , 'Language' , 'extinction' , 'Split' ,  
 'and' , 'evolved' , 'into' , 'the' , 'Galician' , 'and' , 'Portuguese' ,  
 'languages']  
 ['Carolinian' , 'Total' , 'speakers' , '5' , ',' , '700']  
 ['Carolinian' , 'Spoken' , 'in' , 'Northern' , 'Mariana' , 'Islands']  
 ['Carolinian' , 'Official' , 'language' , 'in' , 'Northern' , 'Mariana' ,  
 'Islands']  
 ['Dene' , 'Suline' , 'Total' , 'speakers' , '11' , ',' , '895']  
 ['Forro' , 'Total' , 'speakers' , '69' , ',' , '899']  
 ['Angolar' , 'Total' , 'speakers' , '5' , ',' , '000']  
 ['Principense' , 'Total' , 'speakers' , '200' , ',' , '4000']  
 ['Macanese' , 'Total' , 'speakers' , 'Unknown' , ';' , 'perhaps' , 'several' ,  
 'dozen' , 'in' , 'Macau' , 'and' , 'Hong' , 'Kong' , 'and' , 'several' ,  
 'hundreds' , 'or' , 'more' , 'than' , 'a' , 'thousand' , 'among' , 'the' ,  
 'Macanese' , 'Diaspora' , ';' , 'virtually' , 'all' , 'speakers' , 'at' , 'least' ,  
 'bilingual']  
 ['Gwich\xe2\x80\x99in' , 'Total' , 'speakers' , '770' , '(' , '1998' ,  
 'Statistics' , 'Canada' , ')']  
 ['Gwich\xe2\x80\x99in' , 'Writing' , 'system' , 'Latin' , '(' , 'Northern' ,  
 'Athabaskan' , 'variant' , ')']  
 ['Penang' , 'Hokkien' , 'Total' , 'speakers' , '1' , ',' , '824' , ',' , '741']  
 ['Amorite' , 'Language' , 'extinction' , '2nd' , 'millennium' , 'BC']  
 ['Dogrib' , 'Total' , 'speakers' , '2' , ',' , '640']  
 ['Lepontic' , 'Language' , 'extinction' , 'ca' , '.' , '400' , 'BC' , '?']  
 ['Messapian' , 'Messapic' , 'Language' , 'extinction' , 'around' , '1st' ,  
 'century' , 'BC']  
 ['Hurrian' , 'Language' , 'extinction' , 'Ca' , '1000' , 'BC']  
 ['Ongota' , 'Total' , 'speakers' , '6' , '(' , '2008' , ')']

['Ongota', 'Region', 'Southern', 'Omo', 'Zone', ',', 'Southern']  
 ['Menominee', 'Regulated', 'by', 'Menominee', 'Language', '&', 'amp', ';',  
 'Culture', 'Commission']  
 ['Urartian', 'Language', 'extinction', 'c', '.', '6th', 'century', 'BCE']  
 ['Inuvialuktun', 'Total', 'speakers', '400', '\xe2', '\x80', '\x93', '700']  
 ['Inuvialuktun', 'Regulated', 'by', 'Inuvialuit', 'Cultural', 'Resource',  
 'Centre', 'and', 'Inuit', 'Tapiriit', 'Kanatami']  
 ['Semigallian', 'Language', 'extinction', '16th', 'century']  
 ['Extremaduran', 'Total', 'speakers', '200', ',', '000']  
 ['Lezgian', 'Total', 'speakers', '783', ',', '720']  
 ['Lushootseed', 'Region', 'Southern', 'British', 'Columbia', 'into',  
 'northern', 'Washington']  
 ['Punic', 'Spoken', 'in', 'Formerly', 'spoken', 'in', 'North', 'Africa',  
 'Tunisia', 'Malta']  
 ['N'Ko', 'Total', 'speakers', 'c', '.', '27', 'million', '(', 'mother',  
 'tongue', 'and', 'second', 'language', ')', ';', 'see', 'Bambara', ',',  
 'Mandinka', ',', 'Maninka', ',', 'and', 'Dioula']  
 ['N'Ko', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']  
 ['N'Ko', 'Regulated', 'by', 'Kurukan', 'fuwa', 'gbara', '(', 'Association',  
 'Kurukan', 'fuwa', 'gbara', '-', ')']  
 ['Fang', 'Total', 'speakers', '858', ',', '000']  
 ['Swiss', 'Total', 'speakers', '1', '.', '5', 'million']  
 ['Swiss', 'Official', 'language', 'in', 'Switzerland', '(', 'as', 'French',  
 ')']  
 ['Mingrelian', 'Total', 'speakers', '500', ',', '000', '(', '1989', ')']  
 ['Moksha', 'Total', 'speakers', '~', '500', ',', '000']  
 ['Svan', 'Total', 'speakers', '15', ',', '000']  
 ['Svan', 'Writing', 'system', 'Seldom', '(', 'if', 'ever', ')', 'written',  
 ',', 'but', 'Georgian', 'alphabet', 'is', 'used']  
 ['Me\xc3\xa4nkieli', 'Total', 'speakers', '40', ',', '000-70', ',', '000']  
 ['Han'er', 'Language', 'extinction', 'Evolved', 'into', 'Chinese',  
 'dialects']  
 ['Arawak', 'Total', 'speakers', '2', ',', '500']  
 ['Khariboli', 'Region', 'Delhi', ',', 'Western', 'Uttar', 'Pradesh']  
 ['Khariboli', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Philippine', 'Creole', 'Spanish', 'Total', 'speakers', '2', ',', '500',  
 ',', '000', ';', '7th', 'most', 'spoken', 'native', 'language', 'in',  
 'the', 'Philippines']  
 ['Philippine', 'Region', 'Filipino', 'diaspora', 'and', 'other', 'regions',  
 'with', 'Chavacano', 'communities']  
 ['Massachusett', 'Language', 'extinction', 'Late', '19th', 'century', ',',  
 'revival', 'efforts', 'underway']  
 ['Massachusett', 'Region', 'Wampanoag', 'people', ',', 'Massachusett',

'people']  
 ['Greenlandic', 'Total', 'speakers', 'approximately', '57', ',', '000']  
 ['Greenlandic', 'Regulated', 'by', 'Oqaasileriffik', '(', 'The',  
 'Greenland', 'Language', 'Secretariat', ')']  
 ['Seychellois', 'Total', 'speakers', '72', ',', '700']  
 ['Seychellois', 'Language', 'codes', 'Linguasphere', '51-AAC-cea', '(',  
 '&', 'amp', ';', '-ceb', ')']  
 ['Venetic', 'Language', 'extinction', '1st', 'century', 'AD']  
 ['Bouyei', 'Total', 'speakers', '2', ',', '649', ',', '205', '(', '2000',  
 ')']  
 ['Inuktitut', 'Spoken', 'in', 'Canada', '(', 'Nunavut', ',', 'Quebec', '(',  
 'Nunavik', ')', ',', 'Northwest', 'Territories', ',', 'Newfoundland',  
 'and', 'Labrador', '(', 'Nunatsiavut', ')', ')']  
 ['Inuktitut', 'Regulated', 'by', 'Inuit', 'Tapiriit', 'Kanatami', 'and',  
 'various', 'other', 'local', 'institutions']  
 ['Inuktitut', 'Language', 'codes', 'ISO', '639-3', ':', 'ike', '\xe2',  
 '\x80', '\x93', 'Eastern', 'Canadian', 'Inuktitut']  
 ['Hiligaynon', 'Total', 'speakers', '11', 'million', 'total', '(', 'first',  
 'language', ':', '7', 'million', ',', 'second', 'language', ':', '4',  
 'million', '(', 'est', '.', ')', ')', '3rd', 'most', 'spoken', 'native',  
 'language', 'in', 'the', 'Philippines']  
 ['Hiligaynon', 'Region', 'Western', 'Visayas', 'and', 'Mindanao']  
 ['Hiligaynon', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant',  
 ')', ';']  
 ['Hiligaynon', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Sedang', 'Total', 'speakers', '101', ',', '434']  
 ['Kapampangan', 'Total', 'speakers', '2', '.', '9', 'million', ';', '6th',  
 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']  
 ['Kapampangan', 'Writing', 'system', 'Latin', '(', 'Kapampangan', 'or',  
 'Spanish', 'variant', ')', ';']  
 ['Kapampangan', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Gallo', 'Total', 'speakers', '28', ',', '000']  
 ['Gallo', 'Language', 'codes', 'Linguasphere', 'The', 'Gallo', 'language',  
 'area', '(', 'in', 'grey', ')', 'corresponding', 'to', 'Upper', 'Brittany',  
 '(', 'in', 'colours', ',', 'Lower', 'Brittany', ')']  
 ['Middle', 'French', 'Language', 'extinction', 'evolved', 'into', 'Modern',  
 'French', 'by', 'the', 'early', '17th', 'century']  
 ['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', ':', 'yhd', '\xe2',  
 '\x80', '\x93', 'Judeo-Iraqi', 'Arabic']  
 ['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'aju', '\xe2',  
 '\x80', '\x93', 'Judeo-Moroccan', 'Arabic']  
 ['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'yud', '\xe2',  
 '\x80', '\x93', 'Judeo-Tripolitanian', 'Arabic']  
 ['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'ajt', '\xe2',



'\x80', '\x93', 'Judeo-Tunisian', 'Arabic']  
 ['Judeo-Arabic', 'Language', 'codes', 'ISO', '639-3', 'jye', '\xe2',  
 '\x80', '\x93', 'Judeo-Yemeni', 'Arabic']  
 ['Guern\x3\xa9siais', 'Total', 'speakers', '1', ',', '327', '(', '2', '%',  
 ')', 'fluent']  
 ['Elfdalian\x3\x96vdalian', 'Total', 'speakers', 'Around', '3000']  
 ['Elfdalian\x3\x96vdalian', 'Region', 'Elfdalia', '(', 'Dalecarlia',  
 'County', ')', '(', 'Northern', 'Europe', ')']  
 ['Elfdalian\x3\x96vdalian', 'Writing', 'system', 'Latin', 'Alphabet',  
 '(', 'Elfdalian', 'variant', ')', ',', 'Historically', 'Dalecarlian',  
 'runes']  
 ['Bats', 'Total', 'speakers', '3', ',', '420', '(', '2000', 'WCD', ')']  
 ['Andalusian', 'Arabic', 'Language', 'extinction', '17th', 'century']  
 ['Andalusian', 'Language', 'extinction', '(', 'after', 'the', 'expulsion',  
 'from', 'Spain', 'and', 'Portugal', ')']  
 ['Egyptian', 'Arabic', 'Total', 'speakers', '54', ',', '000', ',', '000',  
 '(', 'native', 'speakers', 'only', ')']  
 ['Yemeni', 'Language', 'codes', 'ISO', '639-3', ':', 'ayh', '\xe2', '\x80',  
 '\x93', 'Hadrami', 'Arabic']  
 ['Yemeni', 'Language', 'codes', 'ISO', '639-3', 'ayn', '\xe2', '\x80',  
 '\x93', 'Sanaani', 'Arabic']  
 ['Yemeni', 'Language', 'codes', 'ISO', '639-3', 'acq', '\xe2', '\x80',  
 '\x93', 'Ta'izzi-Adeni', 'Arabic']  
 ['Washo', 'Total', 'speakers', 'USA', ':', '252', '(', '2000', ')']  
 ['Washo', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact',  
 'distribution', 'of', 'the', 'Washo', 'language']  
 ['Levantine', 'Total', 'speakers', '35', ',', '000', ',', '000']  
 ['Levantine', 'Language', 'codes', 'ISO', '639-3', ':', 'apc', '\xe2',  
 '\x80', '\x93', 'North', 'Levantine', 'Arabic']  
 ['Levantine', 'Language', 'codes', 'ISO', '639-3', 'ajp', '\xe2', '\x80',  
 '\x93', 'South', 'Levantine', 'Arabic']  
 ['(Eastern)', 'Total', 'speakers', 'less', 'than', '60']  
 ['\xc7\x80Xam', 'Language', 'extinction', '19', '?', '?']  
 ['Seroa', 'Language', 'extinction', '19', '?', '?']  
 ['Ju\x3\x80\xca\xbcchoan', 'Total', 'speakers', '35', ',', '000']  
 ['K'iche'', 'Total', 'speakers', 'approx', '.', '1', ',', '000', ',', '000']  
 ['Jiwarli', 'Language', 'extinction', 'April', '1986', 'with', 'the',  
 'death', 'of', 'Jack', 'Butler']  
 ['\xc7\x83\xca\xbc0\xc7\x83Kung', 'Total', 'speakers', '6', ',', '000']  
 ['\xc7\x82Kx\xca\xbcbauc\x7\x81\xca\xbccein', 'Total', 'speakers', '7', ',', '000']  
 ['Naro', 'Total', 'speakers', '14', ',', '000']  
 ['Kxoe', 'Total', 'speakers', '11', ',', '000']  
 ['A'tong', 'Total', 'speakers', '15', ',', '000']  
 ['A'tong', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',

'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']

['Aari', 'Total', 'speakers', '285', ',', ',', '473']

['Abanyom', 'Total', 'speakers', '12', ',', ',', '500']

['Abau', 'Total', 'speakers', '4', ',', ',', '500', '\xe2', '\x80', '\x93', '5',  
 ',', ',', '000']

['Abau', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Oirat', 'Spoken', 'in', "People's", 'Republic', 'of', 'China']

['Oirat', 'Writing', 'system', 'Clear', 'script', '(', 'China', ':',  
 'unofficial', ')', ',', ',', 'Mongolian', 'script', '(', 'Mongolia', ':',  
 'unofficial', ',', ',', 'China', ':', 'official', ')', ',', ',', 'Cyrillic', '(',  
 'Russia', ':', 'official', ',', ',', 'Mongolia', ':', 'official', ')']

['Oirat', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-BAA-b']

['Salar', 'Total', 'speakers', '70', ',', ',', '000']

['Western', 'Yugur', 'Total', 'speakers', '4', ',', ',', '600']

['Kuman', 'Total', 'speakers', '80', ',', ',', '000', '(', '1994', ')']

['Kuman', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Achang', 'Total', 'speakers', '63', ',', ',', '000']

['Kurama', 'Total', 'speakers', 'c', '.', '40', ',', ',', '000', '(', '2000',  
 'est', '.', ')']

['Sungor', 'Total', 'speakers', 'approximately', '38', ',', ',', '000', '(',  
 '1993', ')']

['Kurrama', 'Total', 'speakers', 'Possibly', '20']

['Kurrama', 'Region', 'Pilbara', 'region', 'of', 'Western', 'Australia',  
 ',', ',', 'particularly', 'around', 'the', 'Fortescue', 'valley', 'east', 'of',  
 'Millstream', 'and', 'south', 'of', 'the', 'Fortescue', 'River', 'to',  
 'Rocklea', 'Station']

['Taiwanese', 'Mandarin', 'Total', 'speakers', 'About', '19', '.', '3',  
 'million', 'in', 'Taiwan', 'Area', 'of', 'the', 'ROC']

['Canadian', 'Total', 'speakers', 'Canada', '148', ',', ',', '090', '-', '175',  
 ',', ',', '000']

['Lingua', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']

['Lingua', 'Category', '(', 'purpose', ')', 'Lingua', 'sistemfrater']

['Lingua', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']

['Chimariko', 'Language', 'extinction', 'ca', '.', '1910']

['Polish', 'Language', 'codes', 'ISO', '639-3', 'pso', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']

['Yinggarda', 'Total', 'speakers', 'Possibly', '1']

['Yinggarda', 'Region', 'Gascoyne', 'coast', 'area', 'of', 'Western',  
 'Australia']

['Mentawai', 'Total', 'speakers', '64', ',', ',', '000']

['Mentawai', 'Region', 'Mentawai', 'Islands', ',', ',', 'West', 'Sumatra', '(',

'Indonesia', ')']  
 ['Chuj', 'Total', 'speakers', 'approx', '.', '50', ',', '000']  
 ['Chuj', 'Region', 'Northern', 'Huehuetenango', '(', 'And', 'Chiapas', ',', 'Mexico', ')']  
 ['Solano', 'Language', 'extinction', '18th', 'century']  
 ['Solano', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'Solano', 'language']  
 ['Timucua', 'Language', 'extinction', 'second', 'half', '18th', 'century']  
 ['Timucua', 'Writing', 'system', 'No', 'native', 'writing', 'system', ';', 'language', 'recorded', 'in', 'Latin', 'alphabet']  
 ['Timucua', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Timucua', 'language']  
 ['Timucua', 'Language', 'codes', 'ISO', '639-3', 'The', 'Tawasa', 'dialect']  
 ['Tamahaq', 'Total', 'speakers', '62', ',', '000']  
 ['Wappo', 'Language', 'extinction', 'Laura', 'Fish', 'Somersal', '(', 'd', '.', '1990', ')', 'last', 'fluent', 'speaker']  
 ['Wappo', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Wappo', 'language']  
 ['Iwaidja', 'Spoken', 'in', 'Croker', 'Island', ',', 'Northern', 'Territory']  
 ['Yana', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact', 'distribution', 'of', 'the', 'Yana', 'language']  
 ['M\u03a1ku', 'Language', 'extinction', '2000', '\u00e2', '\u0080', '\u0093', '2002']  
 ['Kitanemuk', 'Language', 'extinction', 'Last', 'spoken', 'in', 'the', '1940s', 'by', 'Marcelino', 'Rivera']  
 ['Kitanemuk', 'Language', 'extinction', 'Isabella', 'Gonzales']  
 ['Kitanemuk', 'Language', 'extinction', 'and', 'Refugia', 'Duran']  
 ['Papi\u03a1', 'Tugu', 'Language', 'extinction', '1970s']  
 ['Portugis', 'Language', 'extinction', 'mid-20th', 'century']  
 ['Portugis', 'Language', 'codes', 'Linguasphere', '51-AAC-ahg', '&', 'amp', ';', '-ahh']  
 ['Bidau', 'Creole', 'Portuguese', 'Language', 'extinction', '1960s']  
 ['Dutch', 'Total', 'signers', '20', ',', '000', '(', '1986', ')']  
 ['Dutch', 'Language', 'codes', 'ISO', '639-3', 'dse', 'sign', 'language', '\u00e2', '\u0080', '\u0094', 'list', 'of', 'sign', 'languages', '\u00e2', '\u0080', '\u0094', 'legal', 'recognition']  
 ['Aklanon', 'Total', 'speakers', '394', ',', '545']  
 ['Aklanon', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')', ';']  
 ['Bodo', 'Total', 'speakers', '1', ',', '350', ',', '478', '+']  
 ['Bodo', 'Total', 'speakers', '1', ',', '350', ',', '478', 'in', 'India', '(', '2001', ')', ',', '3', ',', '301', 'in', 'Nepal', '(', '2001', ')']  
 ['Kven', 'Total', 'speakers', '2000', '-', '8000', '(', '2005', ')']  
 ['Kven', 'Official', 'language', 'in', '(', 'Minority', 'language', ')']

['Kven', 'Regulated', 'by', 'Kven', 'language', 'board']  
 ['Ifugao', 'Total', 'speakers', '43', ',', '000']  
 ['Bontoc', 'Total', 'speakers', '40', ',', '000']  
 ['Karkin', 'Language', 'extinction', '19th', 'century']  
 ['Plains', 'Language', 'codes', 'ISO', '639-3', 'psd', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Quzhou', 'Region', 'Quzhou', 'prefecture', ',', 'Zhejiang', 'province']  
 ['Quzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Italiano', 'Language', 'codes', 'Linguasphere', '51-AAA-ra', '.', '.',  
 '.', '-rba', 'Central', 'italian']  
 ['Alur', 'Total', 'speakers', '1', '.', '2', 'million']  
 ['Saanich', 'Region', 'British', 'Columbia', ',', 'Washington']  
 ['Puquina', 'Language', 'extinction', '18th', 'century']  
 ['Nakanai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Nakanai', 'Total', 'speakers', '13', ',', '000']  
 ['Me'en', 'Total', 'speakers', '145', ',', '333', 'of', 'Me'en', 'dialect']  
 ['Chaha', 'Total', 'speakers', '440', ',', '000']  
 ['Semai', 'Total', 'speakers', '33', ',', '000']  
 ['Libido', 'Total', 'speakers', '58', ',', '788', '(', '2007', 'census',  
 ')']  
 ['Maria', 'Spoken', 'in', 'Central', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Maria', 'Total', 'speakers', '1', ',', '350', '(', '2000', ')']  
 ['Seto', 'Total', 'speakers', '5', ',', '000']  
 ['Takelma', 'Language', 'extinction', '19th', 'century']  
 ['Takelma', 'Language', 'codes', 'ISO', '639-3', 'Takelma', '(', 'south',  
 ')', ',', 'with', 'the', 'Kalapuyan', 'languages', 'to', 'the', 'north']  
 ['West', 'Central', 'K'iche'', '/', 'Quich', '\xc3', '\xa9', 'Total',  
 'speakers', '250', ',', '000']  
 ['Ottawa', 'Language', 'codes', 'Linguasphere', '62-ADA-dd', '(', 'Odawa',  
 ')']  
 ['Ottawa', 'Language', 'codes', 'Linguasphere', 'Ottawa', 'population',  
 'areas', 'in', 'Ontario', ',', 'Michigan', 'and', 'Oklahoma', '.',  
 'Reserves', '/', 'Reservations', 'and', 'communities', 'shown', 'in',  
 'red']  
 ['Israeli', 'Total', 'signers', '5', ',', '000', '(', '1986', ')']  
 ['Israeli', 'Language', 'codes', 'ISO', '639-3', 'isr', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Tayart', 'Tamajeq', 'Total', 'speakers', '250', ',', '000', '(', '1998',  
 ')']  
 ['Buyeo', 'Language', 'extinction', '7th', 'century', '?']  
 ['Buyeo', 'Language', 'codes', 'ISO', '639-3', 'The', 'Three', 'Kingdoms',  
 'of', 'Korea', ',', 'with', 'Goguryeo', 'and', 'Buyeo', 'in', 'blue']  
 ['Philistine', 'Language', 'extinction', 'ca', '.', '9th', 'century']

'BCE']  
 ['Z\xc3\xa1paro', 'Total', 'speakers', '1', '(', 'SIL', ',', '2000', ')']  
 ['New', 'High', 'German', 'Total', 'speakers', 'Native', 'speakers', ':', '95', 'to', '100', 'million']  
 ['New', 'Region', 'Central', 'Europe', ',', 'Western', 'Europe']  
 ['New', 'Writing', 'system', 'Latin', 'alphabet', '(', 'German', 'variant', ')']  
 ['New', 'Official', 'language', 'in', 'co-official', 'language', 'of', 'Namibia', 'until', '1990']  
 ['Seediq', 'Total', 'speakers', '4', ',', '750']  
 ['Hunde', 'Total', 'speakers', '200', ',', '000']  
 ['Wirangu', 'Total', 'speakers', '&', 'lt', ';', '10']  
 ['Wirangu', 'Region', 'West', 'coast', 'of', 'South', 'Australia']  
 ['Krio', 'Dayak', 'Total', 'speakers', '500']  
 ['P\xc3\xaalez', 'Total', 'speakers', '&', 'gt', ';', '71', ',', '400']  
 ['Koti', 'Total', 'speakers', '64', ',', '200']  
 ['Koti', 'Region', 'Koti', 'Island', 'and', 'Angoche', ',', 'Nampula', 'Province']  
 ['Sonjo', 'Total', 'speakers', '30', ',', '000', '(', '2002', 'SIL', ')']  
 ['Sonjo', 'Region', 'near', 'the', 'Kenyan', 'border']  
 ['Mbo', 'Total', 'speakers', '45', ',', '000', '(', '1995', ')']  
 ['Hwana', 'Total', 'speakers', '32', ',', '000', '(', '1992', ')']  
 ['Azari', 'Language', 'extinction', 'gradually', '1100-1600AD']  
 ['Azari', 'Region', 'Middle', 'East', ',', 'Central', 'Asia']  
 ['Cueva', 'Language', 'extinction', '16th', 'century']  
 ['Cueva', 'Region', 'Dari', '\xc3', '\xa9', 'n', 'Ethnicity']  
 ['Colonial', 'French', 'Language', 'extinction', '"', 'virtually', '"', 'extinct', ';', 'melded', 'with', 'Cajun', 'French']  
 ['Abinomn', 'Total', 'speakers', '~', '300', '(', '1999', ')']  
 ['Isirawa', 'Total', 'speakers', '2', ',', '000']  
 ['Karkar-Yuri', 'Total', 'speakers', '1', ',', '142', '(', '1994', 'SIL', ')']  
 ['Karkar-Yuri', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Karkar-Yuri', 'Region', '3', '\xc2', '\xb0', '44', '\xe2', '\x80', '\xb2', 'S', '141', '\xc2', '\xb0', '5', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '3', '.', '733', '\xc2', '\xb0', 'S', '141', '.', '083', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-3', '.', '733', ';', '141', '.', '083']  
 ['Yal\xc3\xab', 'Nagatman', 'Total', 'speakers', '600']  
 ['Yal\xc3\xab', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Busa', '(', 'Odiai', ')', 'Total', 'speakers', '244', '(', '2000', 'census', ')']  
 ['Busa', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Busa', 'Region', '3', '\xc2', '\xb0', '49', '\xe2', '\x80', '\xb2', 'S', '141', '\xc2', '\xb0', '20', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '3', '.', '817', '\xc2', '\xb0', 'S',

'141', '.', '333', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-3',  
 '.', '817', ';', '141', '.', '333']  
 ['Busa', 'Region', 'north', 'of', 'Upper', 'Sepik', 'River']  
 ['Busa', 'Region', 'Abau', 'is', 'south', 'and', 'west']  
 ['Sulka', 'Total', 'speakers', '3', ',', '000']  
 ['Kol', 'Total', 'speakers', '4', ',', '000', '(', '1991', 'SIL', ')', '.',  
 'Population', 'includes', '1', ',', '300', 'Sui', ',', 'Kol', '(', '1987',  
 ')']  
 ['Kuat', 'Kuat', ',', 'Panaris', 'Total', 'speakers', '2000', '\xe2',  
 '\x80', '\x93', '3000']  
 ['Kuat', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Porome', 'Total', 'speakers', '1', ',', '100', '(', '1977', 'SIL', ')']  
 ['Porome', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Porome', 'Region', 'southern', 'Papua', 'New', 'Guinea', 'Coordinates']  
 ['Porome', 'Region', '7', '\xc2', '\xb0', '27', '\xe2', '\x80', '\xb2',  
 'S', '144', '\xc2', '\xb0', '17', '\xe2', '\x80', '\xb2', 'E', '\xef',  
 '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '7', '.', '45', '\xc2',  
 '\xb0', 'S', '144', '.', '283', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
 '\xbf', '/', '-7', '.', '45', ';', '144', '.', '283']  
 ['Riverense', 'Spoken', 'in', 'Northern', 'Uruguay', ',', 'near',  
 'Brazilian', 'border']  
 ['Spokil', 'Date', 'founded', '1887', 'or', '1890']  
 ['Nkore', 'Total', 'speakers', '2', ',', '330', ',', '000']  
 ['An\xc3\xaam', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Ata', 'Total', 'speakers', '1900', '(', '1991', 'SIL', ')']  
 ['Ata', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Y\xc3\xa9l\xc3\xae', 'Dnye', 'Total', 'speakers', '3750', '(', '1998',  
 ')']  
 ['Y\xc3\xa9l\xc3\xae', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Gitxsan', 'Total', 'speakers', '1', ',', '000']  
 ['Gitxsan', 'Region', 'Skeena', 'region', ',', 'British', 'Columbia']  
 ['Bora', 'Total', 'speakers', '2', ',', '828', '(', '2000', 'WCD', ')']  
 ['Tacana', 'Total', 'speakers', '1', ',', '800']  
 ['Phong', 'Total', 'speakers', '1000', '(', 'in', '1981', ')']  
 ['Bahrani', 'Spoken', 'in', 'Bahrain', ',', 'Oman', 'and', 'Saudi',  
 'Arabia']  
 ['Bahrani', 'Total', 'speakers', 'about', '310', ',', '000']  
 ['Upper', 'Kuskokwim', ',', 'Kolchan', 'Total', 'speakers', '40']  
 ['Upper', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan',  
 'variant', ')']  
 ['Tolowa', 'Total', 'speakers', '4', 'in', '1994']  
 ['Tsetsaut', 'Language', 'extinction', 'mid', '20th', 'century']  
 ['Yuki', 'Language', 'extinction', '20th', 'century']  
 ['Konkow', 'Total', 'speakers', '3', 'to', '6']  
 ['Wambule', 'Total', 'speakers', '5', ',', '000']  
 ['Culina', 'Total', 'speakers', '1', ',', '303', '(', '2000', 'WCD', ',',

'2002', 'Boyer', ')']  
 ['Mysian', 'Language', 'extinction', '1st', 'c', '.', 'BC']  
 ['South', 'Picene', 'Language', 'extinction', '6th', 'to', '4th',  
 'century', 'BC']  
 ['South', 'Language', 'codes', 'ISO', '639-3', 'spx', 'Approximate',  
 'distribution', 'of', 'languages', 'in', 'Iron', 'Age', 'Italy', 'during',  
 'the', '6th', 'century', 'BC']  
 ['Naba', 'Total', 'speakers', '232', ',', '000', '(', '1993', 'census',  
 ')']  
 ['Cof\xc3\xa1n', 'Total', 'speakers', '1', ',', '000-1', ',', '600']  
 ['Cof\xc3\xa1n', 'Region', 'Oriente', 'or', 'Ecuadorian', 'Amazon']  
 ['Uab', 'Meto', 'Total', 'speakers', '586', ',', '000', '(', '1997', ')']  
 ['Uab', 'Spoken', 'in', 'West', 'Timor', ',', 'Indonesia']  
 ['Oroqen', 'Total', 'speakers', '1', ',', '200']  
 ['Makalero', 'Total', 'speakers', '70', ',', '000', '(', '1989', ')', '.',  
 'Ethnic', 'population', ':', '70', ',', '000']  
 ['Makalero', 'Region', '8', '\xc2', '\xb0', '39', '\xe2', '\x80', '\xb2',  
 'S', '126', '\xc2', '\xb0', '30', '\xe2', '\x80', '\xb2', 'E', '\xef',  
 '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '8', '.', '65', '\xc2',  
 '\xb0', 'S', '126', '.', '5', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf',  
 '/', '-8', '.', '65', ';', '126', '.', '5']  
 ["Babine\xe2\x80\x93Witsuwit'en", 'Total', 'speakers', '500', '(', '1997',  
 'S', '.', 'Hargus', ')']  
 ['Sichuanese', 'Total', 'speakers', '~', '100', 'million']  
 ['Sichuanese', 'Language', 'codes', 'ISO', '639-3', 'Sichuanese', 'in',  
 'China']  
 ['Talysh', 'Region', 'The', 'Western', 'and', 'Southwestern', 'Caspian',  
 'Sea', 'coastal', 'strip']  
 ['Talysh', 'Regulated', 'by', 'Academy', 'of', 'Persian', 'Language',  
 'and', 'Literature']  
 ['Arwi', 'Language', 'extinction', 'Early', '19th', 'Century']  
 ['Arwi', 'Spoken', 'in', 'India', ',', 'Sri', 'Lanka']  
 ['Kissi', 'Total', 'speakers', '~', '500', ',', '000']  
 ['Kissi', 'Language', 'codes', 'ISO', '639-3', ':', 'kss', '\xe2', '\x80',  
 '\x93', 'Southern', 'Kissi']  
 ['Kissi', 'Language', 'codes', 'ISO', '639-3', 'kqs', '\xe2', '\x80',  
 '\x93', 'Northern', 'Kissi']  
 ['Kono', 'Total', 'speakers', '190', ',', '000']  
 ['Tocodede', 'Total', 'speakers', '63', ',', '000', '(', '1981', 'Wurm',  
 'and', 'Hattori', ')']  
 ['T\xc3\xbcbatulabal', 'Total', 'speakers', 'Possibly', '6']  
 ['Cham', 'Total', 'speakers', '323', ',', '100', '(', 'Ethnologue', ',',  
 '2002', ')']  
 ['Cham', 'Spoken', 'in', 'China', '(', 'Hainan', 'Island', ')']  
 ['Cham', 'Writing', 'system', 'Cham', 'alphabet', '(', 'Vietnam', ')', ',',  
 'Arabic', 'alphabet', '(', 'Cambodia', ')']

['Cham', 'Language', 'codes', 'ISO', '639-3', ':', 'cja', '\xe2', '\x80',  
 '\x93', 'Western', 'Cham']  
 ['Cham', 'Language', 'codes', 'ISO', '639-3', 'cjm', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Cham']  
 ['Etchemin', 'Language', 'extinction', '18th', 'century']  
 ['Zapotec', 'Total', 'speakers', 'ca', '500', ',', '000']  
 ['Zapotec', 'Language', 'codes', 'ISO', '639-3', 'zap', '(', 'see', 'List',  
 'of', 'Zapotec', 'languages', ')']  
 ['Bunak', 'Total', 'speakers', '100', ',', '000', '(', '1977', ')']  
 ['Bunak', 'Spoken', 'in', 'Indonesia', ',', 'East', 'Timor']  
 ['Kemak', 'Total', 'speakers', '100', ',', '000', '(', '1981', 'Wurm',  
 'and', 'Hattori', ')']  
 ['Chiwere', 'Total', 'speakers', 'Uncertain', 'but', 'fewer', 'than', '40']  
 ['Tagol', 'Murut', 'Total', 'speakers', '15', ',', '000']  
 ['Yavapai', 'Spoken', 'in', '1420', 'Yavapai', 'people']  
 ['Galoli', 'Total', 'speakers', '50', ',', '000']  
 ['Galoli', 'Spoken', 'in', 'Manatuto', 'District', ',', 'East', 'Timor']  
 ['Wetarese', 'Language', 'codes', 'ISO', '639-3', 'ilu', '\xe2', '\x80',  
 '\x93', 'Ili'uun']  
 ['Quichua', 'Total', 'speakers', '2', ',', '500', ',', '000']  
 ['Sidetic', 'Language', 'extinction', 'Third', 'century', 'BCE']  
 ['Sidetic', 'Spoken', 'in', 'Ancient', 'southwestern', 'Anatolia']  
 ['Pisidian', 'Language', 'extinction', 'Early', 'CE', '?']  
 ['Pisidian', 'Spoken', 'in', 'Ancient', 'southwestern', 'Anatolia']  
 ['Asturian', 'Total', 'speakers', '150', ',', '000', '\xe2', '\x80',  
 '\x93', '400', ',', '000']  
 ['Asturian', 'Regulated', 'by', 'Academy', 'of', 'the', 'Asturian',  
 'Language', '(', 'Asturian', ')']  
 ['Bezhta', 'Total', 'speakers', '6', ',', '200', '(', '2002', ')']  
 ['Kuanua', 'Total', 'speakers', '~', '100', ',', '000']  
 ['Kuanua', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kuanua', 'Region', 'Gazelle', 'Peninsula', ',', 'East', 'New', 'Britain',  
 'Province']  
 ['Minigir', 'Total', 'speakers', '~', '100', ',', '000']  
 ['Minigir', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Minigir', 'Region', 'Gazelle', 'Peninsula', ',', 'East', 'New',  
 'Britain', 'Province']  
 ['Patpatar', 'Total', 'speakers', '7', ',', '000']  
 ['Patpatar', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kamviri', 'Region', 'and', 'Southern', 'Chitral', 'District']  
 ['Kamviri', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Thai', 'Language', 'codes', 'ISO', '639-3', 'tsq', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',



'\x94', 'legal', 'recognition']  
 ['Sydney', 'Language', 'extinction', 'Late', '19th', '/', 'early', '20th', 'century']  
 ['Bonan', 'Total', 'speakers', '9', ',', '000']  
 ['Eastern', 'Yugur', 'language', 'Total', 'speakers', '3', ',', '500']  
 ['Ogea', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Waray', 'Region', 'Adelaide', 'River', ',', 'Northern', 'Territory']  
 ['Kiga', 'Total', 'speakers', '1', ',', '400', ',', '000']  
 ['Mewati', 'Total', 'speakers', '5', '.', '0', 'million', '(', '2002', ')']  
 ['Shekhawati', 'Spoken', 'in', 'India', '(', 'Shekhawati', 'region', 'of', 'Rajasthan', ')']  
 ['Shekhawati', 'Total', 'speakers', '3', '.', '05', 'million', '(', '1997', ')']  
 ['Domari', 'Total', 'speakers', 'roughly', '1', '.', '9', 'million', ',', 'of', 'which', '1', '.', '3', 'million', 'in', 'Iran']  
 ['Godwar', 'Spoken', 'in', 'India', '(', 'Godwar', 'region', 'of', 'Rajasthan', ')']  
 ['Romanesco', 'Total', 'speakers', '~', '3', ',', '000', ',', '000']  
 ['Woiwurrung', 'Language', 'codes', 'ISO', '639-3', '08e', '\xe2', '\x80', '\x93']  
 ['Media', 'Total', 'speakers', '1000', '(', 'P', '.', 'Bakker', '1999', ')']  
 ['Darkinjung', 'Language', 'codes', 'ISO', '639-3', '0hk', '\xe2', '\x80', '\x93']  
 ['Jicarilla', 'Total', 'speakers', '812', '(', '1990', ')']  
 ['Karaj\xc3\xa1', 'Total', 'speakers', '3600', '(', '1999', ')']  
 ['Wallisian', 'Total', 'speakers', '29', ',', '968']  
 ['West', 'Total', 'speakers', '1', ',', '110']  
 ['Futunan', 'Spoken', 'in', 'Futuna', 'Island', ',', 'Wallis', 'and', 'Futuna', 'and', 'New', 'Caledonia']  
 ['Futunan', 'Total', 'speakers', '6', ',', '600']  
 ['Niuafou', 'Total', 'speakers', '690', '(', 'SIL', '1981', ')']  
 ['Toba', 'Total', 'speakers', '20', ',', '656']  
 ['Nyah', 'Kur', 'Total', 'speakers', '1', ',', '500']  
 ['Zuruah\xc3\xa1', 'Total', 'speakers', '130', '(', '1995', 'AMTB', ')']  
 ['Mocov\xc3\xad', 'Total', 'speakers', '4', ',', '525']  
 ['Kuna', 'Total', 'speakers', 'ca', '.', '58', ',', '700']  
 ['Kuna', 'Language', 'codes', 'ISO', '639-3', ':', 'cuk', '\xe2', '\x80', '\x93', 'San', 'Blas', 'Kuna']  
 ['Chontal', 'Maya', 'Total', 'speakers', 'approx', '.', '50', ',', '000']  
 ['Chontal', 'Region', 'North', 'central', 'and', 'southern', 'Tabasco']  
 ['Gaagudju', 'Language', 'extinction', 'May', '2002', 'when', 'Big', 'Bill', 'Neidjie', 'died']  
 ['Laragiya', 'Total', 'speakers', '6', '(', 'Black', ',', '1983', ')']  
 ['Enindhilyagwa', 'Spoken', 'in', 'Groote', 'Eylandt', 'in', 'the', 'Gulf', 'of', 'Carpentaria']

['Enindhilyagwa', 'Total', 'speakers', '&', 'gt', ';;', '1', ',', '000']  
 ['Garawa', 'Language', 'extinction', 'Late', '20th', 'century']  
 ['Garawa', 'Spoken', 'in', 'Northern', 'Territory', 'and', 'Queensland',  
 ',', 'Australia']  
 ['Kott', 'Language', 'extinction', '19th', 'Century']  
 ['Tataviam', 'Language', 'extinction', '&', 'lt', ';;', '1916']  
 ['Aka', 'Total', 'speakers', '~', '300']  
 ['Ersu', 'Total', 'speakers', '9', ',', '000']  
 ['Rengao', 'Total', 'speakers', '16', ',', '000']  
 ["Fe'fe'", 'Total', 'speakers', '~', '123', ',', '700', '(', '1982', ')']  
 ['Zyphe', 'Total', 'speakers', '20', ',', '000']  
 ['Zumbun', 'Total', 'speakers', '2', ',', '000']  
 ['Barupu', 'Spoken', 'in', 'Sandaun', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Barupu', 'Total', 'speakers', '~', '300']  
 ['Koch', 'Spoken', 'in', 'India', '&', 'amp', ';;', 'Bangladesh']  
 ['Saka', 'Language', 'extinction', '10th', 'century']  
 ['Powhatan', 'Language', 'extinction', 'Late', '18th', 'century']  
 ['Cotabato', 'Total', 'speakers', '~', '30', ',', '000']  
 ['Kabiy\xc3\xa8', 'Total', 'speakers', '~', '730', ',', '000']  
 ['Walmajarri', 'Total', 'speakers', '1', ',', '000']  
 ['Djabugay', 'Total', 'speakers', '3', '(', '1981', ')']  
 ['Inor', 'Total', 'speakers', '280', ',', '000']  
 ['Persian', 'Language', 'codes', 'ISO', '639-3', 'psc', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['North', 'Picene', 'Language', 'extinction', '1st', 'millenium', 'BC']  
 ['Warlmanpa', 'Total', 'speakers', '50', '(', '1981', ')']  
 ['W\xc3\xa8', 'Total', 'speakers', 'over', '300', ',', '000']  
 ['W\xc3\xa8', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Nukuoro', 'Total', 'speakers', 'First', 'language', ':', 'about', '860']  
 ['Sakau', 'Region', 'Big', 'Bay', ',', 'Espiritu', 'Santo', 'Island']  
 ['Doukhorbor', 'Total', 'speakers', 'Up', 'to', '30', '000', 'in', 'its',  
 'heyday', ',', 'hardly', 'any', 'by', 'the', '2000s']  
 ['Valencian', 'Language', 'codes', 'ISO', '639-3', 'vsv', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Vlach', 'Total', 'speakers', '54', ',', '818', '(', '2002', ')']  
 ['Blin', 'Total', 'speakers', '70', ',', '000']  
 ['North', 'Total', 'speakers', '\xe2', '\x89', '\x88', '20']  
 ['Lummi', 'Language', 'extinction', 'c', '.', '2000']  
 ['Ixil', 'Total', 'speakers', 'approx', '.', '70', ',', '000']  
 ['Achi', 'Total', 'speakers', 'approx', '.', '85', ',', '500']  
 ['Awakatek', 'Total', 'speakers', '18', ',', '000', '(', '1998', ')']  
 ['Tojolabal', 'Total', 'speakers', 'approx', '.', '20', ',', '000']  
 ["Itza'", 'Total', 'speakers', 'approx', '.', '10']

['Mopan', 'Total', 'speakers', 'Belize', ',', 'approx', '.', '5', ',', '000']  
 ['Mopan', 'Total', 'speakers', 'Guatemala', ',', 'approx', '.', '2', ',', '000']  
 ['Uspanteko', 'Total', 'speakers', 'approx', '.', '3', ',', '000']  
 ["Tz'utujil", '(', 'Tz'utujiil', ')', 'Total', 'speakers', 'approx', '.', '83', ',', '800']  
 ['Tektitek', '(', 'B'a'aj", ')', 'Total', 'speakers', 'approx', '.', '1', ',', '000']  
 ['Meru', 'Total', 'speakers', '2', '.', '74', 'million']  
 ['Meru', 'Region', 'Mount', 'Kenya', 'and', 'Eastern', 'Province']  
 ['Mwotlap', 'Region', 'Mota', 'Lava', 'island', ',', 'Banks', 'Islands']  
 ['Torwali', 'Total', 'speakers', '60', ',', '000', '(', '1987', 'estimate', ',', '90', ',', '000', '(', '2011', 'Source', 'Khowar', 'Academy', ')']  
 ['Ushoji', 'Total', 'speakers', '500', '(', '2007', ')']  
 ['Ushoji', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Munduruku', 'Total', 'speakers', '10', ',', '000']  
 ['Muruwari', 'Spoken', 'in', 'Queensland', 'and', 'New', 'South', 'Wales', ',', 'Australia']  
 ['Saisiyat', 'Total', 'speakers', '4750', '(', '2002', ')']  
 ['Sop', 'Total', 'speakers', '2', ',', '500']  
 ['Sop', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Chicomuceltec', 'Language', 'extinction', 'late', '20th', 'century']  
 ['Chicomuceltec', 'Region', 'several', 'communities', 'in', 'the', 'Fronteriza', 'and', 'Sierra', 'demographic', 'regions', 'of', 'the', 'southeastern', 'Chiapas', 'highlands', ';', 'isolated', 'villages', 'across', 'the', 'Guatemalan', 'border']  
 ['Kalami', 'Total', 'speakers', '60', ',', '000', '\xe2', '\x80', '\x93', '70', ',', '000', '(', '1995', ')']  
 ['Huichol', 'Total', 'speakers', 'approx', '.', '20', ',', '000', '(', '1990', ')']  
 ['Sempan', 'Spoken', 'in', 'Middle', 'south', 'coast', 'of', 'Western', 'New', 'Guinea']  
 ['Sempan', 'Total', 'speakers', '1', ',', '000', '(', '1987', ')']  
 ['Chocho', 'Chocholteco', ',', 'Chochotec', 'Total', 'speakers', '&', 'lt', ';', '1000']  
 ['Worimi', 'Language', 'extinction', 'It', 'is', 'kept', 'alive', 'by', 'the', 'elders', 'of', 'the', 'Worimi', ',', 'and', 'young', 'members', 'of', 'the', 'tribe', 'are', 'learning', 'it']  
 ['Gadang', 'Total', 'speakers', '2', ',', '500', '(', '1997', ')']  
 ['Khandeshi', 'Total', 'speakers', '2', '.', '5', 'million']  
 ['Susquehannock', 'Spoken', 'in', 'Northeastern', 'United', 'States']  
 ['Pitjantjatjara', 'Spoken', 'in', 'Northwest', 'South', 'Australia']

['Pitjantjatjara', 'Total', 'speakers', '2', ',', '500']  
 ['Derung', 'Total', 'speakers', '11', ',', '300']  
 ['Grebo', 'Total', 'speakers', '265', ',', '000']  
 ['Grebo', 'Language', 'codes', 'ISO', '639-3', ':', 'grj', '\xe2', '\x80',  
 '\x93', 'Southern', 'Grebo', '(', 'incl', '.', 'Jabo', ')']  
 ['Grebo', 'Language', 'codes', 'ISO', '639-3', 'grv', '\xe2', '\x80',  
 '\x93', 'Central', 'Grebo']  
 ['Grebo', 'Language', 'codes', 'ISO', '639-3', 'gbo', '\xe2', '\x80',  
 '\x93', 'Northern', 'Grebo']  
 ['Adnyamathana', 'Total', 'speakers', '20', '(', '1990', ')']  
 ['Pitta', 'Total', 'speakers', '2', '(', '1981', ')']  
 ['Eggon', 'Total', 'speakers', '140', ',', '000', '(', '1990', ')']  
 ['Manam', 'Spoken', 'in', 'Northern', 'New', 'Guinea']  
 ['Manam', 'Total', 'speakers', '7', ',', '000', '(', '1998', ')']  
 ['Nalik', 'Spoken', 'in', 'New', 'Ireland', 'Province', ',', 'Papua',  
 'New', 'Guinea']  
 ['Nalik', 'Total', 'speakers', '5', ',', '138', '(', '1990', ')']  
 ['Garadjari', 'Total', 'speakers', '12', '(', '1991', ')']  
 ['Shubi', 'Total', 'speakers', '150', ',', '000']  
 ['Yurumangu\x3\xad', 'Language', 'extinction', '(', 'attested', '1768',  
 ')']  
 ['Salinan', 'Language', 'extinction', 'ca', '.', '1960']  
 ['Shuswap', 'Total', 'speakers', '1', ',', '650']  
 ['Shuswap', 'Region', 'Central', 'and', 'Southern', 'Interior', 'of',  
 'British', 'Columbia']  
 ['Rekhta', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Kata', 'Total', 'signers', '2', ',', '200', '(', 'of', 'which', '50',  
 'are', 'deaf', ')']  
 ['Kata', 'Language', 'codes', 'ISO', '639-3', 'bqy', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Ban', 'Language', 'codes', 'ISO', '639-3', 'bfk', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Saa', 'Total', 'speakers', '2', ',', '500']  
 ['Maninka', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']  
 ['Maninka', 'Total', 'speakers', '3', ',', '300', ',', '000']  
 ['Maninka', 'Language', 'codes', 'ISO', '639-3', ':', 'myq', '\xe2',  
 '\x80', '\x93', 'Maninka', ',', 'Forest']  
 ['Maninka', 'Language', 'codes', 'ISO', '639-3', 'mku', '\xe2', '\x80',  
 '\x93', 'Maninka', ',', 'Konyanka']  
 ['Otuho', 'Total', 'speakers', '185', ',', '000']  
 ['Yimas', 'Spoken', 'in', 'East', 'Sepik', 'Province', ',', 'Papua', 'New',

'Guinea']  
 ['Barrow', 'Total', 'speakers', '1', '(', '1989', ')']  
 ['Dhundhari', 'Total', 'speakers', '9', 'million', '(', '1997', ')']  
 ['Ngan\xe2\x80\x99gityemerri', 'Total', 'speakers', '275', '(', '1988', ')']  
 ['Naasioi', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Naasioi', 'Total', 'speakers', '~', '10', ',', '000', '(', '1990', ')']  
 ['Oksapmin', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Oksapmin', 'Total', 'speakers', '8', ',', '000', '(', '1991', ')']  
 ['Mizo', 'Total', 'speakers', '700', ',', '000', '+']  
 ['Mizo', 'Total', 'speakers', '674', ',', '756', 'in', 'India', '(', '2001', 'census', ')', ';', '12', ',', '500', 'in', 'Myanmar', '(', '1983', ')', ';', '1', ',', '041', 'in', 'Bangladesh', '(', '1981', 'census', ')']  
 ['Garó', 'Total', 'speakers', '990', ',', '000', '+']  
 ['Garó', 'Total', 'speakers', '889', ',', '479', 'in', 'India', '(', '2001', ')', ';', '102', ',', '000', 'in', 'Bangladesh', '(', '1993', ')']  
 ['Kuku-Thaypan', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Meriam', 'Total', 'speakers', '~', '300', '\xe2', '\x80', '\x93', '400', '(', '1991', ')']  
 ['Mer', 'Total', 'speakers', '85', '(', '2000', ')']  
 ['Dadibi', 'Spoken', 'in', 'Simbu', 'Province', 'and', 'Southern', 'Highlands', 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Dadibi', 'Total', 'speakers', '~', '10', ',', '000', '(', '1991', ')']  
 ['Zumaya', 'Total', 'speakers', '25', '(', '1987', ')']  
 ['Warembori', 'Total', 'speakers', '~', '600']  
 ['Babuza', 'Total', 'speakers', '3', '\xe2', '\x80', '\x93', '4']  
 ['Dimasa', 'Total', 'speakers', '106', ',', '000']  
 ['Lavukaleve', 'Total', 'speakers', '1', ',', '783', '(', '1999', 'SIL', ')']  
 ['Venture\xc3\xb1o', 'Language', 'extinction', 'mid', '20th', 'century']  
 ['Venture\xc3\xb1o', 'Spoken', 'in', 'Southern', 'Californian', 'coastal', 'areas']  
 ['Purari', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Purari', 'Total', 'speakers', '7', ',', '000', '(', '1991', ')']  
 ['Mailu', 'Spoken', 'in', 'Central', 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Mailu', 'Total', 'speakers', '~', '6', ',', '000', '(', '1980', ')']  
 ['Sentani', 'Total', 'speakers', '~', '30', ',', '000', '(', '1996', ')']  
 ['Golin', 'Total', 'speakers', '51', ',', '105', '(', '1981', ')']  
 ['Wambon', 'Total', 'speakers', '~', '3', ',', '000', '(', '1987', ')']  
 ['Maisin', 'Spoken', 'in', 'Oro', 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Maisin', 'Total', 'speakers', '2', ',', '612', '(', '2000', 'census', ')']  
 ['Ghari', 'Total', 'speakers', '12', ',', '119', '(', '1999', ')']  
 ['Oroha', 'Spoken', 'in', 'South', 'Malaita', ',', 'Solomon', 'Islands']

['Oroha', 'Total', 'speakers', '38', '(', '1999', ')']  
 ['Namakura', 'Total', 'speakers', '2', ',', '850', '(', '1983', ')']  
 ['East', 'Total', 'speakers', '~', '5', ',', '000', '(', '1991', ')']  
 ['Aulua', 'Spoken', 'in', 'East', 'Malekula', ',', 'Vanuatu']  
 ['Aulua', 'Total', 'speakers', '~', '300', '(', '1983', ')']  
 ['North', 'Total', 'speakers', '~', '3', ',', '000', '(', '1983', ')']  
 ['Hadiyya', 'Total', 'speakers', '253', ',', '894', '(', '2007', ')']  
 ['Aghu', 'Total', 'speakers', '~', '3', ',', '000']  
 ['Tunggare', 'Total', 'speakers', '~', '500', '(', '1993', ')']  
 ['Mixtecan', 'Total', 'speakers', '550', ',', '000']  
 ['Aghu', 'Tharrnggala', 'Language', 'extinction', 'Unknown']  
 ['Aghu', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Vietnamese', 'Language', 'codes', 'ISO', '639-3', ':', 'hos', '\xe2',  
 '\x80', '\x93', 'Ho', 'Chi', 'Minh', 'City', 'Sign', 'Language']  
 ['Vietnamese', 'Language', 'codes', 'ISO', '639-3', 'hab', '\xe2', '\x80',  
 '\x93', 'Hanoi', 'Sign', 'Language', 'sign', 'language', '\xe2', '\x80',  
 '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal',  
 'recognition']  
 ['Isthmus', 'Zapotec', 'Total', 'speakers', '75', ',', '000', '\xe2',  
 '\x80', '\x93', '100', ',', '000', '(', '2000', 'census', ')']  
 ['Old', 'Javanese', 'Language', 'extinction', 'developed', 'into',  
 'Middle', 'Javanese', 'by', 'the', '13th', 'century']  
 ['Ngkoth', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Ngkoth', 'Language', 'codes', 'ISO', '639-3', '0gt', '\xe2', '\x80',  
 '\x93']  
 ['Uradhi', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Wik-Ngatharr', 'Total', 'speakers', '86', '(', '1981', ')']  
 ['Wik-Ngatharr', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Wurrugu', 'Language', 'extinction', 'Attested', 'in', 'the', '19th',  
 'century']  
 ['Ngardi', 'Spoken', 'in', 'Northern', 'Territory', 'and', 'Western',  
 'Australia']  
 ['Ngardi', 'Language', 'codes', 'ISO', '639-3', '08r', '\xe2', '\x80',  
 '\x93']  
 ['Goldogrin', 'Date', 'founded', 'c', '.', '1910', '\xe2', '\x80', '\x93',  
 'c', '.', '1920']  
 ['Goldogrin', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Elvish',  
 'writing', 'systems', ':', '(', 'mainly', ')', 'sarati']  
 ['Jarai', 'Total', 'speakers', '332', ',', '557']  
 ['Jarai', 'Writing', 'system', 'Vietnam', ':', 'modified', 'Vietnamese',  
 'alphabet', 'Cambodia', ':', 'None']  
 ['Gudang', 'Spoken', 'in', 'Cape', 'York', 'Peninsula']  
 ['Wuthati', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Luthigh', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Yinwum', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Mpakwithi', 'Spoken', 'in', 'Cape', 'York', 'Peninsula']

['Atampaya', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Linngithigh', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Alngith', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Anguthimri', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Yir-Yoront', 'Total', 'speakers', '15', '(', '1991', ')']  
 ['Yir-Yoront', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Northern', 'Paiute', 'Total', 'speakers', '500', 'to', '1600']  
 ['Tangoa', 'Spoken', 'in', 'Tangoa', 'Island', ',', 'Vanuatu']  
 ['Tangoa', 'Total', 'speakers', '~', '800', '(', '2001', ')']  
 ['Bundjalung', 'Total', 'speakers', '10', '(', '1983', ')']  
 ['Ngarinyin', 'Total', 'speakers', '82', '(', '1981', ')']  
 ['Wunambal', 'Total', 'speakers', '20', '(', '1990', ')']  
 ['Kavalan', 'Total', 'speakers', '24', '(', 'Paul', 'Jen-kuei', 'Li', '2000', ')']  
 ['Bahing', 'Total', 'speakers', '2', ',', '765', 'in', 'Nepal', '(', '2001', 'census', ')']  
 ['Kariti\xc3\xa2na', 'Total', 'speakers', '~', '200', '(', '2000', ')']  
 ['Nyikina', 'Spoken', 'in', 'Lower', 'Fitzroy', 'River', ',', 'Western', 'Australia']  
 ['Nyikina', 'Total', 'speakers', '50', '(', '1981', ')']  
 ['Tsez', 'Total', 'speakers', '15', ',', '000']  
 ['Tregami', 'Total', 'speakers', '1', ',', '000', '(', '1994', ')']  
 ['Tregami', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Waigali', 'Total', 'speakers', '1', ',', '500', '(', '2000', ',', 'Van', 'Driem', ')']  
 ['Kata-vari', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Vasi-vari', 'Total', 'speakers', '1', ',', '000', '(', '2000', ')']  
 ['Vasi-vari', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Woi', 'Total', 'speakers', '~', '1', ',', '300']  
 ['Matngele', 'Total', 'speakers', '15', '\xe2', '\x80', '\x93', '20', 'in', '1983']  
 ['Warluwara', 'Total', 'speakers', '3', 'in', '1981']  
 ['Domaaki', 'Total', 'speakers', 'ca', '.', '350', '(', '2004', ')']  
 ['Domaaki', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']

['Pashayi', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Eonavian', 'Total', 'speakers', '45', ',', '000']  
 ['Shumashti', 'Total', 'speakers', '1', ',', '000', '(', '1994', ')']  
 ['Shumashti', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Itonama', 'Total', 'speakers', '&', 'lt', ';', '5']  
 ['Dan', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', "d'Ivoire"]  
 ['Dan', 'Total', 'speakers', '~', '951', ',', '600']  
 ['Ikwerre', 'Total', 'speakers', '200', ',', '000']  
 ['Gogo', 'Total', 'speakers', '1', ',', '300', ',', '000', '(', '1992',  
 'UBS', ')', '(', 'Ethnologue', ')']  
 ['Gogo', 'Writing', 'system', 'Roman', 'alphabet', '(', 'Gogo', 'variant',  
 ')']  
 ['Pongu', 'Total', 'speakers', 'c', '.', '30', ',', '000', '(', '2003',  
 'SIL', ')']  
 ['Gondi', 'Language', 'codes', 'ISO', '639-3', ':', 'ggo', '\xe2', '\x80',  
 '\x93', 'Southern', 'Gondi']  
 ['Gondi', 'Language', 'codes', 'ISO', '639-3', 'gno', '\xe2', '\x80',  
 '\x93', 'Northern', 'Gondi']  
 ['monastic', 'Language', 'codes', 'ISO', '639-3', 'mzg', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Communicationssprache', 'Category', '(', 'purpose', ')', 'International',  
 'auxiliary', 'language']  
 ['Communicationssprache', 'Category', '(', 'sources', ')', 'English',  
 'and', 'German']  
 ['Communicationssprache', 'Setting', 'and', 'usage', 'International',  
 'auxiliary', 'language']  
 ['Kolami', 'Total', 'speakers', '115', ',', '000', '(', '1997', ')']  
 ['Kituba', 'Language', 'codes', 'ISO', '639-3', ':', 'mkw', '\xe2', '\x80',  
 '\x93', 'Kituba', '(', 'Congo-Brazzaville', ')']  
 ['Kituba', 'Language', 'codes', 'ISO', '639-3', 'ktu', '\xe2', '\x80',  
 '\x93', 'Kituba', '(', 'Congo-Kinshasa', ')']  
 ['Mozabite', 'Total', 'speakers', '70', ',', '000', '(', '1995', ')', '(',  
 'source', ':', 'ethnologue', '.', 'com', ')']  
 ['Mozabite', 'Region', "M'zab", '(', 'wilaya', 'of', 'Gharda', '\xc3',  
 '\xaf', 'a', ')']  
 ['Mozabite', 'Language', 'codes', 'ISO', '639-3', 'and', 'Oued', 'Righ']  
 ['Arb\xc3\xabresh', 'Total', 'speakers', '80', ',', '000', '(', 'out',  
 'of', 'an', 'ethnic', 'population', 'of', '260', ',', '000', ')']  
 ['Arb\xc3\xabresh', 'Language', 'codes', 'Linguasphere', '55-AAA-aha',



'to', '55-AAA-ahé']  
['Arrernte', 'Total', 'speakers', '1', ',', '500']  
['Arrernte', 'Language', 'codes', 'ISO', '639-3', 'aer', '\xe2', '\x80', '\x93', 'Eastern', 'Arrernte']  
['Arrernte', 'Language', 'codes', 'ISO', '639-3', 'are', '\xe2', '\x80', '\x93', 'Western', 'Arrernte']  
['Upper', 'Tanana', ',', 'Nabesna', 'Total', 'speakers', '100']  
['Upper', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan', 'variant', ')']  
['Borujerdi', 'dialect', 'Total', 'speakers', 'ca', '.', '0', '.', '8', 'million']  
['Borujerdi', 'Language', 'codes', 'ISO', '639-3', 'lrc', '\xe2', '\x80', '\x93', 'Northern', 'Luri', '(', 'Borujerdi', ')']  
['Tera', 'Total', 'speakers', '~', '100', ',', '620']  
['Gunai', 'Language', 'extinction', 'progressive', 'revival']  
['Wolaytta', 'Total', 'speakers', '1', ',', '627', ',', '955', '(', '2007', 'census', ')']  
['Wolaytta', 'Region', 'Wolaytta', 'region', ',', 'Lake', 'Abaya', 'area']  
['Bakhtiari', 'Total', 'speakers', 'ca', '.', '600', ',', '000']  
['Chinese', 'Language', 'codes', 'ISO', '639-3', 'csl', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
['Nagamese', 'Total', 'speakers', '30', ',', '000']  
['Malak-Malak', 'Total', 'speakers', '9', 'to', '11', 'in', '1988']  
['Tobelo', 'Spoken', 'in', 'North', 'Halmahera', ',', 'Indonesia']  
['Tobelo', 'Total', 'speakers', '27', ',', '720', '(', '2000', ')']  
['Tobelo', 'Language', 'codes', 'ISO', '639-3', 'The', 'Tobelo', 'speaking', 'region', '(', 'grey', ')', 'on', 'Halmahera']  
['Western', 'Total', 'speakers', '~', '10', ',', '804', '(', '2005', ')']  
['(Lower)', 'Tanana', 'Total', 'speakers', '30']  
['(Lower)', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan', 'variant', ')']  
['Koyukon', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan', 'variant', ')']  
['Ogh-Undjan', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
['Ekari', 'Total', 'speakers', '100', ',', '000', '(', '1985', ')']  
['Mumviri', 'Language', 'codes', 'Linguasphere', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
['Betta', 'Kurumba', 'Total', 'speakers', '32', ',', '000', '(', '2003', ')', ',', 'increasing']  
['Betta', 'Region', 'Nilgiri', 'mountains', '(', 'Tamil', 'Nadu', ',', 'Karnataka', ',', 'Kerala', ')']  
['Betta', 'Writing', 'system', 'Kannada', 'script', '.', 'Malayalam', 'script', '.', 'Tamil', 'script']

['Betta', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Samaritan', 'Aramaic', 'Total', 'speakers', 'fewer', 'than', '1', ',',  
 '000', ';', 'liturgical', 'only']  
 ['Samaritan', 'Spoken', 'in', 'Israel', 'and', 'Palestinian', 'Authority',  
 'territories', ',', 'predominantly', 'in', 'Samaria', 'and', 'Holon']  
 ['Gwere', 'Total', 'speakers', '410', ',', '000', '(', '2002', 'census',  
 ')']  
 ['Hlai', 'Total', 'speakers', '615', '000', '(', 'as', 'of', '1999', ')']  
 ['Hlai', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Huli', 'Spoken', 'in', 'Southern', 'Highlands', ',', 'Papua', 'New',  
 'Guinea']  
 ['Huli', 'Total', 'speakers', '70', ',', '000']  
 ['Kom', 'Total', 'speakers', '170', ',', '000']  
 ['Sansiboli', 'Total', 'speakers', 'ca', '.', '60', ',', '000', '(',  
 '2002', ')']  
 ['Uruava', 'Spoken', 'in', 'Bougainville', ',', 'Papua', 'New', 'Guinea']  
 ['Melpa', 'Spoken', 'in', 'Western', 'Highlands', 'Province']  
 ['Melpa', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Melpa', 'Total', 'speakers', '130', ',', '000', '(', '1991', ')']  
 ['Arop-Lokep', 'Spoken', 'in', 'Vitiaz', 'Strait', ',', 'Papua', 'New',  
 'Guinea']  
 ['Arop-Lokep', 'Total', 'speakers', '3', ',', '015', '(', '2000', 'census',  
 ')']  
 ['Seereer-Siin', 'Total', 'speakers', '1', ',', '183', ',', '120']  
 ['Tamambo', 'Total', 'speakers', '3', ',', '000']  
 ['Sissano', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Amaimon', 'Spoken', 'in', 'Madang', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Amaimon', 'Total', 'speakers', '1', ',', '781']  
 ['\xe2\x80\x99Auhelawa', 'Spoken', 'in', 'Milne', 'Bay', 'Province', ',',  
 'Papua', 'New', 'Guinea']  
 ['\xe2\x80\x99Auhelawa', 'Total', 'speakers', '1', ',', '200']  
 ['Adzera', 'Spoken', 'in', 'Morobe', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Adzera', 'Total', 'speakers', '30', ',', '000']  
 ['Waris', 'Spoken', 'in', 'Papua', 'New', 'Guinea', ';', 'Papua',  
 'province']  
 ['Waris', 'Total', 'speakers', '4', ',', '000']  
 ['Amanab', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Amanab', 'Total', 'speakers', '4', ',', '419']  
 ['Haryanvi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',

'of', 'conjuncts', '.', 'More']  
 ['Kannauji', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Afshar', 'Total', 'speakers', 'about', '600', ',', '000']  
 ['Afshar', 'Language', 'codes', 'ISO', '639-3', 'azb', '\xe2', '\x80',  
 '\x93', '(', 'South', 'Azeri', ')']  
 ['Cupe\xc3\xb1o', 'Spoken', 'in', 'Southern', 'California', ',', 'USA']  
 ['Hrangkhol', 'Total', 'speakers', '27', ',', '000']  
 ['\xc3\x84iwoo', 'Total', 'speakers', '8', ',', '400', '(', '1999', 'SIL',  
 ')', '.', 'Population', 'includes', '4', ',', '500', 'on', 'Reef',  
 'Islands', ',', 'others', 'on', 'Santa', 'Cruz', ',', 'and', '1', ',',  
 '000', 'in', 'Honiara', '(', '2000', 'John', 'Rentz', 'SIL', ')']  
 ['\xc3\x84iwoo', 'Region', '10', '\xc2', '\xb0', '13', '\xe2', '\x80',  
 '\xb2', 'S', '166', '\xc2', '\xb0', '12', '\xe2', '\x80', '\xb2', 'E',  
 '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '10', '.', '217',  
 '\xc2', '\xb0', 'S', '166', '.', '2', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
 '\xbf', '/', '-10', '.', '217', ';', '166', '.', '2']  
 ['Luwian', 'Language', 'extinction', 'around', '600', 'BC']  
 ['Luwian', 'Language', 'codes', 'ISO', '639-3', 'Two', 'Luwian',  
 'hieroglyphic', 'texts', 'from', 'the', 'city', 'of', 'Carchemish']  
 ['Fiji', 'Total', 'speakers', '460', ',', '000']  
 ['Fiji', 'Writing', 'system', 'Latin', ',', 'Devanagari', 'script']  
 ['Fiji', 'Language', 'codes', 'ISO', '639-3', 'Countries', 'where', 'Fiji',  
 'Hindi', 'is', 'spoken', '.', 'This', 'page', 'contains', 'Indic', 'text',  
 '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular',  
 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Italian', 'Language', 'codes', 'ISO', '639-3', 'ise', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Broome', 'Spoken', 'in', 'Broome', ',', 'Western', 'Australia']  
 ['Otomi', 'Total', 'speakers', '239', ',', '850', '(', '2005', 'Census',  
 ')']  
 ['Otomi', 'Official', 'language', 'in', 'In', 'Mexico', 'through', 'the',  
 'General', 'Law', 'of', 'Linguistic', 'Rights', 'of', 'Indigenous',  
 'Peoples', '(', 'in', 'Spanish', ')']  
 ['Otomi', 'Language', 'codes', 'ISO', '639-3', ':', 'ote', '\xe2', '\x80',  
 '\x93', 'Mezquital', 'Otomi']  
 ['Otomi', 'Language', 'codes', 'ISO', '639-3', 'otm', '\xe2', '\x80',  
 '\x93', 'Highland', 'Otomi']  
 ['Otomi', 'Language', 'codes', 'ISO', '639-3', 'ott', '\xe2', '\x80',  
 '\x93', 'Temoaya', 'Otomi']  
 ['Cypriot', 'Turkish', 'Total', 'speakers', '177', ',', '000', '(',  
 'Cyprus', ')']  
 ['Cypriot', 'Spoken', 'in', 'Northern', 'Cyprus', ',', 'Cyprus']

['Cypriot', 'Spoken', 'in', '(', 'as', 'Turkish', 'Language', ')']  
 ['Cypriot', 'Official', 'language', 'in', 'Northern', 'Cyprus', ',', 'Cyprus']  
 ['Nyulnyul', 'Language', 'extinction', '1999', ',', 'with', 'the', 'death', 'of', 'Carmel', 'Charles']  
 ['Najdi', 'Total', 'speakers', '9', ',', '863', ',', '520']  
 ['Hejazi', 'Spoken', 'in', 'Hejaz', 'region', ',', 'Saudi', 'Arabia']  
 ['Hejazi', 'Total', 'speakers', '6', ',', '000', ',', '000']  
 ['Onhan', 'Total', 'speakers', '86', ',', '000']  
 ['Berti', 'Language', 'extinction', '1990s', '?']  
 ['Cora', 'Total', 'speakers', '8', ',', '000']  
 ['Central', 'Min', 'Total', 'speakers', '3', '.', '5', 'million']  
 ['Central', 'Spoken', 'in', 'Southern', 'China', ',', 'United', 'States', '(', 'mainly', 'California', ')']  
 ['Central', 'Region', 'eastern', 'Fujian', ';', 'Fuzhou', ';', 'Yong'an', ',', 'Sanming']  
 ['Berom', 'Total', 'speakers', '300', ',', '000', '(', '1993', 'SIL', ')']  
 ['Nete', 'Spoken', 'in', 'East', 'Sepik', 'Province']  
 ['Nete', 'Total', 'speakers', '1000', '(', '1982', 'SIL', ')']  
 ['Cita', 'Total', 'speakers', '8', ',', '000', '(', '1985', ')']  
 ['Marba', 'Total', 'speakers', '124', ',', '357']  
 ['Logo', 'Total', 'speakers', '210', ',', '000']  
 ['Naxi', 'Total', 'speakers', '308', ',', '839', '(', '2000', 'census', ')']  
 ['Naxi', 'Official', 'language', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Avikam', 'Spoken', 'in', 'Grand', 'Lahou', ',', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']  
 ['Shughni', 'Spoken', 'in', 'Afghanistan', '(', '20', ',', '000', ')', ',', 'Tajikistan', '(', '40', ',', '000', ')']  
 ['Shughni', 'Total', 'speakers', 'approximately', '60', ',', '000']  
 ['Vanji', 'Total', 'speakers', 'Extinct', 'as', 'of', 'the', 'late', '19th', 'century']  
 ['Mnong', 'Total', 'speakers', '120', ',', '000']  
 ['Mnong', 'Region', '\xc4', '\x90', '\xe1', '\xba', '\xaf', 'k', 'N', '\xc3', '\xb4', 'ng', 'and', 'B', '\xc3', '\xac', 'nh', 'Ph', '\xc6', '\xb0', '\xe1', '\xbb', '\x9b', 'c', 'provinces', ';', 'Mondulkiri', 'in', 'Cambodia']  
 ['Mnong', 'Language', 'codes', 'ISO', '639-3', ':', 'cmo', '\xe2', '\x80', '\x93', 'Central', 'Mnong']  
 ['Mnong', 'Language', 'codes', 'ISO', '639-3', 'mng', '\xe2', '\x80', '\x93', 'Eastern', 'Mnong']  
 ['Mnong', 'Language', 'codes', 'ISO', '639-3', 'mn', '\xe2', '\x80', '\x93', 'Southern', 'Mnong']  
 ['Tindi', 'Total', 'speakers', '5000', '(', '2000', ')']  
 ['African', 'Romance', 'Language', 'extinction', 'may', 'have', 'survived',

'as', 'late', 'as', 'the', '17th', 'century']  
 ['Coast', 'Tsimshian', 'Total', 'speakers', 'less', 'than', '200']  
 ['Coast', 'Region', 'northwest', 'British', 'Columbia', ',', 'southeast',  
 'Alaska']  
 ['Sarikoli', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Sarikoli', 'Total', 'speakers', 'approximately', '20', ',', '500', '(',  
 '1990', ')']  
 ['Sarikoli', 'Language', 'codes', 'Linguasphere', 'Xinjiang', 'province',  
 ',', 'the', 'light', 'blue', 'are', 'marks', 'the', 'place', 'were',  
 'Sarikoli', 'is', 'spoken']  
 ['Amis', 'Total', 'speakers', 'under', '180', ',', '978', '(', 'number',  
 'of', 'ethnic', 'Amis', 'per', 'Council', 'of', 'Indigenous', 'Peoples',  
 'of', 'Executive', 'Yuan', ',', 'Taiwan', ',', 'in', '2009', ')']  
 ['Tsou', 'Total', 'speakers', '3000', '(', 'as', 'of', '1994', ')']  
 ['Puyuma', 'Total', 'speakers', 'fewer', 'than', '8', ',', '000']  
 ['Rukai', 'Total', 'speakers', '~', '10', ',', '000']  
 ['Thao', 'Total', 'speakers', '5', 'or', '6', '(', '2000', ')']  
 ['Kanakanabu', 'Total', 'speakers', '6', '\xe2', '\x80', '\x93', '8']  
 ['Laurentian', 'Language', 'extinction', 'late', '16th', 'century']  
 ['Laurentian', 'Region', 'Saint', 'Lawrence', 'River', 'Valley']  
 ['Laurentian', 'Language', 'codes', 'ISO', '639-3', 'Territory',  
 'occupied', 'by', 'the', 'St', '.', 'Lawrence', 'Iroquoians', ',', 'circa',  
 '1535']  
 ['Zhang-Zhung', 'Language', 'extinction', 'tenth', 'century']  
 ['Zhang-Zhung', 'Spoken', 'in', 'Western', 'Tibet', 'and', 'Central',  
 'Asia']  
 ['Vestinian', 'Language', 'extinction', 'ca', '.', '100', 'BC']  
 ['Northeastern', 'Spoken', 'in', 'Jilin', ',', 'Heilongjiang', 'and',  
 'Liaoning', 'provinces', 'of', 'the', 'People's', 'Republic', 'of',  
 'China', ';;', '(', 'Overseas', ',', 'United', 'States-New', 'York', 'City',  
 ')']  
 ['Southern', 'Regulated', 'by', 'none', '(', 'Academia', 'Mayor', 'de',  
 'la', 'Lengua', 'Quechua', ')']  
 ['Cusco', 'Total', 'speakers', '1', ',', '500', ',', '000']  
 ['Old', 'Dutch', 'Language', 'extinction', 'developed', 'into', 'Middle',  
 'Dutch', 'by', 'the', 'middle', 'of', 'the', '12th', 'century']  
 ['Old', 'Spoken', 'in', 'the', 'Low', 'Countries']  
 ['Ancient', 'North', 'Arabian', 'Language', 'extinction', 'marginalized',  
 'by', 'Classical', 'Arabic', 'from', 'the', '7th', 'century']  
 ['Ngong', 'Total', 'speakers', '2', '(', '1983', ')']  
 ['Central', 'Bicolano', 'Total', 'speakers', '2', '.', '5', 'million', ';;',  
 '7th', 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']  
 ['Central', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')',  
 ';;']  
 ['Central', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']

['Austrian', 'Language', 'codes', 'ISO', '639-3', 'asq', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Namwanga', 'Total', 'speakers', '260', ',', ',', '000']  
 ['Ngulu', 'Total', 'speakers', '132', ',', ',', '000', '(', '1987', ')']  
 ['Wambaya', 'Total', 'speakers', 'Extinct', '(', '2008', ')']  
 ['Wambaya', 'Total', 'speakers', '12', '(', '1981', ')']  
 ['Wambaya', 'Region', 'Barkly', 'Tableland', ',', ',', 'Northern', 'Territory']  
 ['Capiznon', 'Total', 'speakers', '639', ',', ',', '000']  
 ['Qingdao', 'dialect', 'Total', 'speakers', '~', '3', 'million']  
 ['Archi', 'Total', 'speakers', '1', ',', ',', '200']  
 ['Nai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Nai', 'Total', 'speakers', '595', '(', 'SIL', '2003', ')']  
 ['Ona', 'Total', 'speakers', '1', '\xe2', '\x80', '\x93', '3', '(', '1991',  
 ')']  
 ['Wanka', 'Total', 'speakers', '300', ',', ',', '000']  
 ['Myene', 'Region', 'Ogooue-Maritime', 'Province', ',', ',', 'Middle', 'Ogooue',  
 'Province']  
 ['Belhare', 'Spoken', 'in', 'Dhankuta', 'district', ',', ',', 'Nepal']  
 ['Belhare', 'Total', 'speakers', '2', ',', ',', '000']  
 ["P'urh\xc3\xa9pecha,", 'Tarascan', ',', ',', 'Phorh', '\xc3', '\xa9', 'Total',  
 'speakers', '~', '120', ',', ',', '000']  
 ['Tetela', 'Total', 'speakers', '750', ',', ',', '000', '(', '1991', 'est', '.',  
 ')']  
 ['Tetela', 'Region', 'Northern', 'Kasai', 'Oriental', 'Province']  
 ['Ancash', 'Total', 'speakers', '1', ',', ',', '000', ',', ',', '000']  
 ['Bura-Pabir', 'Total', 'speakers', '~', '250', ',', ',', '000', '(', '1987',  
 ')']  
 ['Nara', 'Total', 'speakers', '80', ',', ',', '000']  
 ['Kunama', 'Total', 'speakers', '\xe2', '\x89', '\x88', '195', ',', ',', '000']  
 ['Kunama', 'Total', 'speakers', '(', 'Eritrea', ')', '187', ',', ',', '000']  
 ['Kunama', 'Total', 'speakers', '(', 'Ethiopia', ')', '7', ',', ',', '428']  
 ['Bade', 'Total', 'speakers', '250', ',', ',', '000', '(', '1993', ')']  
 ['Pumpokol', 'Language', 'extinction', '18th', 'century']  
 ['Nyanga-li', 'Total', 'speakers', '48000', '(', '2002', ')']  
 ['Yami', 'Total', 'speakers', '3', ',', ',', '800', '(', 'Rau', '2006', ':',  
 '79', ')']  
 ['Western', 'Apache', 'Total', 'speakers', '12', ',', ',', '693']  
 ['War', 'Total', 'speakers', '28', ',', ',', '000']  
 ['Vitu', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Vitu', 'Total', 'speakers', '7', ',', ',', '000']  
 ['Bube', 'Total', 'speakers', '40', ',', ',', '000']  
 ['Holikachuk', 'Writing', 'system', 'Latin', '(', 'Northern', 'Athabaskan',  
 'variant', ')']  
 ['Waimoa', 'Spoken', 'in', 'Northeast', 'East', 'Timor']  
 ['Waimoa', 'Total', 'speakers', '3', ',', ',', '000']

['Mankanya', 'Total', 'speakers', '68', ',', '955']  
 ['Gajirrabeng', 'Spoken', 'in', 'The', 'Kimberley', ',', 'Western',  
 'Australia']  
 ['Miriwoong', 'Spoken', 'in', 'Kununurra', ',', 'Western', 'Australia']  
 ['Miriwoong', 'Total', 'speakers', 'About', '20']  
 ['Bai', 'Total', 'speakers', '1', '240', '000', '(', '2003', ')']  
 ['Bai', 'Language', 'codes', 'ISO', '639-3', ':', 'bca', '\xe2', '\x80',  
 '\x93', 'Central', 'Bai', ',', 'Jianchuan', 'dialect']  
 ['Bai', 'Language', 'codes', 'ISO', '639-3', 'bfs', '\xe2', '\x80', '\x93',  
 'Southern', 'Bai', ',', 'Dali', 'dialect']  
 ['Bai', 'Language', 'codes', 'ISO', '639-3', 'bfc', '\xe2', '\x80', '\x93',  
 'Northern', 'Bai', ',', 'Bijiang', 'dialect']  
 ['Lungga', 'Total', 'speakers', '2', ',', '767', '(', 'as', 'of', '1997',  
 ')']  
 ['Kija', 'Spoken', 'in', 'From', 'Halls', 'Creek', 'to', 'Kununurra', ',',  
 'Western', 'Australia']  
 ['Ili', 'Turki', 'Total', 'speakers', '120']  
 ['Tofa', 'Total', 'speakers', '28', '(', '2001', ')']  
 ['Palawano', 'Total', 'speakers', '30', ',', '000']  
 ['Agutaynen', 'Total', 'speakers', '13', ',', '700']  
 ['Kagayanen', 'Total', 'speakers', '20', ',', '000']  
 ['Porohanon', 'Total', 'speakers', '23', ',', '000']  
 ['Assan', 'Language', 'extinction', '18th', 'Century']  
 ['Buhid', 'Total', 'speakers', '8', ',', '000']  
 ['Tadyawan', 'Total', 'speakers', '4', ',', '146']  
 ["Hanun\xc3\xb3'o", 'Total', 'speakers', '14', ',', '000']  
 ["Hanun\xc3\xb3'o", 'Writing', 'system', 'Hanun', '\xc3', '\xb3', '"o"]  
 ['Tawbuid', 'Total', 'speakers', '14', ',', '000']  
 ['Tawbuid', 'Language', 'codes', 'ISO', '639-3', ':', 'bnj', '\xe2',  
 '\x80', '\x93', 'Eastern', 'Tawbuid']  
 ['Tawbuid', 'Language', 'codes', 'ISO', '639-3', 'twb', '\xe2', '\x80',  
 '\x93', 'Western', 'Tawbuid']  
 ['Alangan', 'Total', 'speakers', '7', ',', '694']  
 ['Iraya', 'Total', 'speakers', '10', ',', '000']  
 ['Tlapanec', 'Total', 'speakers', 'more', 'than', '98', ',', '000']  
 ['Hani', 'Total', 'speakers', '520', ',', '000']  
 ['Dano', 'Spoken', 'in', 'Eastern', 'Highlands', 'Province']  
 ['Dano', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Dano', 'Total', 'speakers', '30', ',', '000', '(', '1987', 'SIL', ')']  
 ['Murrinh-patha', 'Total', 'speakers', 'Over', '1', ',', '500']  
 ['Nunggubuyu', 'Spoken', 'in', 'Numbulwar', ',', 'Northern', 'Territory']  
 ['Nunggubuyu', 'Total', 'speakers', 'c', '.', '70', '+', 'unknown',  
 'number', 'of', 'partial', 'or', 'second-language', 'speakers', '(',  
 'perhaps', '200', ')', '(', '2001', ')']  
 ['Trique', 'Total', 'speakers', '25', ',', '000']  
 ['Trique', 'Language', 'codes', 'ISO', '639-3', 'trq', '\xe2', '\x80',

'\x93', 'San', 'Mart', '\xc3', '\xad', 'n', 'Itunyoso']  
 ['Subtiaba', 'Total', 'speakers', '0', '(', '5000', 'ethnic', 'Subtiaba', ')']  
 ['Cuicatec', 'Total', 'speakers', '15', ',', '000']  
 ['Big', 'Nambas', 'Total', 'speakers', '1', ',', '800', '(', 'as', 'of', '1983', ')']  
 ['Mixe', 'Total', 'speakers', 'ca', '.', '188', ',', '000']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'mir', '\xe2', '\x80', '\x93', 'Isthmus', 'Mixe']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'mto', '\xe2', '\x80', '\x93', 'Totontepec', 'Mixe']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'mxp', '\xe2', '\x80', '\x93', 'Tlahuitoltepec', 'Mixe']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'mxq', '\xe2', '\x80', '\x93', 'Juquila', 'Mixe']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'neq', '\xe2', '\x80', '\x93', 'North', 'Central', 'Mixe']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'pxm', '\xe2', '\x80', '\x93', 'Quetzaltepec', 'Mixe']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'plo', '\xe2', '\x80', '\x93', 'Oluta', 'Popoluca']  
 ['Mixe', 'Language', 'codes', 'ISO', '639-3', 'pos', '\xe2', '\x80', '\x93', 'Sayula', 'Popoluca']  
 ['Zoque', 'Total', 'speakers', '&', 'gt', ';', '90', ',', '000']  
 ['Zoque', 'Language', 'codes', 'ISO', '639-3', 'poi', '\xe2', '\x80', '\x93', 'Sierra', 'Popoluca']  
 ['Zoque', 'Language', 'codes', 'ISO', '639-3', 'poq', '\xe2', '\x80', '\x93', 'Texistepec', 'Popoluca']  
 ['Nanai', 'Total', 'speakers', '3', ',', '902']  
 ['Guruntum', 'Total', 'speakers', '15', ',', '000', '(', 'As', 'of', '1993', ')']  
 ['Apurin\xc3\xa3', 'Total', 'speakers', '2', ',', '000']  
 ['Yokohama', 'Pidgin', 'Japanese', 'Language', 'extinction', 'End', 'of', 'the', '19th', 'century']  
 ['Kera', 'Total', 'speakers', '50', ',', '000']  
 ['Sivandi', 'Total', 'speakers', '7', '000']  
 ['Jambi', 'Total', 'speakers', '890', '000', '(', 'SIL', '1989', ')']  
 ['Maricopa', 'Total', 'speakers', '181', '(', '1990', ')']  
 ['Maricopa', 'Total', 'speakers', '160', '(', '2000', ')']  
 ['Wardaman', 'Total', 'speakers', '50', 'Black', '(', '1983', ')']  
 ['Pipil', 'Spoken', 'in', 'El', 'Salvador', '(', 'Sonsonate', ',', 'Ahuachapan', ',', 'La', 'Libertad', ',', 'San', 'Salvador', ')']  
 ['Pipil', 'Total', 'speakers', '3', ',', '000']  
 ['Wadikali', 'Language', 'extinction', 'Early', '20th', 'century']  
 ['Wadikali', 'Spoken', 'in', 'Northwestern', 'New', 'South', 'Wales']  
 ['Wadikali', 'Language', 'codes', 'ISO', '639-3', '0u8', '\xe2', '\x80',



'\x93']  
 ['Sre', 'Total', 'speakers', '130', ',', '000']  
 ['Xaasongaxango', 'Total', 'speakers', '128', ',', '000']  
 ['Drehet', 'Spoken', 'in', 'Manus', ',', 'Papua', 'New', 'Guinea']  
 ['Jicaque', 'Total', 'speakers', '350', '(', '1997', ')']  
 ['Jicaque', 'Language', 'codes', 'ISO', '639-3', 'The', 'Jicaque',  
 'languages', 'are', 'in', 'Nicaragua', 'in', 'the', 'center', 'of', 'the',  
 'map']  
 ['Huave', 'Total', 'speakers', '&', 'gt', ';', '18', ',', '000', '(',  
 '1990', ')', '(', 'Ethnologue', ')']  
 ['Huave', 'Language', 'codes', 'ISO', '639-3', ':', 'hue', '\xe2', '\x80',  
 '\x93', 'San', 'Francisco', 'del', 'Mar']  
 ['Huave', 'Language', 'codes', 'ISO', '639-3', 'huv', '\xe2', '\x80',  
 '\x93', 'San', 'Mateo', 'del', 'Mar']  
 ['Huave', 'Language', 'codes', 'ISO', '639-3', 'hve', '\xe2', '\x80',  
 '\x93', 'San', 'Dionisio', 'del', 'Mar']  
 ['Huave', 'Language', 'codes', 'ISO', '639-3', 'hvv', '\xe2', '\x80',  
 '\x93', 'Santa', 'Mar', '\xc3', '\xad', 'a', 'del', 'Mar', ',', 'Oaxaca']  
 ['Sierra', 'Popoluca', 'Total', 'speakers', '&', 'gt', ';', '30', ',', '000']  
 ['Oluta', 'Popoluca', ',', 'Olutec', 'Total', 'speakers', '&', 'gt', ';',  
 '~', '30']  
 ['Chatino', 'Total', 'speakers', '&', 'lt', ';', '23', ',', '000']  
 ['Sayula', 'Popoluca', 'Total', 'speakers', '&', 'gt', ';', '4', ',', '000']  
 ['Texistepec', 'Popoluca', 'Total', 'speakers', '&', 'gt', ';', '400']  
 ['Modern', 'Category', '(', 'purpose', ')', 'Modern', 'Indo-European']  
 ['Betawi', 'Total', 'speakers', '2', '.', '7', 'million', '(', '1993', ')']  
 ['Mats\xc3\xa9s', 'Total', 'speakers', '~', '2', ',', '000']  
 ['Baraba', 'Total', 'speakers', '8', ',', '000']  
 ['Isnag', 'Total', 'speakers', '55', ',', '000', '(', '2007', 'est', '.',  
 ')']  
 ['Anyin', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire"]  
 ['Anyin', 'Total', 'speakers', '860', ',', '000']  
 ['Southern', 'Spoken', 'in', 'California', ',', 'western', 'slopes', 'of',  
 'Sierra', 'Nevadas', 'Ethnicity']  
 ['Southern', 'Spoken', 'in', 'Valley', 'and', 'Sierra', 'Miwok']  
 ['Southern', 'Total', 'speakers', '7', 'in', '1994']  
 ['Aghul', 'Total', 'speakers', '28', ',', '300', '(', '2002', 'census',  
 ')']  
 ['Naao', 'Language', 'extinction', '19th', 'century']  
 ['Naao', 'Spoken', 'in', 'Eyre', 'Peninsula', ',', 'South', 'Australia']  
 ['Siculo-Arabic', 'Total', 'speakers', 'Substantially', 'extinct', 'by',  
 '1300']  
 ['Tarahumara', 'Total', 'speakers', '~', '70', ',', '000', '(', 'All',  
 'varieties', ')']

['Thiin', 'Language', 'codes', 'ISO', '639-3', 'djl', '\xe2', '\x80',  
 '\x93', '(', 'Jiwarli', ')']  
 ['Camsa', 'Total', 'speakers', '4', ',', '800']  
 ['Chong', 'Total', 'speakers', '5', ',', '500']  
 ['Korean', 'Signed', 'in', 'South', 'Korea']  
 ['Korean', 'Language', 'codes', 'ISO', '639-3', 'kvk', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Bolinao', 'language', 'Total', 'speakers', '~', '50', ',', '000']  
 ['Bolinao', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Bari', 'Total', 'speakers', '480', ',', '000']  
 ['Old', 'Korean', 'Language', 'extinction', 'Evolved', 'into', 'Middle',  
 'Korean', 'by', 'the', '10th', 'century']  
 ['Silla', 'Language', 'extinction', 'gave', 'rise', 'to', 'Middle',  
 'Korean']  
 ['Silla', 'Language', 'codes', 'ISO', '639-3', 'The', 'Three', 'Kingdoms',  
 'of', 'Korea', ',', 'with', 'Silla', 'in', 'red']  
 ['Mambwe', 'Total', 'speakers', '680', ',', '000']  
 ["S\x3\xb6l'ring", 'North', 'Frisian', 'Total', 'speakers', 'unknown',  
 ',', 'fewer', 'than', '500']  
 ['Tepehu\x3\xa1n', 'Total', 'speakers', '~', '25', ',', '000', '(', 'All',  
 'varieties', ')']  
 ['Tepehu\x3\xa1n', 'Language', 'codes', 'ISO', '639-3', ':', 'ntp',  
 '\xe2', '\x80', '\x93', 'Northern', 'Tepehu', '\xc3', '\xa1', 'n']  
 ['Tepehu\x3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'stp', '\xe2',  
 '\x80', '\x93', 'Southeastern', 'Tepehu', '\xc3', '\xa1', 'n']  
 ['Tepehu\x3\xa1n', 'Language', 'codes', 'ISO', '639-3', 'tla', '\xe2',  
 '\x80', '\x93', 'Southwestern', 'Tepehu', '\xc3', '\xa1', 'n']  
 ['Pima', 'Bajo', 'Total', 'speakers', '~', '1000']  
 ['Mexicanero', 'Total', 'speakers', '~', '1000']  
 ['Mexicanero', 'Region', 'Durango', '(', 'Towns', 'of', 'San', 'Pedro',  
 'J', '\xc3', '\xad', 'cora', 'and', 'San', 'Juan', 'Buenaventura', ')']  
 ['Zande', 'Total', 'speakers', '1', ',', '200', ',', '000']  
 ['Lenca', 'Language', 'extinction', 'presumed', 'extinct', '.', 'One',  
 'speaker', 'of', 'Chilanga', '(', 'Salvadoran', ')', 'Lenca', 'reported',  
 'in', 'the', '1970s', '.', 'Some', 'semi-speakers', 'of', 'Honduran',  
 'Lenca', 'reported', 'in', 'the', '1990s']  
 ['Marka', 'Total', 'speakers', '25', '000']  
 ['Manyá', 'Total', 'speakers', '70', '000']  
 ['Yora', 'Total', 'speakers', '&', 'gt', ';', '450']  
 ['Chorotega', 'Spoken', 'in', 'Honduras', ',', 'Costa', 'Rica', 'and',  
 'Nicaragua', 'Ethnicity']  
 ['Pame', 'Total', 'speakers', '&', 'lt', ';', '10000']  
 ['Pame', 'Region', 'San', 'Luis', 'Potos', '\xc3', '\xad', ',', 'Puebla']  
 ['Pame', 'Language', 'codes', 'ISO', '639-3', ':', 'pbs', '\xe2', '\x80',

'\x93', 'Central', 'Pame']  
 ['Pame', 'Language', 'codes', 'ISO', '639-3', 'pmq', '\xe2', '\x80',  
 '\x93', 'Northern', 'Pame']  
 ['Pame', 'Language', 'codes', 'ISO', '639-3', 'pmz', '\xe2', '\x80',  
 '\x93', 'Southern', 'Pame']  
 ['Chichimeca', 'Jonaz', 'Total', 'speakers', '&', 'lt', ';', '200']  
 ['Imbo', 'Ungu', 'Total', 'speakers', '42', ',', '500', '(', '2000',  
 'census', ')', '.', '15', '%', 'monolingual']  
 ['Imbo', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Imbo', 'Region', 'Southern', 'Highlands', 'Province', ',', 'Ialibu',  
 'District', '.', 'Coordinates']  
 ['Imbo', 'Region', '6', '\xc2', '\xb0', '8', '\xe2', '\x80', '\xb2', 'S',  
 '144', '\xc2', '\xb0', '1', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb',  
 '\xbf', '/', '\xef', '\xbb', '\xbf', '6', '.', '133', '\xc2', '\xb0', 'S',  
 '144', '.', '017', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-6',  
 '.', '133', ';', '144', '.', '017']  
 ['Lango', 'Total', 'speakers', '1', '.', '8', 'million']  
 ['Lango', 'Total', 'speakers', '20', ',', '000', '(', '1987', ')']  
 ['Koryo-mar', 'Total', 'speakers', '216', ',', '811', 'as', 'mother',  
 'tongue', '(', '1989', ')', ';', 'additional', 'second-language',  
 'speakers', 'number', ',', 'largely', 'ethnic', 'Korean', ',', 'number',  
 'in', 'the', 'tens', 'of', 'thousands']  
 ['Huambisa', 'Total', 'speakers', '9', ',', '333', '(', '2000', 'WCD', ')']  
 ['Bahamas', 'Total', 'speakers', '400', ',', '000']  
 ['Wik-Mungknh', 'Total', 'speakers', '840', '(', '1996', ')']  
 ['Wik-Mungknh', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Wik-Ngathan', 'Total', 'speakers', '130', '(', '1981', ')']  
 ['Wik-Ngathan', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland',  
 'Coordinates']  
 ['Wik-Ngathan', 'Region', '13', '\xc2', '\xb0', '52', '\xe2', '\x80',  
 '\xb2', 'S', '141', '\xc2', '\xb0', '31', '\xe2', '\x80', '\xb2', 'E',  
 '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '13', '.', '867',  
 '\xc2', '\xb0', 'S', '141', '.', '517', '\xc2', '\xb0', 'E', '\xef',  
 '\xbb', '\xbf', '/', '-13', '.', '867', ';', '141', '.', '517']  
 ['Pochutec', 'Language', 'extinction', '20th', 'century']  
 ['Canela', 'Total', 'speakers', '1', ',', '420', '(', '1995', ')']  
 ['Matlatzinca', 'Total', 'speakers', '&', 'lt', ';', '1', ',', '500']  
 ['Animere', 'Total', 'speakers', 'Estimates', 'vary', ':', '250', '(',  
 'Sommer', '1992', ')', ';', '700', '(', 'Ethnologue', '2003', ')', ';',  
 '30', '(', 'Blench', '2006', ')']  
 ['Animere', 'Region', 'East', 'Central', 'Ghana', ',', 'Kecheibi', 'and',  
 'Kunda', 'villages']  
 ['Tiriy\xc3\xb3', 'Total', 'speakers', '~', '2', ',', '000', '(', 'in',  
 '2005', ')']  
 ['Tiriy\xc3\xb3', 'Region', 'Northern', 'Amazonia', ',', 'Guianas',  
 'Plateau', 'Ethnicity']

['Sangsari', 'Total', 'speakers', '36', ',', ',', '000', '(', '2006', ')']  
 ['Sangsari', 'Spoken', 'in', 'Semnan', '&', 'amp', ';', 'Mazandaran', '&',  
 'amp', ';', 'Golestan', '&', 'amp', ';', 'Markazi', 'provinces', 'of',  
 'Iran']  
 ['Iwal', 'Total', 'speakers', '1', ',', ',', '500', '(', '1987', 'SIL', ')']  
 ['Iwal', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Boro', 'Language', 'extinction', '20th', 'century']  
 ['Santiago', 'Total', 'speakers', '66', ',', ',', '000']  
 ['Batak', 'Toba', 'Total', 'speakers', '2', ',', ',', '000', ',', ',', '000', 'as',  
 'of', 'year', '2000']  
 ['Batak', 'Region', 'Samosir', 'Island', '(', '2', '\xc2', '\xb0', '30',  
 '\xe2', '\x80', '\xb2', 'N', ',', ',', '99', '\xc2', '\xb0', ')', ',', ',', 'and',  
 'to', 'the', 'east', ',', ',', 'south', ',', ',', 'and', 'west', 'of', 'Toba',  
 'Lake', 'in', 'north', 'Sumatra']  
 ['Sora', 'Total', 'speakers', '288', ',', ',', '000', '(', '1997', ')']  
 ['Korku', 'Spoken', 'in', 'Central', 'India', '(', 'Madhya', 'Pradesh',  
 ',', ',', 'Maharashtra', ')']  
 ['Korku', 'Total', 'speakers', '478', ',', ',', '000', '(', 'as', 'of', '1997',  
 ')']  
 ['Hangzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Kugu-Uwanh', 'Total', 'speakers', '40', '(', '1981', ')']  
 ['Kugu-Uwanh', 'Region', 'Cape', 'York', 'Peninsula', ',', ',', 'Queensland',  
 'Coordinates']  
 ['Kugu-Uwanh', 'Region', '14', '\xc2', '\xb0', '4', '\xe2', '\x80', '\xb2',  
 'S', '141', '\xc2', '\xb0', '43', '\xe2', '\x80', '\xb2', 'E', '\xef',  
 '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '14', '.', '067', '\xc2',  
 '\xb0', 'S', '141', '.', '717', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
 '\xbf', '/', '-14', '.', '067', ';', '141', '.', '717']  
 ['Palpa', 'Spoken', 'in', 'Palpa', '(', 'Nepal', ')', ',', ',', 'Kumaon']  
 ['Palpa', 'Total', 'speakers', '7', ',', ',', '562', '(', '2000', ')']  
 ['Wayuu', 'Total', 'speakers', '305', ',', ',', '000']  
 ['Wayuu', 'Region', 'Northern', 'South', 'America', 'Ethnicity']  
 ['Wayuu', 'Language', 'codes', 'ISO', '639-3', 'Extent', 'of', 'both',  
 'the', 'Wayuu', 'people', 'and', 'language']  
 ['Argentine', 'Language', 'codes', 'ISO', '639-3', 'aed', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Central', 'Total', 'speakers', '2', ',', ',', '227', ',', ',', '642']  
 ['Chirip\xc3\xa1', 'Total', 'speakers', '11', ',', ',', '900', '+']  
 ['Iyo'wujwa', 'Total', 'speakers', '2', ',', ',', '008']  
 ['Mby\xc3\xa1', 'Total', 'speakers', '27', ',', ',', '000']  
 ['Kaiw\xc3\xa1', 'Total', 'speakers', '20', ',', ',', '512']  
 ['Nivacl\xc3\xa9', 'Total', 'speakers', '8', ',', ',', '600']  
 ['Pilag\xc3\xa1', 'Total', 'speakers', '2', ',', ',', '000']  
 ['Wich\xc3\xad', 'Total', 'speakers', '15', ',', ',', '000']  
 ['Wich\xc3\xad', 'Total', 'speakers', '2', ',', ',', '081']

['Wich\xc3\xad', 'Total', 'speakers', '25', ',', '000']  
 ['Chilean', 'Language', 'codes', 'ISO', '639-3', 'csg', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Changsha', 'Region', 'Changsha', ',', 'Hunan', 'province']  
 ['Kunza', 'Language', 'extinction', 'ca', '.', '1950s']  
 ['Koromfe', 'Total', 'speakers', '10', ',', '000']  
 ['Shuangfeng', 'Region', 'Shuangfeng', ',', 'Hunan', 'province']  
 ['Wenzhou', 'Region', 'and', 'in', 'Wenzhou', 'immigrant', 'populations',  
 'in', 'New', 'York', 'City', ';;', 'Paris', ';;', 'Milan', 'and', 'Prato']  
 ['Ovambo', 'Total', 'speakers', '670', ',', '000']  
 ['Pohnpeian', 'Total', 'speakers', '29', ',', '000']  
 ['Pingelapese', 'Total', 'speakers', 'approx', '.', '750']  
 ['Bolivian', 'Total', 'signers', '350-400', '(', '1988', ')']  
 ['Bolivian', 'Language', 'codes', 'ISO', '639-3', 'bvl', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Cavine\xc3\xb1a', 'Total', 'speakers', '1', ',', '180']  
 ['Eastern', 'Total', 'speakers', '50', ',', '000', '(', '45', ',', '000',  
 'in', 'Bolivia', ',', '15', ',', '000', 'in', 'Argentina', ',', '300',  
 'in', 'Paraguay', ')']  
 ['Macushi', 'Total', 'speakers', '24', ',', '600']  
 ['Ter\xc3\xa1ana', 'Total', 'speakers', '15', ',', '000']  
 ['Avokaya', 'Spoken', 'in', 'Southern', 'Sudan', ':', 'West', 'Equatoria',  
 'and', 'Democratic', 'Republic', 'of', 'the', 'Congo', ':', 'Haut-Uele']  
 ['Avokaya', 'Total', 'speakers', '40', ',', '000', 'in', 'Sudan', '(',  
 '2002', ')', ';;', '25', ',', '000', 'in', 'Congo', '(', '1989', ')', ';;',  
 '65', ',', '000', 'in', 'all', 'countries']  
 ['Eskayan', 'Language', 'codes', 'ISO', '639-2']  
 ['Leizhou', 'Total', 'speakers', '4', ',', '000', ',', '000']  
 ['Poqomchi', 'Total', 'speakers', 'approx', '.', '92', ',', '000']  
 ['Poqomam', 'Total', 'speakers', 'approx', '.', '49', ',', '000']  
 ['Sakapultek', '(', 'Sacapulteco', ')', 'Total', 'speakers', 'approx', '.',  
 '15', ',', '000', '(', '2006', 'SIL', ')']  
 ['Sakapultek', 'Region', 'El', 'Quich', '\xc3', '\xa9']  
 ['Sipakapense', '(', 'Sipacape', '\xc3', '\xb1', 'o', ')', 'Total',  
 'speakers', 'approx', '.', '8', ',', '000']  
 ['Q'anjo'bal', '(', 'Kanjobal', ')', 'Total', 'speakers', 'approx', '.',  
 '77', ',', '700']  
 ['Mocho\xe2\x80\x99', 'Total', 'speakers', 'approx', '.', '&', 'gt', ';;',  
 '150']  
 ['Mocho\xe2\x80\x99', 'Region', 'Eastern', 'Chiapas', '(', 'villages',  
 'of', 'Tuzatl', '\xc3', '\xa1', 'n', 'and', 'Mototzintla', ')']  
 ['Middle', 'Armenian', 'Language', 'extinction', 'developed', 'into',  
 'Armenian']  
 ['Middle', 'Spoken', 'in', 'Armenian', 'Highlands', ',', 'Cilicia']

['Nyanga', 'Total', 'speakers', '27', ',', '000', '(', 'Biebuyck', '&', 'amp', ';', 'Matheene', '1970', ')', '\xe2', '\x80', '\x93', '150', ',', '000', '(', '1994', 'census', ')']

['Ancient', 'North', 'Arabian', 'Language', 'extinction', 'marginalized', 'by', 'Classical', 'Arabic', 'from', 'the', '7th', 'century']

['Ancient', 'North', 'Arabian', 'Language', 'extinction', 'marginalized', 'by', 'Classical', 'Arabic', 'from', 'the', '7th', 'century']

['Tidore', 'Total', 'speakers', '26', ',', '000', '(', '1981', 'Wurm', 'and', 'Hattori', ')']

['Tidore', 'Region', '0', '\xc2', '\xb0', '1', '\xe2', '\x80', '\xb2', 'N', '127', '\xc2', '\xb0', '44', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '0', '.', '017', '\xc2', '\xb0', 'N', '127', '.', '733', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '0', '.', '017', ';', '127', '.', '733']

['Udege', 'Total', 'speakers', '230', 'out', 'of', '1', ',', '657', 'in', 'ethnic', 'group', '(', '1991', ')']

['Ancient', 'North', 'Arabian', 'Language', 'extinction', 'marginalized', 'by', 'Classical', 'Arabic', 'from', 'the', '7th', 'century']

['Nabataean', 'Language', 'extinction', 'merged', 'with', 'Arabic', 'during', 'the', '4th', 'century', 'AD']

['Nabataean', 'Spoken', 'in', 'Formerly', 'in', 'the', 'southern', 'Middle', 'East']

['Nupe', 'Total', 'speakers', '800', ',', '000', '(', '1990', ')']

['Nupe', 'Dialects', 'Nupe', 'Tako', '(', 'Bassa', 'Nge', ')']

['English', 'Total', 'speakers', 'Second', 'language', ':', '199', 'million', '\xe2', '\x80', '\x93', '1', '.', '4', 'billion']

['English', 'Total', 'speakers', 'Overall', ':', '500', 'million', '\xe2', '\x80', '\x93', '1', '.', '8', 'billion']

['English', 'Pronunciation', '/', '\xcb', '\x88', '\xc9', '\xaa', '\xc5', '\x8b', '\xc9', '\xa1', 'l', '\xc9', '\xaa', '\xca', '\x83', '/', '1']

['Sentinel', 'Total', 'speakers', 'perhaps', '100', '\xe2', '\x80', '\x93', '250']

['Tsakhur', 'Total', 'speakers', '22', ',', '843']

['Tsakhur', 'Writing', 'system', 'Latin', 'alphabet', 'in', 'Azerbaijan', ',', 'Cyrillic', 'alphabet', 'in', 'Russia']

['Lahu', 'Total', 'speakers', '577', ',', '178']

['Lahu', 'Official', 'language', 'in', 'Lancang', 'Lahu', 'Autonomous', 'County', ',', 'Yunnan']

['Xibe', 'Total', 'speakers', '30', ',', '000']

['Ngala', 'Total', 'speakers', 'few', 'native', ',', '3', '.', '5', 'million', 'as', 'trade', 'language', '(', '1991', ')']

['Ipai', 'Total', 'speakers', '25', '(', '1994', ')']

['Ipai', 'Region', 'San', 'Diego', 'County', ',', 'California']

['Kumeyaay', 'Total', 'speakers', '50', '(', 'Hinton', '1994', ')']

['Tiipai', 'Total', 'speakers', '&', 'lt', ';', '200']

['Sui', 'Total', 'speakers', '346', ',', '000', '(', '1990', ')']

['Sui', 'Region', 'Guizhou', '(', '93', '%', ')']  
 ['Uru', 'Spoken', 'in', 'Bolivian', 'Andes', 'Ethnicity']  
 ['Digo', 'Total', 'speakers', 'estimated', '360', ',', '000']  
 ['Digo', 'Language', 'codes', 'ISO', '639-2']  
 ['Brianz\xc3\xb6\xc3\xb6/Brianzoeu', 'Total', 'speakers', '~', '300', ',', '000']  
 ['Canz\xc3\xa9s', 'Total', 'speakers', '~', '2', ',', '000']  
 ['Numbami', 'Total', 'speakers', '270', '(', '1978', 'McElhanon', ')']  
 ['Numbami', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Chakma', 'Total', 'speakers', '612', ',', '207']  
 ['Chakma', 'Total', 'speakers', '312', ',', '207', 'in', 'Bangladesh', '(', '2000', 'WCD', ')', ',', '300', ',', '000', 'in', 'India', '(', '1987', ')']  
 ['Scandoromani', 'Total', 'speakers', '31', ',', '000']  
 ['Scandoromani', 'Official', 'language', 'in', 'Norway', '(', '1993', ')']  
 ['Scandoromani', 'Official', 'language', 'in', 'Sweden', '(', '1999', ')']  
 ['Portuguese', 'Total', 'signers', 'many', 'of', 'the', '8000', 'deaf', '(', '1986', ')']  
 ['Portuguese', 'Language', 'codes', 'ISO', '639-3', 'psr', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['\xc3\x93pata', '(', 'also', ':', 'Language', 'extinction', 'pronounced', 'extinct', 'ca', '.', '1930', 'but', '15', 'speakers', 'were', 'found', 'by', 'INI', 'in', '1993']  
 ['\xc3\x93pata', 'Spoken', 'in', 'Mexico', '(', 'traditionally', 'Sonora', 'state', '-', 'now', 'D', '.', 'F', '.', ')']  
 ['Tepehua', 'Total', 'speakers', '10', ',', '000']  
 ['Tepehua', 'Language', 'codes', 'ISO', '639-3', ':', 'tee', '\xe2', '\x80', '\x93', 'Tepehua', 'of', 'Huehuetla']  
 ['Tepehua', 'Language', 'codes', 'ISO', '639-3', 'tpp', '\xe2', '\x80', '\x93', 'Tepehua', 'of', 'Pisaflores']  
 ['Tepehua', 'Language', 'codes', 'ISO', '639-3', 'tpt', '\xe2', '\x80', '\x93', 'Tepehua', 'of', 'Tlachichilco']  
 ['French', 'Language', 'codes', 'Linguasphere', '51-AAC-cd', '(', 'varieties']  
 ['French', 'Language', 'codes', 'Linguasphere', ':', '51-AAC-cda', 'to', '-cdd', ')']  
 ['Aguaruna', 'Total', 'speakers', '38', ',', '290', '(', '2000', 'WCD', ')']  
 ['Amahuaca', 'Total', 'speakers', '220-250', '(', '2', ',', '000', 'SIL', ')']  
 ['Amarakaeri', 'Total', 'speakers', '500', '(', '1987', ',', 'SIL', ')']  
 ['Lampung', 'Total', 'speakers', '1', ',', '500', ',', '000', '(', '1981', ')']  
 ['Arabela', 'Total', 'speakers', '50', '(', '2002', 'SIL', ')']  
 ['Ash\xc3\xa1ninka', 'Total', 'speakers', '23', ',', '750', '-', '28', ',',

'500', '(', '2000', 'SIL', ')']  
 ['Ash\xc3\xa9ninka', 'Total', 'speakers', '12', ',', '000', '(', '2002', 'SIL', ')']  
 ['Ash\xc3\xa9ninka', 'Total', 'speakers', '5', ',', '500', '(', '2001', ',', 'SIL', ')']  
 ['Ash\xc3\xa9ninka', 'Total', 'speakers', '12', ',', '000', '(', '2001', 'SIL', ')']  
 ['Ash\xc3\xa9ninka', 'Total', 'speakers', '13', ',', '000', '(', '2002', 'SIL', ')']  
 ['Ash\xc3\xa9ninka', 'Total', 'speakers', '7', ',', '212', '(', '2001', 'SIL', ')']  
 ['Awngi', 'Total', 'speakers', '488', ',', '507']  
 ['Khwarshi', 'Total', 'speakers', '8', ',', '500', '(', '2009', ')']  
 ['Khwarshi', 'Total', 'speakers', '1', ',', '870', '(', '2002', ')']  
 ['Judeo-Moroccan', 'Total', 'speakers', '258', ',', '925']  
 ['Mazahua', 'Total', 'speakers', 'ca', '350', ',', '000']  
 ['Mazahua', 'Language', 'codes', 'ISO', '639-3', 'maz', '\xe2', '\x80', '\x93', 'Central', 'Mazahua']  
 ['Cahuarano', 'Total', 'speakers', '5', '(', '1976', 'SIL', ')', ',', 'may', 'be', 'extinct']  
 ['Spaniard', 'Total', 'signers', '102', ',', '000', '(', '1994', ')']  
 ['Spaniard', 'Language', 'codes', 'ISO', '639-3', 'ssp', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Catalan', 'Total', 'signers', '18', ',', '000', '(', '1994', ')']  
 ['Catalan', 'Language', 'codes', 'ISO', '639-3', 'csc', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Candoshi', 'Total', 'speakers', '3', ',', '000', '(', '1981', 'SIL', ')']  
 ['Heiltsuk-Oowekyala', 'Spoken', 'in', 'Northern', 'Central', 'Coast']  
 ['Capanahua', 'Total', 'speakers', '387', '(', '2000', ',', 'WCD', ')']  
 ['Makah', 'Language', 'extinction', '2002', '.', 'Last', 'native', 'speaker', 'was', 'Ruth', 'E', '.', 'Claplanhoo']  
 ['Makah', 'Region', 'Northwestern', 'corner', 'of', 'the', 'Olympic', 'Peninsula', 'of', 'Washington', 'state', ',', 'on', 'the', 'south', 'side', 'of', 'the', 'Strait', 'of', 'Juan', 'de', 'Fuca']  
 ['Squamish', 'Total', 'speakers', '\xe2', '\x89', '\x88', '12-15']  
 ['Caquinte', 'Total', 'speakers', '300', '(', '2000', ',', 'SIL', ')']  
 ['Klamath', 'Region', 'Southern', 'Oregon', 'and', 'northern', 'California', 'Ethnicity']  
 ['Estonian', 'Total', 'signers', '4', ',', '500', '(', '1998', ')']  
 ['Estonian', 'Language', 'codes', 'ISO', '639-3', 'eso', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Chiquitano', 'Spoken', 'in', 'Santa', 'Cruz', ',', 'Bolivia']  
 ['Chiquitano', 'Total', 'speakers', '6', ',', '000', 'out', 'of', 'an',



'ethnic', 'population', 'of', '50', ',', '000']  
 ['Araki', 'Region', 'Araki', 'Island', ',', 'Espiritu', 'Santo']  
 ['Are'are', 'Total', 'speakers', '17', ',', '800', '(', '1999', ')']  
 ['Cashibo', 'Total', 'speakers', '5', ',', '000', '(', '1999', ',', 'Ethnologue', ')']  
 ['Bafut', 'Total', 'speakers', '105', ',', '000', '(', '2005', ')']  
 ['Eastern', 'Lombard', 'Total', 'speakers', '~', '1', ',', '500', ',', '000']  
 ['Eastern', 'Language', 'codes', 'ISO', '639-3', 'lmo', '\xe2', '\x80', '\x93', 'Lombard', 'language']  
 ['Eastern', 'Language', 'codes', 'Linguasphere', '51-AAA-oda', ';', '-odb', ';', '-odc']  
 ['Banjarese', 'Total', 'speakers', '5', ',', '900', ',', '000']  
 ['Banjarese', 'Region', 'South', 'Kalimantan', '(', 'Indonesia', ')', ',', 'Malaysia']  
 ['Chipaya', 'Total', 'speakers', '1200', 'in', '1995']  
 ['Cashinahua', 'Total', 'speakers', '2', ',', '000', '(', '2003', ',', 'Ethnologue', ')']  
 ['Cantabrian', 'Region', 'Autonomous', 'community', 'of', 'Cantabria', 'and', 'Asturian', 'municipalities', 'of', 'Pe', '\xc3', '\xb1', 'amellera', 'Alta', ',', 'Pe', '\xc3', '\xb1', 'amellera', 'Baja', 'and', 'Ribadedeva']  
 ['Nzema', 'Total', 'speakers', 'over', '300', ',', '000']  
 ['Early', 'Middle', 'Japanese', 'Language', 'extinction', 'Evolved', 'into', 'Late', 'Middle', 'Japanese', 'at', 'the', 'end', 'of', 'the', '12th', 'century']  
 ['Classical', 'Gaelic', 'Language', 'extinction', '18th', 'century', '(', 'replaced', 'by', 'Modern', 'Gaelic', ')']  
 ['Wagiman', 'Total', 'speakers', '10', '(', '2000', ')']  
 ['Wagiman', 'Language', 'codes', 'ISO', '639-3', 'Location', '(', 'in', 'red', ')', 'of', 'land', 'traditionally', 'associated', 'with', 'Wagiman', '.', 'Dark', 'grey', 'area', 'shows', 'highlighted', 'area', 'on', 'a', 'map', 'of', 'Australia']  
 ['Marovo', 'Total', 'speakers', '8', ',', '094', '(', 'As', 'of', '1999', ')']  
 ['Susuami', 'Spoken', 'in', 'Upper', 'Watut', 'valley']  
 ['Susuami', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Susuami', 'Total', 'speakers', '10', '(', '2000', ')']  
 ['Cowlitz', 'Total', 'speakers', 'extinct', '(', 'ethnic', 'population', ':', '200', ')']  
 ['Kwaza', 'Total', 'speakers', '25', '(', 'as', 'of', '2005', ')']  
 ['Javindo', 'Language', 'extinction', '20th', 'century']  
 ['Anejom\xcc\x83', 'Total', 'speakers', '600', '(', 'as', 'of', '1983', ')']  
 ['Judeo-Yemeni', 'Total', 'speakers', '51', ',', '000']  
 ['Ambai', 'Total', 'speakers', '10', ',', '000', '(', '2000', 'WCD', ')']

['Ugong', 'Total', 'speakers', '80', '(', '2000', 'David', 'Bradley', ')']  
 ['Phuan', 'Total', 'speakers', '204', ',', '704', '(', '2000', ')']  
 ['Phuan', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Chamicuro', 'Total', 'speakers', '2', '(', '2000', ',', 'Adelaar', ')']  
 ['Boon', 'Total', 'speakers', '59', '(', '2000', ')']  
 ['Boon', 'Region', 'Jilib', 'District', ',', 'Middle', 'Jubba']  
 ['Jurruru', 'Language', 'extinction', '2', 'speakers', 'left', 'in',  
 '1967']  
 ['Jurruru', 'Language', 'extinction', 'No', 'speakers', 'by', '1986']  
 ['Jurruru', 'Spoken', 'in', 'Pilbara', ',', 'Western', 'Australia']  
 ['Totoli', 'Spoken', 'in', 'parts', 'of', 'Central', 'Sulawesi', ',',  
 'Indonesia']  
 ['Totoli', 'Total', 'speakers', '25-30', ',', '000']  
 ['Aigon', 'Spoken', 'in', 'parts', 'of', 'West', 'New', 'Britain',  
 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Aigon', 'Total', 'speakers', '2000', '(', '2003', 'SIL', ')']  
 ['Miu', 'Spoken', 'in', 'southwest', 'interior', 'of', 'West', 'New',  
 'Britain', 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Miu', 'Total', 'speakers', '500', '(', '1998', 'NTM', ')']  
 ['Kaulong', 'Spoken', 'in', 'parts', 'of', 'West', 'New', 'Britain',  
 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Kaulong', 'Total', 'speakers', '4000', '(', '2000', 'Tryon', ')']  
 ['Sengseng', 'Spoken', 'in', 'southwest', 'interior', 'of', 'West', 'New',  
 'Britain', 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Sengseng', 'Total', 'speakers', '1750', '(', '2003', 'SIL', ')']  
 ['Karore', 'Spoken', 'in', 'parts', 'of', 'West', 'New', 'Britain',  
 'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Karore', 'Total', 'speakers', '550', '(', '2003', 'SIL', ')']  
 ['Karnai', 'Spoken', 'in', 'a', 'handful', 'of', 'villages', 'in',  
 'Morobe', 'and', 'Madang', 'Provinces', ',', 'Papua', 'New', 'Guinea']  
 ['Karnai', 'Total', 'speakers', '915', '(', '2000', 'D', '.', 'Tryon', ')']  
 ['Malasanga', 'Spoken', 'in', 'Morobe', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Malasanga', 'Total', 'speakers', '900', '(', '2000', 'census', ')']  
 ['Maleu-Kilenge', 'Spoken', 'in', 'West', 'New', 'Britain', 'Province']  
 ['Maleu-Kilenge', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Maleu-Kilenge', 'Total', 'speakers', '5200', '(', '1983', 'census', ')']  
 ['Amanay', 'Total', 'speakers', '60', '(', '2000', ',', 'C', '.',  
 'Jensen', ')']  
 ['Anamb', 'Total', 'speakers', '7', '(', '1991', 'SIL', ')']  
 ['Apiac', 'Total', 'speakers', '2', '(', '1986', 'Rodrigues', ')']  
 ['Hadrami', 'Total', 'speakers', '1', ',', '000', ',', '000']  
 ['Coeur', 'd'Alene', 'Total', 'speakers', '5', '(', '1999', 'R', '.',

'McDonald', ')]  
 ['Arikap\xc3\xba', 'Total', 'speakers', '6', '(', '1998', 'SIL', ')]'  
 ['Arutani', 'Total', 'speakers', '19', '(', '1986', 'and', '2001', 'SIL', ')]'  
 ['Aur\xc3\xa1', 'Total', 'speakers', '2', '(', '2004', 'SIL', ')]'  
 ['Western', 'Abnaki', 'Language', 'extinction', '2009']  
 ['Hajong', 'Total', 'speakers', '19', ',', '000', 'in', 'India', '(', '1997', ')', ',', '31', ',', '381', '(', '2001', ')', 'Meghalaya', ',', 'India']  
 ['Hajong', 'Region', 'Arunachal', 'Pradesh', 'and', 'West', 'Bengal']  
 ['Halbi', 'Total', 'speakers', '500', ',', '000', '(', '2000', ')]'  
 ['Harsusi', 'Total', 'speakers', '1', ',', '000', 'to', '2', ',', '000', '(', '1998', 'H', '.', 'Mutzafi', ')]'  
 ['Karakhawana', 'Total', 'speakers', '40', '(', '1995', 'SIL', ')]'  
 ['Karipun\xc3\xa1', 'Total', 'speakers', '12', 'to', '15', '(', '2000', 'SIL', ')]  
 ['Katawixi', 'Total', 'speakers', 'extinct', '']  
 ['Kanamar\xc3\xad', 'Total', 'speakers', '647', '(', '1995', 'SIL', ')]'  
 ['Katuk\xc3\xadna', 'Total', 'speakers', '1', '(', '1976', 'SIL', ')]'  
 ['Kinamigin', 'Total', 'speakers', '60', ',', '000', '(', '1973', ')]'  
 ['Kayort', 'Total', 'speakers', '22', ',', '000', '(', '2002', ')]'  
 ['Kharia', 'Region', 'Jharkhand', ',', 'Orissa', 'and', 'West', 'Bengal', 'states']  
 ['Guatemalan', 'Sign', 'Language', 'Region', 'Central', 'America']  
 ['Guatemalan', 'Language', 'codes', 'ISO', '639-3', 'gsm', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Kreye', 'Total', 'speakers', '30', '(', '1995', 'SIL', ')]'  
 ['Leco', 'Total', 'speakers', '20', '(', '2001', 'Simon', 'van', 'de', 'Kerke', ')]'  
 ['Mal', 'Paharia', 'Total', 'speakers', '51', ',', '000', 'to', '71', ',', '000', '(', '1994', ')]'  
 ['Mal', 'Region', 'Jharkhand', 'and', 'West', 'Bengal']  
 ['Mapoyo', 'Total', 'speakers', '3', '(', '2000', 'Muller', ')]'  
 ['Nahari', 'Total', 'speakers', '108', '(', '1961', 'census', ')]'  
 ['Nk\xe1\xbb\x8d\xe1\xbb\x8d\xe1\xbb\x8d', 'Total', 'speakers', '4', ',', '550', '(', '1989', 'UBS', ')]'  
 ['Pacahuara', 'Total', 'speakers', '17', '(', '2000', 'W', '.', 'Adelaar', ')]'  
 ['Ch\xc3\xa1cobo', 'Total', 'speakers', '550', '(', '2000', 'SIL', ')]'  
 ['P\xc3\xa9mono', 'Total', 'speakers', '1', '(', '2000', 'M-C', 'Mattei', 'Muller', ')]'  
 ['Reyesano', 'Total', 'speakers', 'Possibly', 'a', 'few', 'speakers', '(', '2000', 'W', '.', 'Adelaar', ')]'  
 ['Sikiana', 'Total', 'speakers', '48', '(', '1986', 'SIL', ',', '2001', 'Carlin', ')]'

['Yabarana', 'Total', 'speakers', '20', 'to', '50', '(', '1977',  
 'Migliazza', ')']  
 ['Labu', 'Total', 'speakers', '1', ',', '600', '(', '1989', 'SIL', ')']  
 ['Labu', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Macagu\xc3\xa1n', 'Total', 'speakers', '405', '(', '1993', 'census', ')']  
 ['Guahibo', 'Total', 'speakers', '23', ',', '006', 'in', 'Colombia', '(',  
 '2005', 'census', ')', ',', '8', ',', '428', 'in', 'Venezuela', '(',  
 '2001', 'census', ')']  
 ['Cuiba', 'Total', 'speakers', '2', ',', '343', 'in', 'Colombia', '(',  
 '1993', 'census', ')', ',', '650', 'in', 'Venezuela', '(', '1995', 'SIL',  
 ')']  
 ['Guayabero', 'Total', 'speakers', '1', ',', '237', '(', '1993', 'census',  
 ')']  
 ['Guayabero', 'Region', 'Upper', 'Guaviar', '\xc3', '\xa9', 'River']  
 ['Playero', 'Total', 'speakers', '244', '(', '2000', 'WCD', ')']  
 ['Playero', 'Language', 'codes', 'ISO', '639-2']  
 ['Awad', 'Bing', 'Total', 'speakers', '1', ',', '150', '(', '2000', 'SIL',  
 ')']  
 ['Awad', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Mindiri', 'Total', 'speakers', '80', '(', '2000', 'Wurm', ')']  
 ['Mindiri', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Mindiri', 'Region', 'along', 'Rai', 'Coast', ',', 'Madang', 'Province']  
 ['Wab', 'Total', 'speakers', '120', '(', '2000', 'Wurm', ')']  
 ['Wab', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Wab', 'Region', 'Wab', 'and', 'Sau'i', 'villages']  
 ['Marik', 'Total', 'speakers', '3', ',', '500', '(', '1998', 'SIL', ')']  
 ['Marik', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Bil', 'Bil', 'Total', 'speakers', '1', ',', '246', '(', '2003', 'SIL',  
 ')']  
 ['Bil', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Gedaged', 'Spoken', 'in', 'Madang', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Gedaged', 'Total', 'speakers', '6', ',', '954', '(', '2003', 'SIL', ')']  
 ['Matukar', 'Total', 'speakers', '434', '(', '2003', 'SIL', ')']  
 ['Matukar', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Takia', 'Total', 'speakers', '19', ',', '619', '(', '2003', 'SIL', ')']  
 ['Takia', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['U'wa', 'Spoken', 'in', 'Northern', 'slopes', 'of', 'the', 'Sierra',  
 'Nevada', 'del', 'Cocuy']  
 ['U'wa', 'Total', 'speakers', '1', ',', '800', '\xe2', '\x80', '\x93', '3',  
 ',', '600']  
 ['U'wa', 'Language', 'codes', 'ISO', '639-3', 'tuf', '\xe2', '\x80',  
 '\x93', 'Central', 'Tunebo']  
 ['U'wa', 'Language', 'codes', 'ISO', '639-3', 'tnb', '\xe2', '\x80',  
 '\x93', 'Western', 'Tunebo']  
 ['Amara', 'Spoken', 'in', 'parts', 'of', 'West', 'New', 'Britain',

'Province', ',', 'Papua', 'New', 'Guinea']  
 ['Amara', 'Total', 'speakers', '1170', '(', '2000', 'D', '.', 'Tryon', ')']  
 ['Lamogai', 'Total', 'speakers', '3', ',', '653', '(', '1980', 'Johnston', ')']  
 ['Lamogai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Lamogai', 'Region', 'parts', 'of', 'West', 'New', 'Britain', 'Province']  
 ['Mouk-Aria', 'Total', 'speakers', '626', '(', '1982', 'SIL', ')']  
 ['Mouk-Aria', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Mouk-Aria', 'Region', 'coastal', 'West', 'New', 'Britain', 'Province']  
 ['Shabaki', 'Total', 'speakers', '10', ',', '000-20', ',', '000']  
 ['Polar', 'Inuit', 'Total', 'speakers', 'approximately', '1000']  
 ['Bukawa', 'Total', 'speakers', '9', ',', '694', '(', '1978', 'McElhanon', ')']  
 ['Bukawa', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Bukawa', 'Region', 'Huon', 'Gulf', ',', 'Morobe', 'Province']  
 ['Yabem', 'Total', 'speakers', '2', ',', '084', '(', '1978', 'McElhanon', ')']  
 ['Yabem', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Yabem', 'Region', 'Huon', 'Gulf', ',', 'Morobe', 'Province']  
 ['Kela', 'Total', 'speakers', '2', ',', '420', '(', '1997', ')']  
 ['Kela', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kela', 'Region', 'Huon', 'Gulf', ',', 'Morobe', 'Province']  
 ['Mangareva', 'Total', 'speakers', '1', ',', '600', '(', '1987', ')']  
 ['Mangareva', 'Region', 'Gambier', 'Islands', ',', 'Mangareva', 'Island']  
 ['Mangareva', 'Language', 'codes', 'ISO', '639-2']  
 ['Austral', 'Total', 'speakers', '8', ',', '000', '(', '1987', ')']  
 ['Penrhyn', 'Total', 'speakers', '600', '(', '1981', 'Wurm', 'and', 'Hattori', ')']  
 ['Penrhyn', 'Region', 'Penrhyn', 'Island', ',', 'Northern', 'Cook', 'Islands']  
 ['Rakahanga-Manihiki', 'Total', 'speakers', '5000', '(', '1981', 'Wurm', 'and', 'Hattori', ')']  
 ['Maay', 'Spoken', 'in', 'Kenya', 'significant', 'communities', 'in', 'North', 'America']  
 ['Gone', 'Dau', 'Total', 'speakers', '686', '(', '2000', ')']  
 ['Gone', 'Region', 'Gone', 'and', 'Dau', 'Islands', ',', 'Eastern', 'Fiji']  
 ['Lauan', 'Total', 'speakers', '16', ',', '000', '(', '1981', 'P', '.', 'Geraghty', ')']  
 ['Lomaiviti', 'Total', 'speakers', '1', ',', '627', '(', '2000', 'WCD', ')']  
 ['Kachari', 'Total', 'speakers', '~', '59', ',', '000', '(', 'as', 'of', '1997', ')']  
 ['Western', 'Fijian', 'Total', 'speakers', '57', ',', '000', '(', '1977', 'Lincoln', ')']  
 ['Namosi-Naitasiri-Serua', 'Total', 'speakers', '1', ',', '627', '(', '2000', 'WCD', ')']

['Vangunu', 'Total', 'speakers', '907', '(', '1999', 'SIL', ')']  
 ['Duke', 'Total', 'speakers', '2', ',', '312', '(', '1999', 'SIL', ')']  
 ['Botolan', 'Total', 'speakers', '32', ',', '867']  
 ['Botolan', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Ghanongga', 'Total', 'speakers', '2', ',', '508', '(', '1999', 'SIL',  
 ')']  
 ['Hoava', 'Total', 'speakers', '459', '(', '1999', 'SIL', ')']  
 ['Hoava', 'Region', 'Marovo', 'Lagoon', ',', 'New', 'Georgia', 'Island']  
 ['Ambala', 'Total', 'speakers', '~', '2', ',', '000']  
 ['Kusaghe', 'Total', 'speakers', '2', ',', '395', '(', '1999', 'SIL', ')']  
 ['Kusaghe', 'Region', 'north', 'New', 'Georgia', 'Island']  
 ['Mag-indi/Baloga', 'Total', 'speakers', '~', '5', ',', '000']  
 ['Roviana', 'Total', 'speakers', '9', ',', '871', '(', '1999', 'SIL', ')']  
 ['Roviana', 'Region', 'north', 'central', 'New', 'Georgia', 'Island']  
 ['Simbo', 'Total', 'speakers', '2', ',', '701', '(', '1999', 'SIL', ')']  
 ['Ughele', 'Total', 'speakers', '1', ',', '202', '(', '1999', 'SIL', ')']  
 ['Sambal', 'Total', 'speakers', '~', '200', ',', '000']  
 ['Sambal', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Uneapa', 'Total', 'speakers', '10', ',', '000', '(', '1998', 'SIL', ')']  
 ['Uneapa', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Uneapa', 'Region', 'Bali', 'Island', ',', 'West', 'New', 'Britain']  
 ['Mag-antsi', 'Total', 'speakers', '~', '4', ',', '200']  
 ['Abellen', 'Total', 'speakers', '~', '3', ',', '500']  
 ['Hausa', 'Region', 'Northern', 'states', 'of', 'Nigeria', '(', 'Hausa-  
 speaking', 'areas', ')']  
 ['Hausa', 'Language', 'codes', 'ISO', '639-3', 'hsl', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Vivaro-Alpine', 'Language', 'codes', 'Linguasphere', '51-AAA-gf', '&',  
 'amp', ';', '51-AAA-gg']  
 ['Marivele\xc3\xb1o', 'Total', 'speakers', '~', '500']  
 ['Guatuso', 'Total', 'speakers', '&', 'lt', ';', '750']  
 ['Adabe', 'Spoken', 'in', 'Atauro', 'Island', ',', 'East', 'Timor']  
 ['Adabe', 'Total', 'speakers', '1', ',', '000', '(', 'as', 'of', '1981',  
 ')']  
 ['Jewish', 'Babylonian', 'Aramaic', '(', 'ca', '.', '200-1200', 'CE', ')',  
 'Language', 'extinction', 'about', '1200', 'CE']  
 ['Jewish', 'Spoken', 'in', 'Ancient', 'Near', 'East']  
 ['sign', 'Language', 'codes', 'ISO', '639-3', 'sls', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Bali', 'Total', 'speakers', '100', ',', '000', '(', 'as', 'of', '2006',  
 ')']  
 ['Gojri/Gujari', 'Total', 'speakers', '20', 'million', '(', '2010', ')']

['Southern', 'Italian', 'Total', 'speakers', 'approx', '.', '11',  
 'million']  
 ['Chayahuita', 'Total', 'speakers', '11', ',', '384', '(', '2000', ',', 'WCD', ')']  
 ['Cocama', 'Total', 'speakers', '2', ',', '050', '(', '2000', ',', 'W',  
 '.', 'Aldeaar', ',', 'SIL', ')']  
 ['Mursi', 'Total', 'speakers', '7', ',', '500', '(', '2007', 'census', ')']  
 ['Rama', 'Total', 'speakers', 'approx', '.', '36']  
 ['Klao', 'Total', 'speakers', '223', ',', '000', '(', '2003', ')']  
 ['Iteri', 'Total', 'speakers', '475', '(', '2003', 'SIL', ')']  
 ['Iteri', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Iteri', 'Region', '4', '\xc2', '\xb0', '7', '\xe2', '\x80', '\xb2', 'S',  
 '141', '\xc2', '\xb0', '35', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb',  
 '\xbf', '/', '\xef', '\xbb', '\xbf', '4', '.', '117', '\xc2', '\xb0', 'S',  
 '141', '.', '583', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-4',  
 '.', '117', ';', '141', '.', '583']  
 ['Mantion', 'Total', 'speakers', '12', ',', '000', '(', '1987', ')']  
 ['Arikara', 'Language', 'codes', 'ISO', '639-3', 'ari', 'Arikara',  
 'language', 'distribution']  
 ['Samoan', 'Plantation', 'Pidgin', 'Language', 'extinction', 'Effectively',  
 'extinct']  
 ['Alutiiq,', 'Pacific', 'Gulf', 'Yupik', 'Total', 'speakers',  
 'approximately', '400']  
 ['Alutiiq,', 'Region', 'coastal', 'Alaska', '(', 'Alaska', 'Peninsula',  
 'to', 'Prince', 'William', 'Sound']  
 ['Central', 'Alaskan', 'Yup'ik", 'Total', 'speakers', 'more', 'than', '14',  
 ',', '000']  
 ['Bribri', 'Total', 'speakers', '11', ',', '000']  
 ['Wom', 'Spoken', 'in', 'East', 'Sepik', ',', 'Papua', 'New', 'Guinea']  
 ['Wom', 'Total', 'speakers', '4', ',', '264', '(', 'as', 'of', '2003', ')']  
 ['Sudanese', 'Spoken', 'in', 'Sudan', '(', 'Southern', 'Sudan', 'is',  
 'the', 'official', 'area', ')']  
 ['Old', 'Polish', 'Language', 'extinction', 'developed', 'into', 'Middle',  
 'Polish', 'by', 'the', '16th', 'century']  
 ['Old', 'Spoken', 'in', 'Central', 'and', 'Eastern', 'Europe']  
 ['Tetelcingo', 'Nahuatl', 'Total', 'speakers', '3', ',', '500', '(',  
 '1990', ')']  
 ['Makhuwa-Shirima', 'Total', 'speakers', '500', ',', '000']  
 ['Misantla', 'Totonac', 'Total', 'speakers', 'approx', '.', '&', 'lt', ';',  
 '500']  
 ['Rutul', 'Spoken', 'in', 'Southern', 'Dagestan', ',', 'Russia', ';',  
 'Azerbaijan']  
 ['Rutul', 'Total', 'speakers', '29', ',', '510', '(', '2002', ')']  
 ['Dagaare', 'Total', 'speakers', '&', 'gt', ';', '1', 'million']  
 ['Dagaare', 'Language', 'codes', 'ISO', '639-3', ':', 'dga', '\xe2',  
 '\x80', '\x93', 'Southern', 'Dagaare']

['Dagaare', 'Language', 'codes', 'ISO', '639-3', 'dgi', '\xe2', '\x80',  
 '\x93', 'Northern', 'Dagara']  
 ['Kadazan', 'Coastal', 'Total', 'speakers', 'first', 'language', ':',  
 '60', ',', '000', '(', '1986', 'SIL', '.', ')']  
 ['Western', 'Total', 'speakers', '30', '000']  
 ["Ts'\xc3\xbcn-Lao", 'Total', 'speakers', '10', ',', '000', '(', '1993',  
 ')']  
 ['Sawi', 'Total', 'speakers', '3', ',', '500', '(', '1993', ')']  
 ['Petats', 'Total', 'speakers', '2', ',', '000', '(', '1981', ')']  
 ['Petats', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Mojave', 'Total', 'speakers', '204', '(', '2000', ')']  
 ['Cocopah', 'Total', 'speakers', '350', '(', '1998', ')']  
 ['Yinjibarndi', 'Region', 'Roebourne', 'region', 'of', 'Western',  
 'Australia']  
 ['Colorado', 'River', 'Total', 'speakers', 'less', 'than', '2000']  
 ['Yerukala', 'Total', 'speakers', '69', ',', '533']  
 ['Batangas', 'Writing', 'system', 'Latin', '(', 'Tagalog', 'or',  
 'Filipino', 'variant', ')', ';']  
 ['Shirvani', 'Arabic', 'Language', 'extinction', 'Second', 'half', 'of',  
 'the', '19th', 'century']  
 ['Central', 'Language', 'codes', 'ISO', '639-3', ':', 'abh', '\xe2',  
 '\x80', '\x93', 'Tajiki', 'Arabic']  
 ['Central', 'Language', 'codes', 'ISO', '639-3', 'auz', '\xe2', '\x80',  
 '\x93', 'Uzbeki', 'Arabic']  
 ['Rama', 'Cay', 'Creole', 'Total', 'speakers', 'approx', '.', '900']  
 ['Rama', 'Language', 'codes', 'ISO', '639-3', 'bzk', '\xe2', '\x80',  
 '\x93', 'Miskito', 'Coast', 'Creole']  
 ['North', 'Total', 'speakers', '6', ',', '300', ',', '000']  
 ['Tajiki', 'Total', 'speakers', '6', ',', '000']  
 ['Uzbeki', 'Arabic', 'Total', 'speakers', '700']  
 ['Judeo-Iraqi', 'Total', 'speakers', '100', ',', '100']  
 ['Aringa', 'Spoken', 'in', 'west', 'of', 'Ma'di']  
 ['Aringa', 'Total', 'speakers', '588', ',', '830', '(', '1991', 'census',  
 ')']  
 ['Twana', 'Total', 'speakers', '0', ',', 'or', '10', '(', 'USA', 'Census',  
 ',', '2000', ')']  
 ['Twana', 'Region', 'East', 'of', 'Puget', 'Sound', ',', 'Washington',  
 'state']  
 ['Labrador', 'Inuit', 'Pidgin', 'French', 'Language', 'extinction',  
 'About', '1760']  
 ['Qabiao', 'Total', 'speakers', '307', 'in', 'China', '(', '2002', ')']  
 ['Qabiao', 'Total', 'speakers', '307', 'in', 'Vietnam', '(', '1990', ')']  
 ['Qabiao', 'Language', 'codes', 'ISO', '639-2']  
 ['Amoy', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Amoy', 'Spoken', 'in', 'Taiwan', '(', 'where', 'it', 'is', 'known', 'as',  
 'Taiwanese', ')']



['Amoy', 'Spoken', 'in', 'Japan', '(', 'due', 'to', 'large', 'Taiwanese',  
 'community', 'in', 'Tokyo-Yokohama', 'Metropolitan', 'Area', ')']  
 ['Amoy', 'Spoken', 'in', 'and', 'other', 'areas', 'of', 'Min', 'Nan',  
 'and', 'Hoklo', 'settlement']  
 ['Wogamusin', 'Spoken', 'in', 'East', 'Sepik', 'Province']  
 ['Wogamusin', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Wogamusin', 'Total', 'speakers', '700', '(', 'as', 'of', '1998', ')']  
 ['Selayar', 'Total', 'speakers', '100', ',', '000', '(', '1986', ')']  
 ['Sangirese', 'Total', 'speakers', '255', ',', '000']  
 ['Bengkulu', 'Total', 'speakers', '2', '.', '4', 'million']  
 ['Huitoto', 'Language', 'codes', 'ISO', '639-3', 'hux', '\xe2', '\x80',  
 '\x93', 'N', '\xc3', '\xbc', 'pode']  
 ['Abom', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Abom', 'Total', 'speakers', '15', '(', '2002', 'census', ')']  
 ['Cisalpine', 'Gaulish', 'Language', 'extinction', 'ca', '.', '1st',  
 'century', 'BC', '?']  
 ["Sa'a", 'Spoken', 'in', 'South', 'Malaita', ',', 'Solomon', 'Islands']  
 ["Sa'a", 'Total', 'speakers', '11', ',', '519', '(', '1999', ')']  
 ["Dori'o", 'Spoken', 'in', 'South', 'Malaita', ',', 'Solomon', 'Islands']  
 ["Dori'o", 'Total', 'speakers', '2', ',', '406', '(', '1999', ')']  
 ['Eastern', 'Total', 'speakers', '410', ',', '000', '(', 'as', 'of',  
 '1991', ')']  
 ['Balochi', 'Total', 'speakers', '5', '.', '5', 'million', '(', '1998',  
 ',', 'Ethnologue', ')', 'not', 'including', 'Northern', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bgp', '\xe2', '\x80',  
 '\x93', 'Eastern', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bgn', '\xe2', '\x80',  
 '\x93', 'Western', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', 'bcc', '\xe2', '\x80',  
 '\x93', 'Southern', 'Balochi']  
 ['Balochi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Lola', 'Total', 'speakers', '830', '(', '1995', ')']  
 ['Temuan', 'Total', 'speakers', '22', ',', '162', '(', '2003', 'COAC', ')']  
 ['Temuan', 'Writing', 'system', 'no', 'native', 'script', ',', 'Latin',  
 'alphabet', 'is', 'used']  
 ['Sri', 'Total', 'speakers', '46', ',', '000']  
 ['Sri', 'Official', 'language', 'in', 'Sri', 'Lanka', '(', 'minority',  
 'language', ')']  
 ['Kunjen', 'Total', 'speakers', '20', '(', '1991', ')']  
 ['Kunjen', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Piro', 'Total', 'speakers', '4', ',', '000']  
 ['Wasco-Wishram', 'Total', 'speakers', '69', '(', '1990', 'census', ')',  
 ',', 'of', 'which', '7', 'monolingual']

['Lamba', 'Total', 'speakers', '220', ',', '000', 'in', 'Zambia', '(', '2006', ')', ',', 'unknown', 'but', 'smaller', 'number', 'in', 'DRC']  
 ['Isoko', 'Total', 'speakers', '423', '000', '(', '2001', ',', 'Ethnologue', ')']  
 ['Zulgo-Gemzek', 'Total', 'speakers', '26', ',', '000', '(', '2002', ')']  
 ['E', 'Total', 'speakers', '30', '000', '(', '1992', 'Ethnologue', '/', 'Edmondson', ')']  
 ['Canadian', 'Spoken', 'in', 'Canada', '(', 'primarily', 'Quebec', ',', 'Ontario', 'and', 'New', 'Brunswick', ',', 'but', 'present', 'throughout', 'the', 'country', ')', ';', 'smaller', 'numbers', 'in', 'emigrant', 'communities', 'in', 'New', 'England']  
 ['Canadian', 'Total', 'speakers', 'Smaller', 'numbers', 'in', 'the', 'U', '.', 'S']  
 ['Canadian', 'Official', 'language', 'in', 'Canada', '(', 'as', 'French', ')']  
 ['Iaai', 'Total', 'speakers', '1', ',', '562', '(', '1996', 'census', ')']  
 ['Huachipaeri', 'Total', 'speakers', '311', '(', '2000', 'WCD', ')']  
 ['Maritime', 'Total', 'signers', 'Few', '-', 'now', 'a', 'moribund', 'language']  
 ['Maritime', 'Language', 'codes', 'ISO', '639-3', 'nsr', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Amal', 'Total', 'speakers', '831', '(', '2003', 'SIL', ')']  
 ['Ari', 'Total', 'speakers', '50', '(', '2000', 'S', '.', 'Wurm', ')']  
 ['Ari', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Ari', 'Region', '7', '\xc2', '\xb0', '57', '\xe2', '\x80', '\xb2', 'S', '142', '\xc2', '\xb0', '24', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '7', '.', '95', '\xc2', '\xb0', 'S', '142', '.', '4', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-7', '.', '95', ';', '142', '.', '4']  
 ['Ari', 'Region', 'Ari', 'and', 'Serea', 'villages']  
 ['Ari', 'Region', 'Western', 'Province', '.', 'Coordinates']  
 ['Gogodala', 'Total', 'speakers', '10', ',', '000', '(', '1991', 'UBS', ')']  
 ['Gogodala', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Gogodala', 'Region', '8', '\xc2', '\xb0', '5', '\xe2', '\x80', '\xb2', 'S', '142', '\xc2', '\xb0', '52', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '8', '.', '083', '\xc2', '\xb0', 'S', '142', '.', '867', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-8', '.', '083', ';', '142', '.', '867']  
 ['Minica', 'Total', 'speakers', '1', ',', '705', '(', '2000', 'WCD', ')']  
 ['Murui', 'Total', 'speakers', '2', ',', '900', '(', '1995', 'SIL', ')']  
 ['N\xc3\xbcpcode', 'Language', 'codes', 'ISO', '639-3', 'Map', 'of', 'the', 'extent', 'of', 'the', 'N', '\xc3', '\xbc', 'pode', 'Huitoto', 'language']  
 ['I\xc3\xblapari', 'Total', 'speakers', '4', '(', '1999', 'SIL', ')']  
 ['Chantyal', 'Total', 'speakers', '2', ',', '000']

['Pisabo', 'Total', 'speakers', '500', '(', 'as', 'of', '2000', ')']  
 ['Chadian', 'Total', 'speakers', '1', ',', '140', ',', '000']  
 ['Talossan', 'Regulated', 'by', 'La', 'Comit', '\xc3', '\xa0', 'per', 'l', '\xc3', '\x9a', 'tzil', 'del', 'Glhe', '\xc3', '\xbe', '(', 'Kingdom', ')', ',', 'L'Icastola', '(', 'Republic', ')']  
 ['Talossan', 'Total', 'speakers', '5', 'fluent', ',', 'est', '.', '&', 'gt', ';', '50']  
 ['Ika', 'Total', 'speakers', '22', ',', '772', '(', 'as', 'of', '2000', ')']  
 ['Kullili', 'Spoken', 'in', 'Queensland', 'and', 'New', 'South', 'Wales', ',', 'Australia']  
 ['Lardil', 'Total', 'speakers', '2', '(', '2000', ')']  
 ['Ankave', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Ankave', 'Total', 'speakers', '1', ',', '600', '(', 'as', 'of', '1987', ')']  
 ['Kayardild', 'Total', 'speakers', '&', 'lt', ';', '10']  
 ['lingua', 'franca', 'Language', 'extinction', '19th', 'century']  
 ['Chintang', 'Total', 'speakers', '~', '5', ',', '000']  
 ['Anglo-Norman', 'Language', 'extinction', 'contributed', 'to', 'Middle', 'English']  
 ['Athpare', 'Spoken', 'in', 'Dhankuta', 'district', ',', 'Nepal']  
 ['Athpare', 'Total', 'speakers', '2', ',', '000']  
 ['Khaling', 'Total', 'speakers', '9', ',', '288']  
 ['Gorontalo', 'Total', 'speakers', '900', '.', '000']  
 ['Guaym\xc3\xad', 'Total', 'speakers', '133', ',', '092', '(', '1990', 'Panama', 'Census', ')']  
 ['Jingpho', 'Total', 'speakers', '900', ',', '000']  
 ['Jingpho', 'Spoken', 'in', 'Burma', '(', 'Myanmar', ')', ',', 'China']  
 ['Luba', 'Total', 'speakers', '1', ',', '505', ',', '000', '(', '1991', ')']  
 ['Central', 'Spoken', 'in', 'M', '\xc3', '\xa9', 'xico', '(', 'Hidalgo', ',', 'Veracruz', ',', 'San', 'Luis', 'Potos', '\xc3', '\xad', ')']  
 ['Central', 'Total', 'speakers', '200', ',', '000', '(', '2000', 'census', ')']  
 ['Mandar', 'Total', 'speakers', '200', ',', '000', '(', '1985', ')']  
 ['Highland', 'Total', 'speakers', '125', ',', '00', '(', '1983', ')']  
 ['Orizaba', 'Total', 'speakers', '120', ',', '000', '(', '1991', 'SIL', ')']  
 ['Bangl-me', 'Total', 'speakers', '2000', '+']  
 ['Western', 'Spoken', 'in', 'M', '\xc3', '\xa9', 'xico', '(', 'Hidalgo', 'and', 'San', 'Luis', 'Potos', '\xc3', '\xad', ')']  
 ['Western', 'Total', 'speakers', '400', ',', '000', '(', '1991', 'SIL', ')']  
 ['Mezquital', 'Total', 'speakers', '100', ',', '000', '(', '1990', 'Census', ')']  
 ['Highland', 'Total', 'speakers', '120', ',', '000', '(', '1982', 'SIL',

')]  
 ['Tzotzil', 'Total', 'speakers', '130', ',', '000', '(', '1990', 'Census',  
 ')']  
 ['Bachaj\xc3\xb3n', 'Total', 'speakers', '100', ',', '000', '(', '1993',  
 'SIL', ')']  
 ['Loun', 'Total', 'speakers', 'less', 'than', '100']  
 ['Bustocco', 'Total', 'speakers', '~', '50', ',', '000']  
 ['Comasco-Lecchese', 'Total', 'speakers', '~', '300', ',', '000']  
 ['Vallassinese', 'Total', 'speakers', '10', ',', '000']  
 ['Nuares', 'Total', 'speakers', '30', ',', '000']  
 ['Varesino', 'Total', 'speakers', '300', ',', '000']  
 ['Amuesha', 'Total', 'speakers', 'primary', 'language', ':', '9', ',', '831']  
 ['Cremonese', 'Total', 'speakers', '30', ',', '000']  
 ['Southwestern', 'Total', 'speakers', '300', ',', '000']  
 ['Tarantino', 'Total', 'speakers', '~', '300', ',', '000']  
 ['Lecchese', 'Total', 'speakers', '30', ',', '000']  
 ['Comasco', 'Total', 'speakers', '30', ',', '000']  
 ['Lugbara', 'Total', 'speakers', '1', ',', '040', ',', '000']  
 ['Lagh\xc3\xa9\xc3\xa9', 'Total', 'speakers', '30', ',', '000']  
 ['Pavese', 'Total', 'speakers', '30', ',', '000']  
 ['Zou', 'Total', 'speakers', '39', ',', '112']  
 ['Zou', 'Region', 'In', 'Myanmar', ':', 'Chin', 'State']  
 ['Zou', 'Language', 'codes', 'ISO', '639-2']  
 ["Zo'\xc3\xa9", 'Total', 'speakers', '150', '(', '1998', ')']  
 ["Zo'\xc3\xa9", 'Region', "Zo'", '\xc3', '\xa9', 'people']  
 ['Nkutu', 'Total', 'speakers', '40', ',', '000', '(', '1972', 'est', '.',  
 ')']  
 ['Nkutu', 'Region', 'Northern', 'Kasai', 'Oriental', 'Province']  
 ['Kara', 'Spoken', 'in', 'New', 'Ireland', 'Province', ',', 'Papua', 'New',  
 'Guinea']  
 ['Kara', 'Total', 'speakers', '5', ',', '000', '(', 'as', 'of', '1998',  
 ')']  
 ['Ashtiani', 'Total', 'speakers', '21', ',', '000']  
 ['Bolango', 'Total', 'speakers', '18', ',', '500']  
 ['Nangalami', 'Total', 'speakers', '5000', 'in', '1994']  
 ['Nangalami', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Puxian', '/', 'Xinghua', 'Total', 'speakers', '2', '.', '6', 'million']  
 ['Puxian', 'Region', 'Fujian', '(', 'Putian', ',', 'Xianyou', ',',  
 'Fuzhou', ',', 'Quanzhou', ')']  
 ['Central', 'Region', 'Central', 'Morocco', '\xe2', '\x80', '\x93',  
 'Middle', 'Atlas']  
 ['Mappila', 'Writing', 'system', 'Malayalam', 'script', '(', 'rarely',

'in', 'Arabi', 'Malayalam', 'script', ')']  
 ['Western', 'Brythonic', 'Language', 'extinction', 'N', '/', 'A']  
 ['Western', 'Spoken', 'in', 'North', 'West', 'England']  
 ['Maasina', 'Spoken', 'in', 'Mali', ';', 'minor', ':', 'Ivory', 'Coast',  
 ', ', 'Ghana']  
 ['Butuanon', 'Total', 'speakers', '35', ', ', '000']  
 ['Swiss', 'Total', 'speakers', '0', '.', '5', 'million']  
 ['Swiss', 'Official', 'language', 'in', 'Switzerland', '(', 'as',  
 'Italian', ')']  
 ['Solombala-English', 'Language', 'extinction', 'Late', '19th', 'century']  
 ['Tigak', 'Spoken', 'in', 'New', 'Ireland', 'Province', ', ', 'Papua',  
 'New', 'Guinea']  
 ['Tigak', 'Total', 'speakers', '6', ', ', '000', '(', 'as', 'of', '1991',  
 ')']  
 ['Wuvulu-Aua', 'Total', 'speakers', '~', '1', ', ', '500']  
 ['Wuvulu-Aua', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Indonesian', 'Language', 'codes', 'ISO', '639-3', 'inl', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Ayoreo', 'Total', 'speakers', '3', ', ', '771']  
 ['Ayoreo', 'Region', 'Santa', 'Cruz', 'department', '(', 'Bolivia', ')']  
 ['Chamacoco', 'Total', 'speakers', '1', ', ', '800']  
 ['Nyole', 'Total', 'speakers', '341', ', ', '000']  
 ['Kui', 'Total', 'speakers', '700', ', ', '000', '(', '1997', ')']  
 ['Muher', 'Total', 'speakers', '90', ', ', '000']  
 ['Koya', 'Total', 'speakers', '330', ', ', '000', '(', '1997', ')']  
 ['Koya', 'Language', 'codes', 'ISO', '639-3', 'kff', 'Image', ':', '250px']  
 ['Sauria', 'Total', 'speakers', '117', ', ', '000', '(', '2000', 'estimate',  
 ')']  
 ['U\u005beice', 'Language', 'extinction', 'suppressed', 'by', 'the',  
 'literary', 'language']  
 ['U\u005beice', 'Region', 'Stari', 'Vlah', '(', 'U', '\u005', '\xbe',  
 'ice', ')']  
 ['Hinukh', 'Total', 'speakers', '200', '\xe2', '\x80', '\x93', '500']  
 ['Banat', 'Spoken', 'in', 'Romania', '(', 'Banat', ', ', 'Transylvania',  
 ')', ', ', 'Serbia', '(', 'Vojvodina']  
 ['Banat', 'Total', 'speakers', 'form', '.', '8', ', ', '000']  
 ['Banat', 'Total', 'speakers', 'ca', '.', '15', ', ', '000']  
 ['Sanum\u003\xa1', 'Total', 'speakers', '5', ', ', '074']  
 ['Karamanli', 'Total', 'speakers', '400', '.', '000']  
 ['Sadri', 'Total', 'speakers', 'ca', '.', '2', 'million', '(', '1997', ')']  
 ['Sadri', 'Region', 'East', '&', 'amp', ';', 'South', 'Chotta', 'Nagpur',  
 'and', 'surrounding', 'states']  
 ['Nafusi', 'Total', 'speakers', '210', ', ', '000', 'inc', '.', 'Zuara',  
 '+', 'Tunisian', 'Berber']  
 ['Ningbo', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']

['Singaporean', 'Total', 'speakers', '2', ',', '502', ',', '209', '(',  
 '2000', 'est', '.', ')']  
 ['Kuku', 'Total', 'speakers', '26', ',', '400']  
 ['Noxilo', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Noxilo', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Mongo', 'Total', 'speakers', '(', 'native', ')', '400', ',', '000']  
 ['Tayo', 'Spoken', 'in', 'France', '(', 'New', 'Caledonia', ')']  
 ['Tayo', 'Total', 'speakers', '2', ',', '000', '(', '1996', 'C', '.',  
 'Corne', ')']  
 ['Gaddang', 'Total', 'speakers', '30', ',', '000']  
 ['Negeri', 'Sembilan', 'Malay', 'Total', 'speakers', '300', ',', '000']  
 ['Negeri', 'Region', 'Negeri', 'Sembilan', '(', 'Malaysia', ')']  
 ['Urhobo', 'Total', 'speakers', '500', ',', '000', '(', '1990', ')']  
 ['Southern', 'Yukaghir', 'Total', 'speakers', '10', '\xe2', '\x80', '\x93',  
 '50']  
 ['Mundari', 'Total', 'speakers', '70', ',', '000', 'to', '100', ',', '000']  
 ['Osing', 'Total', 'speakers', '481', ',', '852', '(', 'as', 'of', '2000',  
 ')']  
 ['Kusunda', 'Total', 'speakers', 'perhaps', '8']  
 ['Colombian', 'Language', 'codes', 'ISO', '639-3', 'csn', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Ormuri', 'Total', 'speakers', '1', ',', '000']  
 ['Beurla', 'Region', 'Scotland', 'Scottish', 'Highlands', '&', 'amp', ';;',  
 'Western', 'Isles']  
 ['Kallawaya', 'Total', 'speakers', '10-20', '(', '2nd', 'language', ')']  
 ['Zaghawa', 'Total', 'speakers', '186', ',', '834']  
 ['Tupuri', 'Total', 'speakers', '216', ',', '000']  
 ['Semelai', 'Total', 'speakers', 'ethnic', 'population', '6', ',', '400']  
 ['Bor\xc3\xb4ro', 'Total', 'speakers', '1020', '(', 'in', '1997', ')']  
 ['Rusenu', 'Language', 'extinction', 'Unknown', ',', 'likely', '20th',  
 'century']  
 ['Rusenu', 'Spoken', 'in', 'eastern', 'East', 'Timor']  
 ['Itawis', 'Total', 'speakers', '134', ',', '000', '(', '1990', ')']  
 ['Salvadoran', 'Signed', 'in', 'El', 'Salvador']  
 ['Salvadoran', 'Language', 'codes', 'ISO', '639-3', 'esn', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Onge', 'Language', 'codes', 'ISO', '639-3', 'oon', 'A', 'map', 'of',  
 'tribal', 'and', 'language', 'divisions', 'in', 'the', 'Andaman',  
 'Islands', 'prior', 'to', 'the', '1850s']  
 ['Boga', 'Total', 'speakers', '10', ',', '000', '(', '1990', ')']  
 ['Ga'anda', 'Total', 'speakers', '43', ',', '000', '(', '1992', ')']  
 ['Jara', 'Total', 'speakers', '46', ',', '251']

['Cibak', 'Total', 'speakers', '100', ',', '000', '(', '1993', ')']  
 ['Pandan', 'Bikol', 'Total', 'speakers', '122', ',', '000']  
 ['Pandan', 'Region', 'Pandan', 'and', 'north-east', 'Catanduanes']  
 ['Iriga', 'Bikol', 'or', 'Rinconada', 'Bikol', 'Total', 'speakers', '234', ',', '361', '(', 'Ethnologue', ',', '2000', ')']  
 ['Iriga', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')', ';']  
 ['Iriga', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Tswa', 'Total', 'speakers', '1', ',', '200', ',', '000']  
 ['Ronga', 'Total', 'speakers', '722', ',', '000']  
 ['Yeyi', 'Total', 'speakers', '24', ',', '000']  
 ['Lau', 'Total', 'speakers', '16', ',', '937', '(', '1999', 'SIL', ')']  
 ['Shark', 'Bay', 'Total', 'speakers', '450']  
 ['Shark', 'Region', 'Litaro', '(', 'Pilot', ')', 'Island', ',', 'Shark', 'Bay', 'coast']  
 ['Kamoro', 'Spoken', 'in', 'Middle', 'south', 'coast', 'of', 'Western', 'New', 'Guinea']  
 ['Yaur', 'Total', 'speakers', '350', '(', '1978', ')']  
 ['Banggai', 'Total', 'speakers', '125', ',', '000']  
 ['Iwam', 'Total', 'speakers', '3', ',', '000', '(', 'as', 'of', '1998', ')']  
 ['Yao', 'Total', 'speakers', '1', ',', '916', ',', '000']  
 ['Parkari', 'Total', 'speakers', '250', ',', '000', '(', '1995', ')']  
 ['Gnau', 'Total', 'speakers', '1', ',', '325', '(', '2000', 'census', ')']  
 ['Gnau', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Gnau', 'Region', '3', '\xc2', '\xb0', '35', '\xe2', '\x80', '\xb2', 'S', '142', '\xc2', '\xb0', '11', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '3', '.', '583', '\xc2', '\xb0', 'S', '142', '.', '183', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-3', '.', '583', ';', '142', '.', '183']  
 ['Juang', 'Total', 'speakers', '50', ',', '000', '(', '2000', ')']  
 ['Tai', 'Bueng', 'Total', 'speakers', '6', ',', '000']  
 ['Ngawun', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Kesen', 'Total', 'speakers', '74', ',', '000']  
 ['Ai-Cham', 'Total', 'speakers', '2', ',', '700', '(', '2000', ')']  
 ['Dhudhuroa', 'Language', 'extinction', 'Early', '20th', 'century']  
 ['Paniya', 'Total', 'speakers', '70', ',', '000', '(', '2003', ')']  
 ['Paniya', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Bushi', 'Total', 'speakers', '806', ',', '200']  
 ['Kedah', 'Total', 'speakers', '2', '.', '6', 'million']  
 ['Rikbakts', '\xc3\xa1', 'erigpaks', '\xc3', '\xa1', 'Total', 'speakers', '910', '(', 'in', '2001', ')']

['Abron', 'Total', 'speakers', '1', ',', '181', ',', '700']  
 ['Burgundian', 'Total', 'speakers', '~', '2000']  
 ['Burgundian', 'Language', 'codes', 'Linguasphere', '51-AAA-hk', '&',  
 'amp', ';', '51-AAA-hl']  
 ['Glosa', 'Date', 'founded', '1972', '\xe2', '\x80', '\x93', '1992']  
 ['Glosa', 'Category', '(', 'sources', ')', 'vocabulary', 'from', 'Latin',  
 'and', 'Greek', 'and', 'grammatical', 'influences', 'from', 'isolating',  
 'languages', 'and', 'creole', 'languages']  
 ['Loko', 'Total', 'speakers', '140', ',', '000']  
 ['Abui', 'Total', 'speakers', '16', ',', '000', '+']  
 ['Beray', 'Total', 'speakers', '11', ',', '200', '(', '2007', 'SIL',  
 'Ethnologue', ')']  
 ['Kako', 'Spoken', 'in', 'Cameroon', ',', 'Central', 'African', 'Republic']  
 ['Kako', 'Total', 'speakers', '120', ',', '000']  
 ['Lop', 'Region', 'Xinjiang', 'Uyghur', 'Autonomous', ',', 'Lop', 'County']  
 ['Daur', 'Total', 'speakers', '96', ',', '085', '(', 'as', 'of', '1999',  
 ')']  
 ['Achagua', 'Total', 'speakers', '400', '(', '1994', ')']  
 ['Bartangi', 'Total', 'speakers', '3', ',', '000']  
 ['Qimant', 'Total', 'speakers', '1', ',', '650', 'out', 'of', 'an',  
 'ethnic', 'population', 'of', '172', ',', '000', '(', '1994', 'Census',  
 ')']  
 ['Greenlandic', 'Norse', 'Language', 'extinction', 'by', 'the', 'late',  
 '15th', 'century', '(', '16th', 'century', 'at', 'the', 'latest', ')']  
 ['Greenlandic', 'Spoken', 'in', 'Greenland', ';', 'Western', 'Settlement',  
 'and', 'Eastern', 'Settlement']  
 ['Nukak', 'Total', 'speakers', 'ca', '.', '570']  
 ['Nukak', 'Official', 'language', 'in', 'In', 'Colombia', 'it', 'is',  
 'recognized', 'as', 'an', 'official', 'language', 'with', 'the', 'Nukak',  
 'territory']  
 ['Kawaiisu', 'Total', 'speakers', '8', 'to', '10']  
 ['Bonda', 'Total', 'speakers', '9', ',', '000', '(', '2002', 'SIL', ')']  
 ['Gataq', 'Total', 'speakers', '3', ',', '055', '(', '1991', 'census', ')']  
 ['Kpala', 'Total', 'speakers', '65', ',', '000']  
 ['Kpala', 'Region', 'Northwestern', 'corner', 'of', 'Congo', '(', 'DRC',  
 ')']  
 ['Cha\xe2\x80\x99', 'Total', 'speakers', '10', ',', '000', '(', '2004',  
 ')']  
 ['Mara', 'Total', 'speakers', 'India', ':', '47', ',', '000', '(', 'MADC',  
 ',', '2003', ')', 'Myanmar', ':', '37', ',', '000', '(', 'MEC', ',', 'BAM',  
 '2005', ')']  
 ['Wellemmedan', 'Total', 'speakers', '640', ',', '000', '(', '1998', ')']  
 ['Burji', 'Region', 'South', 'of', 'Lake', 'Chamo']  
 ['Gedeo', 'Total', 'speakers', '637', ',', '082']  
 ['Dirasha', 'Total', 'speakers', '15', ',', '336', 'of', 'Dirasha',  
 'dialect']



['Nambikwara', 'Total', 'speakers', '1', ',', '150', '(', 'in', '2000', ')']  
 ['Siwu', 'Total', 'speakers', '27', ',', '000']  
 ['Moronene', 'Total', 'speakers', '37', ',', '000']  
 ['Umatilla', 'Total', 'speakers', '50', '(', '1977', ')']  
 ['Prekmurian', 'Total', 'speakers', '80', ',', '000']  
 ['Ollari', 'Total', 'speakers', '15', ',', '000', 'Ollari', ',', '8', ',', '000', 'Kondekor']  
 ['Ollari', 'Language', 'codes', 'ISO', '639-3', ':', 'gdb', '\xe2', '\x80', '\x93', 'Ollari', '(', 'Gadaba', ')']  
 ['Late', 'Middle', 'Japanese', 'Language', 'extinction', 'Evolved', 'into', 'Modern', 'Japanese', 'in', 'the', '17th', 'century']  
 ['Dawera-Daweloor', 'Total', 'speakers', '1', ',', '500']  
 ['Ngaju', 'Total', 'speakers', '800', ',', '000']  
 ['Melanau', 'Total', 'speakers', '1', '500', '000']  
 ['Melanau', 'Region', 'Sarawak', 'and', 'neighboring', 'Brunei']  
 ['Buyang', 'Total', 'speakers', '2', ',', '772']  
 ['Buyang', 'Language', 'codes', 'ISO', '639-3', ':', 'yzg', '\xe2', '\x80', '\x93', 'E'ma', 'Buyang']  
 ['Phula', 'Total', 'speakers', '13', ',', '246']  
 ['Saintongeais', 'Language', 'codes', 'ISO', '639-3', 'Linguistic', 'area', 'of', 'Saintongeais', 'the', 'Charentes', 'and', 'the', 'Northern', 'part', 'of', 'Gironde']  
 ['Filipino', 'Language', 'codes', 'ISO', '639-3', 'psp', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Nguon', 'Total', 'speakers', '2', ',', '000', '(', '1981', ')']  
 ['Nguon', 'Writing', 'system', 'Latin', 'alphabet', '(', 'qu', '\xe1', '\xbb', '\x91', 'c', 'ng', '\xe1', '\xbb', '\xaf', ')']  
 ['Biak', 'Total', 'speakers', 'about', '30', '.', '000', 'speakers']  
 ['Taos', 'Total', 'speakers', '803', '(', '1980', ')']  
 ['Taos', 'Region', 'Taos', 'Pueblo', ',', 'New', 'Mexico']  
 ['Taos', 'Language', 'codes', 'ISO', '639-3', 'Location', 'of', 'city', 'of', 'Taos', '(', '', 'Taos', 'village', '', ')', 'in', 'Taos', 'County', 'New', 'Mexico', '.', 'Taos', 'Pueblo', 'is', 'about', 'two', 'miles', 'north', 'of', 'the', 'city', 'of', 'Taos']  
 ['Bugan', 'Total', 'speakers', '3', ',', '000']  
 ['Arim', 'Language', 'extinction', '18th', 'century']  
 ['Finnish', 'Total', 'speakers', '7', ',', '002', '(', 'of', 'which', '5', ',', '410', 'in', 'Finland', 'and', '1', ',', '592', 'in', 'Sweden', ')']  
 ['Afade', 'Total', 'speakers', '30', ',', '000']  
 ['Afade', 'Region', 'Nigeria', ';', 'Far', 'North', 'Province']  
 ['Lagwan', 'Total', 'speakers', '10', ',', '000']  
 ['Lagwan', 'Region', 'Far', 'North', 'Province', ',', 'Cameroon', ';', 'west', 'Chad']  
 ['Malgbe', 'Total', 'speakers', '6', ',', '000']

['Malgbe', 'Region', 'Far', 'North', 'Province', ',', ',', 'Cameroon', ';',  
 'west', 'Chad']  
 ['Maslam', 'Region', 'Far', 'North', 'Province', ',', ',', 'Cameroon', ';',  
 'west', 'Chad']  
 ['Mpade', 'Total', 'speakers', '18', ',', ',', '366']  
 ['Mpade', 'Region', 'Far', 'North', 'Province', ',', ',', 'Cameroon', ';',  
 'west', 'Chad']  
 ['Mser', 'Total', 'speakers', '6', ',', ',', '000']  
 ['Mser', 'Region', 'Far', 'North', 'Province', ',', ',', 'Cameroon', ';',  
 'west', 'Chad']  
 ['Cakfem-Mushere', 'Total', 'speakers', '5', ',', ',', '000']  
 ['Jorto', 'Total', 'speakers', '17', ',', ',', '284']  
 ['Kofyar', 'Total', 'speakers', '109', ',', ',', '943']  
 ['Miship', 'Total', 'speakers', '6', ',', ',', '000']  
 ['Mwaghavul', 'Total', 'speakers', '295', ',', ',', '000']  
 ['Ngas', 'Total', 'speakers', '400', ',', ',', '000']  
 ['Koenoen', 'Total', 'speakers', '3', ',', ',', '000']  
 ['Montol', 'Total', 'speakers', '21', ',', ',', '858']  
 ['Pyapun', 'Total', 'speakers', '17', ',', ',', '284']  
 ['Tal', 'Total', 'speakers', '10', ',', ',', '000']  
 ['Ciwogai', 'Total', 'speakers', '2', ',', ',', '000']  
 ['Diri', 'Total', 'speakers', '7', ',', ',', '196']  
 ['Kariya', 'Total', 'speakers', '2', ',', ',', '000']  
 ['Mburku', 'Total', 'speakers', '12', ',', ',', '000']  
 ['Miya', 'Total', 'speakers', '30', ',', ',', '000']  
 ['Pa'a', 'Total', 'speakers', '8', ',', ',', '000']  
 ['Warji', 'Total', 'speakers', '77', ',', ',', '665']  
 ['Piruris', 'Total', 'speakers', '803', '(', '1980', ')']  
 ['Piruris', 'Region', 'Taos', 'Pueblo', ',', ',', 'New', 'Mexico']  
 ['Duwai', 'Total', 'speakers', '11', ',', ',', '386']  
 ['Kambaata', 'Total', 'speakers', '614', ',', ',', '807', 'of', 'Kambaata',  
 'dialect', '97', ',', ',', '806', 'of', 'Tambaro', 'dialect']  
 ['Southern', 'Tiwa', 'Total', 'speakers', '1600']  
 ['Arbore', 'Total', 'speakers', '7', ',', ',', '211', '(', '2007', 'census',  
 ')']  
 ['Baiso', 'Total', 'speakers', '1', ',', ',', '010']  
 ['Daasanach', 'Region', 'Lower', 'Omo', 'River', ',', ',', 'Lake', 'Turkana']  
 ['El', 'Molo', 'Total', 'speakers', '8', '(', '1994', ')']  
 ['Alagwa', 'Total', 'speakers', '30', ',', ',', '000', '(', '2001', ')']  
 ['Burunge', 'Total', 'speakers', '13', ',', ',', '000', '(', '2002', ')']  
 ['Gorowa', 'Total', 'speakers', '50', ',', ',', '000', '(', '1999', ')']  
 ['Iraqw', 'Total', 'speakers', '462', ',', ',', '000', '(', '2001', ')']  
 ['Dime', 'Total', 'speakers', '574', '(', '2007', 'census', ')']  
 ['Dime', 'Region', 'Debub', '(', 'South', ')', 'Omo', 'Zone']  
 ['Karo', '(', 'Ethiopia', ')', 'Total', 'speakers', '1', ',', ',', '478', '(',  
 '2007', 'census', ')']

['Bana', 'Total', 'speakers', '13', ',', ',', '000', '(', '1987', ')']  
 ['Hya', 'Total', 'speakers', '940', '(', '2002', ')']  
 ['Hya', 'Region', 'Far', 'North', 'Province', 'and', 'neighboring',  
 'Nigeria']  
 ['Kamwe', 'Total', 'speakers', '300', ',', ',', '000', '(', '1992', ')']  
 ['Psikye', 'Total', 'speakers', '52', ',', ',', '500', '(', '1982', ')']  
 ['Psikye', 'Region', 'Far', 'North', 'Province', ',', ',', 'Adamawa', 'State']  
 ['Baldemu', 'Total', 'speakers', '3-6', '(', '2003', ')']  
 ['Cuvok', 'Total', 'speakers', '5', ',', ',', '000', '(', '1983', ')']  
 ['Dugwor', 'Total', 'speakers', '5', ',', ',', '000', '(', '2001', ')']  
 ['Mada', '(', 'Cameroon', ')', 'Total', 'speakers', '17', ',', ',', '000', '(',  
 '1982', ')']  
 ['Mafa', 'Total', 'speakers', '140', ',', ',', '907', '(', '1982', ')']  
 ['Mafa', 'Region', 'Far', 'North', 'Province', ',', ',', 'Borno', 'State']  
 ['Matal', 'Total', 'speakers', '18', ',', ',', '000', '(', '1982', ')']  
 ['Mbuko', 'Total', 'speakers', '13', ',', ',', '000', '(', '2002', ')']  
 ['Mefele', 'Total', 'speakers', '11', ',', ',', '000', '(', '2002', ')']  
 ['Merey', 'Total', 'speakers', '10', ',', ',', '000', '(', '1982', ')']  
 ['Mofu-Gudur', 'Total', 'speakers', '60', ',', ',', '000', '(', '1998', ')']  
 ['Moloko', 'Total', 'speakers', '8', ',', ',', '500', '(', '1992', ')']  
 ['Muyang', 'Total', 'speakers', '15', ',', ',', '000', '(', '1982', ')']  
 ['North', 'Giziga', 'Total', 'speakers', '20', ',', ',', '000', '(', '1982',  
 ')']  
 ['North', 'Mofu', 'Total', 'speakers', '27', ',', ',', '500', '(', '1982', ')']  
 ['South', 'Giziga', 'Total', 'speakers', '60', ',', ',', '000', '(', '1991',  
 ')']  
 ['Vame', 'Total', 'speakers', '8', ',', ',', '500', '(', '1992', ')']  
 ['Wuzlam', 'Total', 'speakers', '10', ',', ',', '500', '(', '1982', ')']  
 ['Bacama', 'Total', 'speakers', '150', ',', ',', '000', '(', '1992', ')']  
 ['Bata', 'Total', 'speakers', '152', ',', ',', '500', '(', '1992', ')']  
 ['Bata', 'Region', 'Adamawa', 'State', ',', ',', 'North', 'Province']  
 ['Fali', 'Total', 'speakers', '20', ',', ',', '000', '(', '1990', ')']  
 ['Gude', 'Total', 'speakers', '96', ',', ',', '000', '(', '1987', ')']  
 ['Gude', 'Region', 'Borno', 'State', ';', 'Far', 'North', 'Province']  
 ['Gudu', 'Total', 'speakers', '5', ',', ',', '000', '(', '1993', ')']  
 ['Jimi', '(', 'Cameroon', ')', 'Total', 'speakers', '3', ',', ',', '500', '(',  
 '1982', ')']  
 ['Ngwaba', 'Total', 'speakers', '10', ',', ',', '000', '(', '1993', ')']  
 ['Nzanyi', 'Total', 'speakers', '86', ',', ',', '000', '(', '1993', ')']  
 ['Sharwa', 'Total', 'speakers', '5', ',', ',', '100', '(', '2000', ')']  
 ['Tsuwan', 'Total', 'speakers', '2', ',', ',', '300', '(', '2000', ')']  
 ['Buduma', 'Total', 'speakers', '54', ',', ',', '800', '(', '1993', ')']  
 ['Jilbe', 'Total', 'speakers', '100', '(', '1999', ')']  
 ['Ngizim', 'Total', 'speakers', '80', ',', ',', '000', '(', '1993', ')']  
 ['Dr\xc3\xa8ents', 'Drents', ',', 'Dreins', 'Total', 'speakers', 'About',  
 'half', 'the', 'population', 'of', 'Drenthe', '(', 'total', '483', '.',

'173', ')', 'and', 'a', 'number', 'of', 'people', 'in', 'Overijssel']  
 ['Dr\xc3\xa8ents', 'Official', 'language', 'in', 'the', 'Netherlands',  
 '(', 'as', 'part', 'of', 'Low', 'Saxon', ')']  
 ['Chara', 'Total', 'speakers', '13', ',', '100', '(', '2007', 'census',  
 ')']  
 ['Buwal', 'Total', 'speakers', '7', ',', '000', '(', '2001', ')']  
 ['Daba', 'Total', 'speakers', '17', ',', '500', '(', '1977', ')']  
 ['Daba', 'Region', 'Far', 'North', 'Province', ';', 'Adamawa', 'State']  
 ['Gavar', 'Total', 'speakers', '7', ',', '000', '(', '2001', ')']  
 ['Mazagway', 'Total', 'speakers', '17', ',', '000', '(', '1997', ')']  
 ['Mazagway', 'Region', 'North', 'Province', ',', 'Far', 'North',  
 'Province']  
 ['Mbedam', 'Total', 'speakers', '6', ',', '000', '(', '2001', ')']  
 ['Mina', 'Total', 'speakers', '10', ',', '987', '(', '2000', ')']  
 ['Bidiyo', 'Total', 'speakers', '14', ',', '000', '(', '1981', ')']  
 ['Dangal\xc3\xa9at', 'Total', 'speakers', '45', ',', '000', '(', '1999',  
 ')']  
 ['Jonkor', 'Bourmataguil', 'Total', 'speakers', '1', ',', '500', '(',  
 '1993', ')']  
 ['Mabire', 'Total', 'speakers', '3', '(', '2001', ')']  
 ['Mawa', 'Total', 'speakers', '6', ',', '560', '(', '2000', ')']  
 ['Migaama', 'Total', 'speakers', '20', ',', '000', '(', '1991', ')']  
 ['Mogum', 'Total', 'speakers', '7', ',', '000', '(', '1997', ')']  
 ['Ubi', 'Total', 'speakers', '1', ',', '100', '(', '1995', ')']  
 ['Official', 'Aramaic', '(', '700-300', 'BCE', ')', 'Language',  
 'extinction', '700-300', 'BCE']  
 ['Official', 'Spoken', 'in', 'Ancient', 'Near', 'East']  
 ['Old', 'Aramaic', '(', 'up', 'to', '700', 'BCE', ')', 'Language',  
 'extinction', 'up', 'to', '700', 'BCE']  
 ['Old', 'Spoken', 'in', 'Ancient', 'Near', 'East']  
 ['Buso', 'Total', 'speakers', '40', '(', '1971', ')']  
 ['Boor', 'Total', 'speakers', '100', '(', '1999', ')']  
 ['Miltu', 'Total', 'speakers', '272', '(', '1993', ')']  
 ['Sarua', 'Total', 'speakers', '2', ',', '000', '(', '1997', ')']  
 ['Birgit', 'Total', 'speakers', '10', ',', '398', '(', '2000', ')']  
 ['Kajakse', 'Total', 'speakers', '10', ',', '000', '(', '1983', ')']  
 ['Masmaje', 'Total', 'speakers', '25', ',', '727', '(', '1993', ')']  
 ['Mubi', 'Total', 'speakers', '35', ',', '277', '(', '1993', ')']  
 ['Toram', 'Total', 'speakers', '8', ',', '456', '(', '2000', ')']  
 ['Zirenkel', 'Total', 'speakers', '2', ',', '237', '(', '1993', ')']  
 ['Barein', 'Total', 'speakers', '14', ',', '000', '(', '1981', ')']  
 ['Saba', 'Total', 'speakers', '1', ',', '335', '(', '2000', ')']  
 ['Sokoro', 'Total', 'speakers', '5', ',', '000', '(', '1994', ')']  
 ['Tamki', 'Total', 'speakers', '500', '(', '1999', ')']  
 ['Karakare', 'Total', 'speakers', '150', ',', '000-200', ',', '000', '(',  
 '1993', ')']

['Beele', 'Total', 'speakers', '120', '(', '1922', ')']  
 ['Bole', 'Total', 'speakers', '100', ',', '000', '(', '1990', ')']  
 ['Deno', 'Total', 'speakers', '6', ',', '000', '(', '1995', ')']  
 ['Galambu', 'Total', 'speakers', '21', ',', '622', 'ethnic', 'population',  
 ',', 'but', 'most', 'do', 'not', 'speak', 'Galambu', '(', '2000', ')']  
 ['Gera', 'Total', 'speakers', '200', ',', '000', '(', '1995', ')']  
 ['Geruma', 'Total', 'speakers', '9', ',', '029', '(', '2000', ')']  
 ['Giiwo', 'Total', 'speakers', '14', ',', '000', '(', '1998', ')']  
 ['Kholok', 'Total', 'speakers', '2', ',', '500', '(', '1977', ')']  
 ['Maaka', 'Total', 'speakers', '10', ',', '000', '(', '1993', ')']  
 ['Ngamo', 'Total', 'speakers', '60', ',', '000', '(', '1993', ')']  
 ['Boghom', 'Total', 'speakers', '50', ',', '000', '(', '1973', ')']  
 ['Kir-Balar', 'Total', 'speakers', '3', ',', '050', '(', '1993', ')']  
 ['Mangas', 'Total', 'speakers', '100', '(', '1995', ')']  
 ['Aweer', 'Total', 'speakers', '3', ',', '500', '(', '1994', ')']  
 ['Aweer', 'Region', 'Coast', 'Province', ',', 'North-Eastern', 'Province']  
 ['Rendille', 'Total', 'speakers', '32', ',', '000', '(', '1994', ')']  
 ['Basketo', 'Total', 'speakers', '92', ',', '586', '(', '2007', ')']  
 ['Male', '(', 'Ethiopia', ')', 'Total', 'speakers', '94', ',', '746', '(',  
 '2007', 'census', ')']  
 ['Boro', 'Total', 'speakers', '37', ',', '459', '(', '2007', 'census', ')']  
 ['Algerian', 'Saharan', 'Arabic', 'Total', 'speakers', '110', ',', '000',  
 '(', '1996', ')']  
 ['Sapo', 'Total', 'speakers', '31', ',', '600', '(', '1991', ')']  
 ['Moose', 'Cree', 'Total', 'speakers', '4', ',', '500']  
 ['Dhofari', 'Total', 'speakers', '70', ',', '000', '(', '1996', ')']  
 ['Yalunka', 'Total', 'speakers', 'over', '100', ',', '000']  
 ['Yalunka', 'Region', 'Southern', 'Guinea', ',', 'Senegal', '\xe2', '\x80',  
 '\x93', 'Mali', 'border']  
 ['Bedawi', 'Total', 'speakers', '1', '610', '000']  
 ['Dewoin', 'Total', 'speakers', '8', ',', '100', '(', '1991', ')']  
 ['Gbii', 'Total', 'speakers', '5', ',', '600', '(', '1991', ')']  
 ['Tajuasohn', 'Total', 'speakers', '9', ',', '600', '(', '1991', ')']  
 ['Glio-Oubi', 'Spoken', 'in', 'Liberia', ',', 'C', '\xc3', '\xb4', 'te',  
 "d'Ivoire"]  
 ['Glio-Oubi', 'Total', 'speakers', '6', ',', '000', '(', '1991', ')']  
 ['Kuwaa', 'Total', 'speakers', '12', ',', '800', '(', '1991', ')']  
 ['Omani', 'Arabic', 'Total', 'speakers', '815', ',', '000']  
 ['Sa`idi', 'Total', 'speakers', '18', ',', '900', ',', '000']  
 ['Sanaani', 'Arabic', 'Total', 'speakers', '7', ',', '600', ',', '000',  
 '(', '1996', ')']  
 ["Ta'izzi-Adeni", 'Arabic', 'Total', 'speakers', '6', ',', '869', ',', '000']  
 ['Shihhi', 'Total', 'speakers', '27', ',', '000']  
 ['Judeo-Tripolitanian', 'Total', 'speakers', '35', ',', '000']  
 ['Judeo-Tunisian', 'Total', 'speakers', '45', ',', '500']

['Hoby\xc3\xb3t', 'Total', 'speakers', 'about', '100', '(', '1998', ')']  
 ['Shehri', 'Total', 'speakers', '25', ',', ',', '000']  
 ['Cineni', 'Total', 'speakers', '3', ',', ',', '000', '(', '1998', ')']  
 ['Dghwede', 'Total', 'speakers', '30', ',', ',', '000', '(', '1980', ')']  
 ['Glavda', 'Total', 'speakers', '31', ',', ',', '265']  
 ['Glavda', 'Region', 'Borno', 'State', ';;', 'Far', 'North', 'Province']  
 ['Guduf-Gava', 'Total', 'speakers', '55', ',', ',', '918', '(', '2000', ')']  
 ['Gvoko', 'Total', 'speakers', '21', ',', ',', '000']  
 ['Gvoko', 'Region', 'Borno', 'State', ';;', 'Far', 'North', 'Province']  
 ['Duhwa', 'Total', 'speakers', '800', '(', '1973', ')']  
 ['Kulere', 'Total', 'speakers', '15', ',', ',', '570', '(', '1990', ')']  
 ['Mundat', 'Total', 'speakers', '1', ',', ',', '000', '(', '1998', ')']  
 ['Ron', 'Total', 'speakers', '115', ',', ',', '000', '(', '1995', ')']  
 ['Sha', 'Total', 'speakers', '3', ',', ',', '000', '(', '1998', ')']  
 ['Carolina', 'Language', 'codes', 'ISO', '639-3', ':', 'crr', '\xe2',  
 '\x80', '\x93', 'Carolina', 'Algonquian']  
 ['Naskapi', 'Total', 'speakers', '1', ',', ',', '177']  
 ['Northern', 'East', 'Cree', 'Total', 'speakers', '5', ',', ',', '308', '(',  
 '1997', ')']  
 ['Southern', 'East', 'Cree', 'Total', 'speakers', '7', ',', ',', '306']  
 ['Woods', 'Cree', 'Total', 'speakers', '35', ',', ',', '000']  
 ['Chippewa', 'Region', 'upper', 'Michigan', 'westward', 'to', 'North',  
 'Dakota']  
 ['Malecite-Passamaquoddy', 'Total', 'speakers', '1', ',', ',', '655']  
 ['Malecite-Passamaquoddy', 'Region', 'New', 'Brunswick', ';;', 'Maine']  
 ['Gros', 'Ventre', 'Total', 'speakers', 'A', 'few', 'semi-speakers', 'in',  
 '2000']  
 ['Dass', 'Total', 'speakers', '8', ',', ',', '830']  
 ['Dass', 'Region', 'Bauchi', 'State', ',', ',', 'Plateau', 'State']  
 ['Jimi', 'Total', 'speakers', '1', ',', ',', '000']  
 ['Gawwada', 'Total', 'speakers', '32', ',', ',', '698', 'including', '27', ',', ',',  
 '477', 'monolinguals', '(', '1998', ')']  
 ['Tsamai', 'Total', 'speakers', '18', ',', ',', '048', '(', '2007', 'census',  
 ')']  
 ['Dabarre', 'Total', 'speakers', '26', ',', ',', '753', '(', '2000', ')']  
 ['Garre', 'Total', 'speakers', '50', ',', ',', '000', '(', '1992', ')']  
 ['Jiiddu', 'Total', 'speakers', '29', ',', ',', '726', '(', '2000', ')']  
 ['Jiiddu', 'Region', 'Lower', 'Shabeelle', 'Bay', 'and', 'Middle', 'Jubba',  
 'regions']  
 ['Tunni', 'Total', 'speakers', '29', ',', ',', '726', '(', '2000', ')']  
 ['Tunni', 'Region', 'Lower', 'Shabeelle', 'and', 'Middle', 'Jubba',  
 'regions']  
 ['Mukulu', 'Total', 'speakers', '12', ',', ',', '000', '(', '1991', ')']  
 ['Dorze', 'Total', 'speakers', '20', ',', ',', '782', ',', ',', 'including', '9',  
 ',', ',', '905', 'monolinguals', '(', '1998', ')']  
 ['Gamo-Gofa-Dawro', 'Total', 'speakers', '1', ',', ',', '236', ',', ',', '637', ',', ',',

'including', '1', ',', '046', ',', '084', 'monolinguals', '(', '1998', ')']  
 ['Melo', 'Total', 'speakers', '20', ',', '151', ',', 'including', '13',  
 ',', '264', 'monolinguals', '(', '1998', ')']  
 ['Oyda', 'Total', 'speakers', '36', ',', '946', '(', '2007', 'census', ')']  
 ['Middle', 'Mongolian', 'Language', 'extinction', 'developed', 'into',  
 'Classical', 'Mongolian', 'by', 'the', '17th', 'century']  
 ['Maipure', 'Language', 'extinction', 'late', '18th', 'century']  
 ['Northern', 'Spoken', 'in', 'Uzbekistan', ',', 'People's', 'Republic',  
 'of', 'China']  
 ['Northern', 'Total', 'speakers', '18', ',', '795', ',', '591']  
 ['Southern', 'Total', 'speakers', '1', ',', '454', ',', '981']  
 ['Northern', 'Altai', 'Total', 'speakers', '&', 'lt', ';', '5', ',', '000']  
 ['Southern', 'Altai', 'Total', 'speakers', '20', ',', '000']  
 ['North', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['South', 'Azeri', 'Total', 'speakers', '12', '-', '18', 'million',  
 'worldwide']  
 ['South', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Salchuq', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']  
 ['Amto', 'Total', 'speakers', '200', '(', '2000', ')']  
 ['Amto', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Siawi', 'Total', 'speakers', '70', '(', '2000', ')']  
 ['Siawi', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Fyer', 'Total', 'speakers', '26', ',', '131', '(', '2000', ')']  
 ['Tambas', 'Total', 'speakers', '3', ',', '000', '(', '2001', ')']  
 ['Mundari', 'Total', 'speakers', '2', ',', '074', ',', '700']  
 ['Rodriguan', 'Total', 'speakers', '40', ',', '000']  
 ['Chagossian', 'Total', 'speakers', '3', ',', '000']  
 ['Agalega', 'Total', 'speakers', '&', 'lt', ';', '1', ',', '000']  
 ['Ma'di', 'Total', 'speakers', '325', ',', '000']  
 ['Batak', 'Karo', 'Total', 'speakers', '600', ',', '000']  
 ['Sunam', 'Total', 'speakers', '560', '(', '1998', ')']  
 ['Gahri', 'Total', 'speakers', '\xe2', '\x89', '\x88', '4000']  
 ['Paama', 'Total', 'speakers', '7', ',', '000']  
 ['Zia', 'Total', 'speakers', '3', ',', '943', '(', '1978', 'McElhanon',  
 ')']  
 ['Zia', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Sheko', 'Total', 'speakers', '38', ',', '911', '(', '2007', 'census',  
 ')']  
 ['Sheko', 'Region', 'Bench', 'Maji', 'Zone', ',', 'Kafa', 'region']  
 ['Camling', 'Spoken', 'in', 'Nepal', '.', 'India', ',', 'Bhutan']  
 ['Camling', 'Total', 'speakers', '10', ',', '000']  
 ['Puinave', 'Total', 'speakers', '3', ',', '000', '(', '1977', ')']  
 ['Jarawa', 'Total', 'speakers', '266', '(', '2001-2002', ')']  
 ['Jemez', 'Language', 'codes', 'ISO', '639-3', 'Towa', 'language',  
 'distribution', 'in', 'the', 'State', 'of', 'New', 'Mexico']  
 ['Awa\xe2\x80\x93Cuaiquer', 'Spoken', 'in', 'Northwestern', 'South',

'America']  
 ['Awa\xe2\x80\x93Cuaiquer', 'Total', 'speakers', 'Colombia', '20', ',', '000', '(', '1986', ')']  
 ['Awa\xe2\x80\x93Cuaiquer', 'Total', 'speakers', 'Ecuador', '1', ',', '000', '(', '1991', ')']  
 ['Cat\xe3\xado', 'Total', 'speakers', '15', ',', '040']  
 ['Cham\xe3\xad', 'Total', 'speakers', '40', ',', '000']  
 ['Kulung', 'Total', 'speakers', '15', ',', '000', '-', '18', ',', '868']  
 ['Ndali', 'Total', 'speakers', '220', ',', '000', '(', '1987', ')']  
 ['Old', 'Pronunciation', 'kaste', '\xce', '\xbb', 'ano', 'aN'ti', '\xce', '\xb3', 'wo']  
 ['Dizin', 'Total', 'speakers', '33', ',', '927', '(', '2007', 'census', ')']  
 ['Andoque', 'Total', 'speakers', '610', '(', '2000', ')']  
 ['Barasana\xe2\x80\x93Eduria', 'Total', 'speakers', '1890', '(', '2001', ')']  
 ['Abui', 'Total', 'speakers', '32', ',', '000']  
 ['Cabiyari', 'Total', 'speakers', '50', '(', 'dated', ')', '(', '1976', ')']  
 ['Kakwa', 'Total', 'speakers', '150', '(', 'dated', ')', '(', '1982', ')']  
 ['Khinalug', 'Total', 'speakers', '1', ',', '500']  
 ['Gumuz', 'Writing', 'system', 'Ethiopic', '/', 'Latin', '(', 'in', 'Ethiopia', ')']  
 ['Ngiemboon', 'Total', 'speakers', '100', '.', '000', 'SIL', ',', '1987']  
 ['Ngiemboon', 'Region', 'Province', 'de', 'l'Ouest', ',', 'Bamboutos']  
 ['Burgundian', 'Language', 'extinction', '5th', 'century']  
 ['Sio', 'Total', 'speakers', '3', ',', '500', '(', '1987', ')']  
 ['Sio', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['K\xe3\xa2te', 'Total', 'speakers', '6', ',', '125', '(', '1978', 'McElhanon', ')']  
 ['K\xe3\xa2te', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Pare', 'Total', 'speakers', '500', ',', '000']  
 ['Davani', 'Total', 'speakers', 'ca', '.', '1000']  
 ['Busa', 'Total', 'speakers', '40', ',', '000']  
 ['Singapore', 'Hokkien', 'Total', 'speakers', '1', ',', '824', ',', '741']  
 ['Kriva', 'Palanka', 'dialect', 'Total', 'speakers', 'approx', '.', '21', ',', '000']  
 ['Kriva', 'Language', 'codes', 'ISO', '639-3', 'The', 'location', 'of', 'the', 'Kriva', 'Palanka', 'dialect', 'dialect', 'among', 'the', 'others', 'Macedonian', 'dialects']  
 ['Mirpur', 'Language', 'codes', 'ISO', '639-3', 'Mirpuri', 'Punjabi', 'is', 'part', 'of', 'the', 'area', 'listed', 'as', 'Lahnda']  
 ['Omi', 'Total', 'speakers', '39', ',', '500', '(', '1989', 'SIL', ')']  
 ['Baima', 'Total', 'speakers', '11', ',', '000', 'in', 'China', '(', '1999', 'Sun', 'Hongkai', ')']  
 ['Bujhyal', 'Total', 'speakers', '5', ',', '000', 'in', 'China', '(',



'1998', ')]  
 ['Chepang', 'Total', 'speakers', '36', ',', '807', '(', '2001', 'census', ')]  
 ['Wayu', 'Total', 'speakers', '1', ',', '743', '(', '2001', 'census', ')]  
 ['Deilami', 'Language', 'extinction', 'gradually', '900-1300AD']  
 ['Deilami', 'Region', 'South', 'Caspian', 'Sea', ',', 'Alborz', 'highlands']  
 ['Mangbutu', 'Total', 'speakers', '15', ',', '000']  
 ['Mangbetu', 'Total', 'speakers', '620', ',', '000']  
 ['Efe', 'Total', 'speakers', '20', ',', '000', '(', '1991', ')]  
 ['Lese', 'Total', 'speakers', '50', ',', '000', '(', '1991', ')]  
 ['Malawi', 'Total', 'speakers', '250', ',', '000']  
 ['Kumarbhag', 'Spoken', 'in', 'Jharkhand', ',', 'West', 'Bengal']  
 ['Kumarbhag', 'Total', 'speakers', '16', ',', '000', '(', '2000', 'estimate', ')]  
 ['Jakati', 'Spoken', 'in', 'Afghanistan', '(', '1', ',', '365', 'speakers', ')]  
 ['Jakati', 'Spoken', 'in', 'Ukraine', '(', '29', ',', '250', 'speakers', ')]  
 ['Jakati', 'Total', 'speakers', '30', ',', '615']  
 ['Jordanian', 'Total', 'speakers', 'First', 'language', ':', '6', '.', '15', 'million']  
 ['Barbare\xc3\xb1o', 'Language', 'extinction', '1965', ',', 'with', 'the', 'death', 'of', 'Mary', 'Yee']  
 ['Barbare\xc3\xb1o', 'Spoken', 'in', 'Southern', 'Californian', 'coastal', 'areas']  
 ['Transpiranto', 'Setting', 'and', 'usage', 'Swedish', 'anthology']  
 ['Shompen', 'Total', 'speakers', '\xe2', '\x89', '\x88', '400']  
 ['Fasu', 'Total', 'speakers', '1', ',', '200', '(', 'Wurm', 'and', 'Hattori', '1981', ')', ':', '750', 'Fasu', ',', '300', 'Namuni', ',', '150', 'Some']  
 ['Tokelauan', 'Language', 'extinction', '19th', 'century']  
 ['Alo-Teqel', 'Language', 'extinction', 'ca', '.', '1900', '?']  
 ['Rerep', 'Spoken', 'in', 'East', 'Malekula', ',', 'Vanuatu']  
 ['Rerep', 'Total', 'speakers', '~', '400', '(', '1983', ')]  
 ['Chan\xc3\xa1', 'Language', 'extinction', 'early', '20th', 'Century']  
 ['Baenan', 'Language', 'extinction', 'ca', '.', '1940']  
 ['Mire', 'Total', 'speakers', '1400', '(', '1990', ')]  
 ['Culle', 'Language', 'extinction', 'mid-20th', 'Century']  
 ['Ndam', 'Total', 'speakers', '6500', '(', '1990', ')]  
 ['Cuitlatec', 'Language', 'extinction', '1960s', ',', 'with', 'the', 'death', 'of', 'Juana', 'Can']  
 ['Somrai', 'Total', 'speakers', '7414', '(', '1993', ')]  
 ['Tumak', 'Total', 'speakers', '25', ',', '249', '(', '1993', ')]  
 ['Abduyi', 'Total', 'speakers', 'ca', '.', '600']  
 ['Abduyi', 'Regulated', 'by', 'Academy', 'of', 'Persian', 'Language',

'and', 'Literature']  
 ['Oriya', 'Total', 'speakers', '33', 'million', '(', '2001', ')']  
 ['Oriya', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains',  
 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may',  
 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of',  
 'conjuncts', '.', 'More']  
 ['Andi', 'Total', 'speakers', '21', ',', '800', '(', '2002', ')']  
 ['Akhvakh', 'Spoken', 'in', 'Southern', 'Dagestan', ',', 'Azerbaijan']  
 ['Akhvakh', 'Total', 'speakers', '6500', '(', '2006', ')']  
 ['Jaminjung', 'Total', 'speakers', '50-150', '(', '2000', ')']  
 ['Jaminjung', 'Total', 'speakers', '30', '(', '1991', ')']  
 ['Nungali', 'Region', 'Upper', 'Daly', 'River', ',', 'Northern',  
 'Territory']  
 ['Movima', 'Total', 'speakers', 'ca', '.', '1400']  
 ['Torgut', 'Total', 'speakers', '106', ',', '000', '(', 'China', ')', '+',  
 '11', ',', '400', '(', 'Mongolia', ')']  
 ['Torgut', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Torgut', 'Writing', 'system', 'Clear', 'script', '(', 'China', ':',  
 'informal', ')', ',', 'Mongolian', 'script', '(', 'China', ':',  
 'official', ')', ',', 'Cyrillic', '(', 'Kalmykia', ')']  
 ['Kol', 'Total', 'speakers', '12', ',', '000', '(', 'including', '1', ',',  
 '000', 'in', 'Bikeng', ')']  
 ['Aja/Adja', 'Total', 'speakers', '740', ',', '000']  
 ['Wuxi', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Wuxi', 'Region', 'Wuxi', ',', 'Jiangsu', 'province']  
 ['Temein', 'Total', 'speakers', '13', ',', '000', '(', '2006', ')']  
 ['Illyrian', 'Language', 'extinction', '2nd-6th', 'centuries']  
 ['Bahnar', 'Total', 'speakers', '160', ',', '000']  
 ['Old', 'Kazeruni', 'Language', 'extinction', '?']  
 ['Old', 'Regulated', 'by', 'Academy', 'of', 'Persian', 'Language', 'and',  
 'Literature']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', '0xw', '\xe2', '\x80', '\x93']  
 ['Urarina', 'Total', 'speakers', 'around', '3000']  
 ['Guat\xc3\xb3', 'Total', 'speakers', '49', '(', 'in', '1993', ')']  
 ['Sechura', 'Language', 'extinction', 'late', '19th', 'Century', '?']  
 ['Sechura', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'Sechura', 'and', 'Tall', '\xc3', '\xa1', 'n', 'in', 'the', 'Piura',  
 'region']  
 ['Laghu', 'Spoken', 'in', 'Santa', 'Isabel', 'Island']  
 ['Gothic', 'Language', 'extinction', 'mostly', 'extinct', 'by', 'the',  
 '8th', 'or', '9th', 'century', ',', 'remnants', 'may', 'have', 'lingered',  
 'into', 'the', '17th', 'century']  
 ['Oran', 'Total', 'speakers', '1', ',', '737']  
 ['Hakha', 'Total', 'speakers', '446', ',', '264']  
 ['Kassite', 'Language', 'extinction', 'by', 'the', 'end', 'of', 'the',  
 '4th', 'century', 'BCE']

['Pasto', 'Spoken', 'in', 'Northwestern', 'South', 'America']  
 ['Wagdi', 'Total', 'speakers', '1', '.', '6', 'million', '(', '1997', ')']  
 ['Caura', 'Total', 'speakers', '2', ',', '000']  
 ['Universal', 'Created', 'by', 'G', '.', 'I', '.', 'Muravskin', '&', 'amp',  
 ';', 'L', '.', 'I', '.', 'Vasilevskij']  
 ['Zhaba', 'Total', 'speakers', '8', ',', '000', '(', '1995', ')']  
 ['Judaéo-Spanish', 'also', 'Judezmo', ',', 'Ladino', ',', 'and', 'other',  
 'names', 'Total', 'speakers', '100', ',', '000', 'in', 'Israel', '(',  
 '1985', ')']  
 ['Judaéo-Spanish', 'Regulated', 'by', 'Autoridad', 'Nasionala', 'del',  
 'Ladino', 'in', 'Israel', '(', 'based', 'on', 'modern', 'Castilian', 'and',  
 'using', 'Latin', 'letters', ')']  
 ['Judaéo-Spanish', 'Language', 'codes', 'Linguasphere', '51-AAB-ba', '.',  
 '.', '51-AAB-bd']  
 ['Cayuse', 'Language', 'extinction', '19th', 'century']  
 ['Lojban', 'Category', '(', 'purpose', ')', '(', 'descended', 'from',  
 'Loglan', ')']  
 ['Lojban', 'Writing', 'system', 'Latin', 'alphabet', 'and', 'others',  
 'constructed', 'languages']  
 ['Lojban', 'Created', 'by', 'Logical', 'Language', 'Group']  
 ['Albanian', 'Total', 'speakers', '7', '.', '6', 'million']  
 ['Albanian', 'Spoken', 'in', 'Primarily', 'in', 'Southeastern', 'Europe',  
 'and', 'by', 'the', 'Albanian', 'diaspora', 'worldwide']  
 ['Albanian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Albanian',  
 'variant', ')']  
 ['Albanian', 'Language', 'codes', 'Linguasphere', '55-AAA-aaa', 'to', '55-  
 AAA-ahe', '(', '25', 'varieties', ')']  
 ['Gciriku', 'Total', 'speakers', '36', ',', '000']  
 ['Hokkien', 'Total', 'speakers', 'c', '.', '47', '.', '3', 'million']  
 ['Hokkien', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Hokkien', 'Official', 'language', 'in', 'None', '(', 'Legislative',  
 'bills', 'have', 'been', 'proposed', 'for', 'Taiwanese', 'to', 'be', 'one',  
 'of', 'the', '"national", "languages"', 'in', 'the', 'Republic', 'of',  
 'China', ')', ';', 'one', 'of', 'the', 'statutory', 'languages', 'for',  
 'public', 'transport', 'announcements', 'in', 'the', 'ROC']  
 ['Hokkien', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'Minnan', 'dialects', '.', 'Hokkien', 'is', 'dark', 'green']  
 ['Puguli', 'Writing', 'system', 'African', 'reference', 'alphabet', '(',  
 'Latin', 'alphabet', ')']  
 ['Lote', 'Total', 'speakers', '5', ',', '500', '(', '2004', 'SIL', ')']  
 ['Lote', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Lote', 'Region', '5', '\xc2', '\xb0', '57', '\xe2', '\x80', '\xb2', 'S',  
 '150', '\xc2', '\xb0', '54', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb',  
 '\xbf', '/', '\xef', '\xbb', '\xbf', '5', '.', '95', '\xc2', '\xb0', 'S',  
 '150', '.', '9', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-5',  
 '.', '95', ';', '150', '.', '9']

['Lote', 'Region', 'East', 'New', 'Britain', 'Province']  
 ['Lote', 'Region', 'southeast', 'coast', 'and', 'inland', 'near', 'Cape',  
 'Dampier', '.', 'Coordinates']  
 ['Ugaritic', 'Language', 'extinction', '12th', 'century', 'BC']  
 ['Aheu', 'Total', 'speakers', 'c', '.', '2', ',', '520', '(', '2000', ')']  
 ['Akha', 'Total', 'speakers', '449', ',', '261']  
 ['Nayi', 'Total', 'speakers', '7', ',', '188', '(', '2007', 'census', ')']  
 ['Alak', 'Total', 'speakers', '4', ',', '000', '(', '2000', ')']  
 ['Arem', 'Total', 'speakers', '40', '(', '1996', ')']  
 ['Bit', 'Total', 'speakers', '2', ',', '030', '(', '1990', ')']  
 ['Pinyin', 'Total', 'speakers', '24', ',', '600', '(', '2001', 'SIL', ')']  
 ['Koroshi', 'Total', 'speakers', 'ca', '.', '1000']  
 ['Koroshi', 'Regulated', 'by', 'Academy', 'of', 'Persian', 'Language',  
 'and', 'Literature']  
 ['Eastern', 'Total', 'speakers', '129', ',', '559', '(', '1999', ')']  
 ['Chut', 'Total', 'speakers', '4', ',', '279']  
 ['Con', 'Total', 'speakers', '1', ',', '000', '(', '1981', ')']  
 ['Halang', 'Total', 'speakers', '4', ',', '346', '(', '2000', ')']  
 ['Korouni', 'Regulated', 'by', 'Academy', 'of', 'Persian', 'Language',  
 'and', 'Literature']  
 ['Tikopia', 'Total', 'speakers', '3', ',', '324', '(', '1999', 'SIL', ')']  
 ['Anuta', 'Total', 'speakers', '267', '(', '1999', 'SIL', ')']  
 ['Alamblak', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Alamblak', 'Total', 'speakers', '1', ',', '527', '(', '2000', 'census',  
 ')']  
 ['Hung', 'Total', 'speakers', '2', ',', '700-3', ',', '700', '(', '1996',  
 ')']  
 ['Ir', 'Total', 'speakers', '4', ',', '420', '(', '2000', ')']  
 ['Jeh', 'Total', 'speakers', '23', ',', '256', '(', '2002', ')']  
 ['Jeng', 'Total', 'speakers', '7', ',', '320', '(', '2000', ')']  
 ['Savosavo', 'Total', 'speakers', '2', ',', '415', '(', '1999', 'SIL', ')']  
 ['Touo', 'Total', 'speakers', '1', ',', '874', '(', '1999', 'census', ')']  
 ['Touo', 'Region', 'South', 'Rendova', 'Island', ',', 'Western',  
 'Province']  
 ['Tetepare', 'Region', 'Tetepare', 'Island', ',', 'Western', 'Province']  
 ['Bilua', 'Total', 'speakers', '8', ',', '740', '(', '1999', 'SIL', ')']  
 ['Bilua', 'Region', 'Vella', 'Lavella', 'Island', ',', 'Western',  
 'Province']  
 ['Santa', 'Cruz', 'Total', 'speakers', '5', ',', '899', '(', '1999', 'SIL',  
 ')', '.', 'Population', 'includes', '4', ',', '276', 'Londai', 'and', '1',  
 ',', '623', 'Nea']  
 ['Santa', 'Region', 'Santa', 'Cruz', 'Islands', ',', 'Eastern', 'Solomons',  
 '.', 'Coordinates']  
 ['Santa', 'Region', '10', '\xc2', '\xb0', '40', '\xe2', '\x80', '\xb2',  
 'S', '165', '\xc2', '\xb0', '50', '\xe2', '\x80', '\xb2', 'E', '\xef',  
 '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '10', '.', '667', '\xc2',

'\xb0', 'S', '165', '.', '833', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
'\xbf', '/', '-10', '.', '667', ';', '165', '.', '833']  
['Nanggu', 'Total', 'speakers', '210', '(', '1999', 'SIL', ')']  
['Nanggu', 'Region', '10', '\xc2', '\xb0', '44', '\xe2', '\x80', '\xb2',  
'S', '165', '\xc2', '\xb0', '53', '\xe2', '\x80', '\xb2', 'E', '\xef',  
'\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '10', '.', '733', '\xc2',  
'\xb0', 'S', '165', '.', '883', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
'\xbf', '/', '-10', '.', '733', ';', '165', '.', '883']  
['Moxos', 'Total', 'speakers', '4', ',', '500', 'and', '5', ',', '500',  
'out', 'of', 'an', 'ethnic', 'population', 'of', '21', ',', '000']  
['Yur\xc3\xad', 'Region', '1', '\xc2', '\xb0', '50', '\xe2', '\x80',  
'\xb2', 'S', '69', '\xc2', '\xb0', '0', '\xe2', '\x80', '\xb2', 'W',  
'\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '1', '.', '833',  
'\xc2', '\xb0', 'S', '69', '\xc2', '\xb0', 'W', '\xef', '\xbb', '\xbf',  
 '/', '-1', '.', '833', ';', '-69']  
['Tuwali', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')',  
';']  
['Parya', 'Total', 'speakers', '2', ',', '500']  
['Larteh', 'Total', 'speakers', '74', '000']  
['Busuu', 'Total', 'speakers', '8', '(', '1986', 'Breton', ')']  
['Siraya', 'Region', '22', '\xc2', '\xb0', '58', '\xe2', '\x80', '\xb2',  
'N', '120', '\xc2', '\xb0', '18', '\xe2', '\x80', '\xb2', 'E', '\xef',  
'\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '22', '.', '967', '\xc2',  
'\xb0', 'N', '120', '.', '3', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf',  
 '/', '22', '.', '967', ';', '120', '.', '3']  
['Orang', 'Kanaq', 'Total', 'speakers', '83']  
['Lun', 'Bawang', 'Total', 'speakers', 'first', 'language', ':', '300',  
'in', 'Brunei', '(', '1987', 'Langub', ')', ',', '12', ',', '800',  
'including', '10', ',', '000', 'in', 'Lun', 'Bawang', 'dialect', 'in',  
'Sarawak', '(', '1987', ')', ',', '2', ',', '800', 'in', 'Lun', 'Daye',  
'in', 'Sabah', '(', '1982', 'SIL', ')']  
['Swedish', 'Sign', 'Language', 'Signed', 'in', 'Sweden']  
['Swedish', 'Language', 'codes', 'ISO', '639-3', 'swl', 'sign', 'language',  
'\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
'\x94', 'legal', 'recognition']  
['Pajapan', 'Total', 'speakers', '7', ',', '000', '(', 'as', 'of', '1990',  
')']  
['Kemezung', 'Total', 'speakers', '4', ',', '500']  
['Sena', 'Total', 'speakers', '1', ',', '340', ',', '000', '(', '2006',  
')']  
['Kokota', 'Spoken', 'in', 'Santa', 'Isabel', 'Island']  
['Marchigiano', 'Total', 'speakers', '~', '900', '.', '000']  
['Marchigiano', 'Language', 'codes', 'Linguasphere', '51-AAA-okl', '&',  
'amp', ';', '51-AAA-rba']  
['Cicipu', 'Total', 'speakers', '20', ',', '000', '(', 'CAPRO', '1995',  
')']

['Irish', 'Total', 'speakers', '1', '.', '66', 'million', 'with', 'some', 'knowledge']

['Irish', 'Total', 'speakers', '(', '2006', 'Republic', 'only', ')']

['Irish', 'Spoken', 'in', 'Ireland', '(', 'Republic', 'of', ')', '(', '1', '.', '66', 'million', ')']

['Irish', 'Spoken', 'in', 'United', 'Kingdom', '(', '95', ',', '000', ')']

['Irish', 'Spoken', 'in', 'USA', '(', '18', ',', '000', ')']

['Irish', 'Spoken', 'in', 'EU', '(', 'Official', 'EU', 'language', ')']

['Vedda', 'Spoken', 'in', 'Sri', 'Lanka', 'Uva', 'Province']

['Kelantanese', 'Malay', 'Total', 'speakers', '1', '.', '5', 'million']

['Baekje', 'Language', 'extinction', '7th', 'century', '?']

['Baekje', 'Language', 'codes', 'ISO', '639-3', 'The', 'Three', 'Kingdoms', 'of', 'Korea', ',', 'with', 'Baekje', 'in', 'green']

['Finnish', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Finnish', 'variant', ')']

['Finnish', 'Regulated', 'by', 'Language', 'Planning', 'Department', 'of', 'the', 'Research', 'Institute', 'for', 'the', 'Languages', 'of', 'Finland']

['Dutch', 'Region', 'mainly', 'Western', 'Europe', ',', 'today', 'also', 'in', 'South', 'America', 'and', 'the', 'Caribbean']

['Dutch', 'Region', 'Afrikaans', 'is', 'spoken', 'in', 'Southern', 'Africa']

['Dutch', 'Spoken', 'in', 'and', 'South', 'Africa']

['Dutch', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Dutch', 'variant', ')']

['Dutch', 'Official', 'language', 'in', 'Union', 'of', 'South', 'American', 'Nations']

['Dutch', 'Regulated', 'by', '(', 'Dutch', 'Language', 'Union', ')']

['Dutch', 'Language', 'codes', 'Linguasphere', '52-ACB-a', '(', 'varieties']

['Dutch', 'Language', 'codes', 'Linguasphere', ':', '52-ACB-aa', 'to', 'an', ')', 'Dutch-speaking', 'world', '.', 'Dutch', 'is', 'also', 'one', 'of', 'the', 'official', 'languages', 'of', 'the', 'European', 'Union', 'and', 'the', 'Union', 'of', 'South', 'American', 'Nations']

['Salish', 'Total', 'speakers', '&', 'lt', ';', '124']

['Icelandic', 'Total', 'speakers', '+', '320', '000']

['Icelandic', 'Spoken', 'in', 'USA', 'and', 'Canada', 'by', '2']

['Icelandic', 'Spoken', 'in', '170', 'people', 'as', 'of', '2006', 'Canadian', 'census']

['Icelandic', 'Regulated', 'by', '\xc3', '\x81', 'rni', 'Magn', '\xc3', '\xba', 'sson', 'Institute', 'for', 'Icelandic', 'Studies', 'in', 'an', 'advisory', 'capacity']

['Sobei', 'Total', 'speakers', '2', ',', '000', '(', 'Stern', '1987', ')']

['Bikol', 'Total', 'speakers', '2', '.', '5', 'million', ';', '7th', 'most', 'spoken', 'native', 'language', 'in', 'the', 'Philippines']

['Bikol', 'Writing', 'system', 'Latin', '(', 'Filipino', 'variant', ')',

';' ]  
 ['Bikol', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino',  
 'Language']  
 ['Bikol', 'Language', 'codes', 'ISO', '639-3', 'bcl', '\xe2', '\x80',  
 '\x93', 'Central', 'Bicolano']  
 ['Jalapa', 'Mazatec', 'Total', 'speakers', '15', ',', '000']  
 ['Jalapa', 'Region', 'San', 'Felipe', 'Jalapa', 'de', 'D', '\xc3', '\xad',  
 'az']  
 ['Axininca', 'Total', 'speakers', '4', ',', '000', '(', '2000', ')']  
 ['Ndrumbea', 'Total', 'speakers', '946', '(', '1996', 'census', ')']  
 ['Ndrumbea', 'Region', 'Southern', 'tip', 'outside', 'Noum', '\xc3',  
 '\xa9', 'a', '(', 'Païta', 'on', 'the', 'west', 'coast', ',', 'Ounïa',  
 'on', 'the', 'east', 'coast', ')']  
 ['Avatime', 'Total', 'speakers', '24', '000', '(', '2003', ')']  
 ['Yan-nha\xc5\x8bu', 'Total', 'speakers', '15', '(', '2008', ')']  
 ['Kairiru', 'Spoken', 'in', 'Northern', 'New', 'Guinea']  
 ['Kairiru', 'Total', 'speakers', '3', ',', '507', '(', 'Wivell', '1981',  
 ')']  
 ['Jiamao', 'Total', 'speakers', '52', '300', '(', 'as', 'of', '1987', ')']  
 ['Jiamao', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Sabino', 'Region', 'Lazio', '(', 'provinces', 'of', 'Rieti', ',', 'Rome',  
 ')', ',', 'Northern', 'Abruzzo', '(', 'province', 'of', 'Aquila', ')']  
 ['Harappan', 'Language', 'extinction', '1900', 'BCE', '?']  
 ['Romano-Greek', 'Total', 'speakers', '30', '(', 'in', '2000', ')']  
 ['Djeoromitxi', 'Total', 'speakers', '5', '(', '1990', 'SIL', ')']  
 ['Karir\xc3\xad', 'Language', 'extinction', 'ca', '.', '1970']  
 ['Achuar', 'Total', 'speakers', '5', ',', '000']  
 ['Idoma', 'Spoken', 'in', 'Benue', 'State', ',', 'Central', 'Nigeria']  
 ['Idoma', 'Total', 'speakers', '600', ',', '000', '(', '1991', ')']  
 ['Ngbandi', 'Spoken', 'in', 'Central', 'and', 'eastern', 'Central',  
 'African', 'Republic', ',', 'parts', 'of', 'Democratic', 'Republic', 'of',  
 'the', 'Congo']  
 ['Ngbandi', 'Language', 'codes', 'ISO', '639-3', ':', 'ngb', '\xe2',  
 '\x80', '\x93', 'Northern', 'Ngbandi']  
 ['Ngbandi', 'Language', 'codes', 'ISO', '639-3', 'nbw', '\xe2', '\x80',  
 '\x93', 'Southern', 'Ngbandi']  
 ['Nimadi', 'Total', 'speakers', '1', '.', '3', 'million', '(', '1997', ')']  
 ['Lambadi', 'Total', 'speakers', '2', '.', '8', 'million', '(', '1997',  
 ')']  
 ['Bajjika', 'Total', 'speakers', '11', ',', '500', ',', '000']  
 ['Bajjika', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Bijago', 'Total', 'speakers', '30', ',', '000']  
 ['Pyu', 'Language', 'extinction', '12th', 'century']

['Maastrichtian', 'Total', 'speakers', '60', '.', '000', '(', 'est', '.', ')']  
 ['Cori', 'Total', 'speakers', '1000', '(', 'in', '2004', ')']  
 ['Khunsari', 'Total', 'speakers', '21', ',', '000']  
 ['Dusun', 'Total', 'speakers', 'first', 'language', ':', '140', ',', '500', '(', '1986', 'SIL', '.', ')']  
 ['Kaingang', 'Total', 'speakers', '18', ',', '000']  
 ['Bambassi', 'Total', 'speakers', '5', ',', '000', '(', '1982', ')']  
 ['Izi', 'Total', 'speakers', '200', ',', '000']  
 ['Khengkha', 'Total', 'speakers', '40', ',', '000', '(', 'as', 'of', '1993', ')']  
 ['Anuak', 'Total', 'speakers', '90', ',', '579', 'in', 'Ethiopia', '(', '2007', 'census', ')']  
 ['Hozo', 'Region', 'in', 'Benishangul-Gumuz', ',', 'South', 'of', 'Asosa']  
 ['Kachama-Ganjule', 'Total', 'speakers', '2', ',', '829', '(', '2007', 'census', ')']  
 ['Kacipo-Balesi', 'Region', 'Southern', 'Ethiopia-Sudan', 'border', ',', 'Boma', 'Plateau', 'in', 'Sudan']  
 ['Kafa', 'Total', 'speakers', '834', ',', '918', '(', '2007', 'census', ')']  
 ['Komo', 'Region', 'North', 'of', 'Gambella', ',', 'across', 'the', 'Ethiopia-Sudan', 'border']  
 ['Loarki', 'Total', 'speakers', '20', ',', '000', '(', '1998', ')']  
 ['Ngarla', 'Spoken', 'in', 'Port', 'Hedland', 'area', 'of', 'Western', 'Australia']  
 ['Warumungu', 'Total', 'speakers', '200', '(', 'as', 'of', '1983', ')']  
 ['Ogba', 'Total', 'speakers', '170', ',', '000', '(', 'as', 'of', '1993', ')']  
 ['Kilit', 'language', 'Language', 'extinction', '20th', 'century']  
 ['Waibuk', 'Total', 'speakers', '1', ',', '000', '(', '1988', ')']  
 ['Waibuk', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Pemon', 'Total', 'speakers', '30', ',', '000']  
 ['Singlish', 'Spoken', 'in', 'Singapore', ',', 'sometimes', 'Malaysia', 'and', 'Brunei']  
 ['Singlish', 'Total', 'speakers', '5', '000', '000']  
 ['Changzhou', 'dialect', 'Total', 'speakers', '3', '.', '5', 'million']  
 ['Changzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Coatl\xca1n', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '5', ',', '000']  
 ['Isthmus', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '20', ',', '000']  
 ['Totontepec', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '5', ',', '200']  
 ['Tlahuitoltepec', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '5', ',', '000']  
 ['Juquila', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '5', ',', '000']  
 ['Mazatl\xca1n', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '19', ',',



'211']  
 ['North', 'Central', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '13',  
 ',,', '000']  
 ['Quetzaltepec', 'Mixe', 'Total', 'speakers', '&', 'gt', ';', '6', ',,',  
 '700']  
 ['Idoma', 'Spoken', 'in', 'Benue', 'State', ',,', 'Central', 'Nigeria']  
 ['Idoma', 'Total', 'speakers', '250', ',,', '000', '(', '1991', ')']  
 ['Medieval', 'Greek', 'Language', 'extinction', 'developed', 'into',  
 'Modern', 'Greek', 'by', '1453']  
 ['Ossetian', 'Spoken', 'in', 'Russia', '(', 'North', 'Ossetia', ')']  
 ['Ossetian', 'Spoken', 'in', 'South', 'Ossetia', '(', 'recognized', 'as',  
 'independent', 'by', 'Russia', ',,', 'Nicaragua', ',,', 'Venezuela', 'and',  
 'Nauru', ')']  
 ['Ossetian', 'Total', 'speakers', 'c', '.', '641', ',,', '450']  
 ['Ossetian', 'Official', 'language', 'in', 'South', 'Ossetia', ',,',  
 'North', 'Ossetia']  
 ['Ossetian', 'Language', 'codes', 'Linguasphere', 'Ossetian', 'text',  
 'from', 'a', 'book', 'published', 'in', '1935', '.', 'Part', 'of', 'an',  
 'alphabetic', 'list', 'of', 'proverbs', '.', 'Latin', 'script']  
 ['Sikule', 'Total', 'speakers', '20', ',,', '000']  
 ['Simeulue', 'Total', 'speakers', '30', ',,', '000']  
 ['Cac\xc3\xa1n', 'Language', 'extinction', '18th', 'century']  
 ['Cac\xc3\xa1n', 'Spoken', 'in', 'Northern', 'Argentina', 'and', 'Chile',  
 'Ethnicity']  
 ['Amdo', 'Total', 'speakers', '800', ',,', '000', '(', '1987', ')']  
 ['Bengali', 'Region', 'Bihar', '.', 'Widely', 'spoken', 'in', 'New',  
 'Delhi']  
 ['Bengali', 'Official', 'language', 'in', 'India', '(', 'West', 'Bengal',  
 ',,', 'Tripura', 'and', 'Barak', 'Valley', ')', '(', 'comprising',  
 'districts', 'of', 'south', 'Assam-', 'Cachar', ',,', 'Karimganj', 'and',  
 'Hailakandi', ')', ',,']  
 ['Bengali', 'Official', 'language', 'in', 'Sierra', 'Leone', '(', 'to',  
 'honour', "Bangladesh's", 'contributions', 'to', 'UN', 'Peacekeeping',  
 'Mission', 'in', 'Sierra', 'Leone', ')']  
 ['Bengali', 'Regulated', 'by', 'Paschimbanga', 'Bangla', 'Akademi', '(',  
 'West', 'Bengal', ')']  
 ['Bengali', 'Language', 'codes', 'Linguasphere', '59-AAF-u', '(',  
 'including', 'Sylhetti', 'etc', '.', ',,', '30', 'varieties', ':', '59-AAF-  
 ua', 'to']  
 ['Bengali', 'Language', 'codes', 'Linguasphere', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['British', 'Romance', 'Language', 'extinction', 'Early', 'Middle', 'Ages']  
 ['British', 'Language', 'codes', 'ISO', '639-2']  
 ['Danish', 'Total', 'signers', '5', ',,', '000']

['Danish', 'Language', 'codes', 'ISO', '639-3', 'dsl', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']  
 ['Bangl', 'Total', 'speakers', '120', ',', ',', '000']  
 ['Nias', 'Total', 'speakers', '770', ',', ',', '000']  
 ['Nias', 'Region', 'Nias', 'and', 'Batu', 'Islands', ',', ',', 'North',  
 'Sumatra']  
 ['Warao', 'Total', 'speakers', '18', ',', ',', '000', '+']  
 ['Kembra', 'Total', 'speakers', '20', '(', '2000', 'Wurm', ')']  
 ['Kembra', 'Spoken', 'in', 'Western', 'New', 'Guinea']  
 ["Olu'bo", 'Total', 'speakers', '15', ',', ',', '000', '(', '1985', ')']  
 ['Dumpu', 'Total', 'speakers', '510', 'in', '2003']  
 ['Dumpu', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Folopa', 'Spoken', 'in', 'Kerabi', 'Valley', 'and', 'Southern',  
 'Highlands', 'Province']  
 ['Folopa', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Folopa', 'Total', 'speakers', '~', '3', ',', ',', '000', '(', '1985', ')']  
 ['Arhuaco', 'Total', 'speakers', '14', ',', ',', '800', '(', '2001', ')']  
 ['Erromintxela', 'Total', 'speakers', 'approx', '500-1', ',', ',', '000', 'in',  
 '1997']  
 ['Bud', 'Total', 'speakers', '180', ',', ',', '000', '(', '1991', ')']  
 ["S'gaw", 'Karen', 'Total', 'speakers', '1', ',', ',', '284', ',', ',', '000']  
 ['Makonde', 'Total', 'speakers', '1', ',', ',', '340', ',', ',', '000']  
 ['Karamay', 'dialect', '/', 'North', 'Xinjiang', 'Area', 'dialect',  
 'Total', 'speakers', '\xe2', '\x89', '\xa4', '500', ',', ',', '000']  
 ['Karamay', 'Spoken', 'in', 'Northern', 'Xinjiang', ',', ',', 'China']  
 ['Karamay', 'Region', 'Northern', 'Xinjiang', ',', ',', 'China', ';']  
 ['Suki', 'Total', 'speakers', '3', ',', ',', '512']  
 ['Suki', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Mednyj', 'Aleut', 'Total', 'speakers', '5', '(', '2004', ')']  
 ['Kumbainggar', 'Total', 'speakers', '33', '(', '2006', ')']  
 ['Kelabit', 'Total', 'speakers', '1750', '(', 'year', '2000', ')']  
 ['Malagasy', 'Total', 'speakers', '+', '20', 'million']  
 ['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'bmm', '\xe2', '\x80',  
 '\x93', 'Northern', 'Betsimisaraka']  
 ['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'plt', '\xe2', '\x80',  
 '\x93', 'Plateau', 'Malagasy']  
 ['Malagasy', 'Language', 'codes', 'ISO', '639-3', 'bjq', '\xe2', '\x80',  
 '\x93', 'Southern', 'Betsimisaraka']  
 ['Zhongshan', 'Min', 'Total', 'speakers', '150', ',', ',', '000']  
 ['Zhenan', 'Min', 'Total', 'speakers', '574', ',', ',', '000']  
 ['Longyan', 'Min', 'Total', 'speakers', '740', ',', ',', '000', '(', 'approx',  
 '.', ')']  
 ['Quanzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Zhangzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Zhanjiang', 'Region', 'Mainly', 'in', 'Zhanjiang', ',', ',', 'southwestern',

'Guangdong', 'province']  
 ['Shantou', 'Region', 'Mainly', 'in', 'Shantou', ',', 'southwestern',  
 'Guangdong', 'province']  
 ['Medan', 'Hokkien', 'Total', 'speakers', '700', ',', '000']  
 ['Medan', 'Region', 'Medan', ',', 'as', 'well', 'as', 'other', 'northern',  
 'city', 'states', 'of', 'North', 'Sumatra', 'surrounding', 'it', 'where',  
 'Chinese', 'communities', 'reside']  
 ['Sanxiang', 'Region', 'Mainly', 'in', 'Sanxiang', ',', 'southern',  
 'Guangdong', 'province']  
 ['Nyangatom', 'Total', 'speakers', '24', ',', '303', '(', '2007', 'census',  
 ')']  
 ["Hui'an", 'Region', 'Mainly', 'in', "Hui'an", ',', 'southern', 'Fujian',  
 'province']  
 ['Nanlang', 'Region', 'Mainly', 'in', 'Nanlang', ',', 'southern',  
 'Guangdong', 'province']  
 ['Telefol', 'Total', 'speakers', '5', ',', '400']  
 ['Telefol', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Uropi', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Uropi', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Fali', 'Total', 'speakers', '36', ',', '000', '(', 'in', '1982', ')']  
 ['Fali', 'Language', 'codes', 'ISO', '639-3', ':', 'fal', '\xe2', '\x80',  
 '\x93', 'South', 'Fali']  
 ['Fali', 'Language', 'codes', 'ISO', '639-3', 'fl1', '\xe2', '\x80',  
 '\x93', 'North', 'Fali']  
 ['Akpes', 'Total', 'speakers', '10', ',', '000']  
 ['(North)', 'Akoko', 'Total', 'speakers', '50', ',', '000']  
 ['Oko', 'Total', 'speakers', '10', ',', '000']  
 ['Western', 'Total', 'speakers', '7', ',', '000']  
 ['Sukuma', 'Total', 'speakers', '5', ',', '430', ',', '000', '(', '2006',  
 ')']  
 ['Nyamwezi', 'Total', 'speakers', '980', ',', '000', '(', '2006', ')']  
 ['Datooga', 'Total', 'speakers', '90', ',', '000', '(', '2000', ')']  
 ['Cajamarca', 'Total', 'speakers', '10', ',', '000-30', ',', '000']  
 ['Inkawasi-Ka\xc3\xblaris', 'Total', 'speakers', '30', ',', '000']  
 ['Chachapoyas', 'Total', 'speakers', '6', ',', '000']  
 ['Lamas', 'Total', 'speakers', '45', ',', '000']  
 ['Kele', 'Total', 'speakers', '160', ',', '000']  
 ['Kabalai', 'Total', 'speakers', '17', ',', '885', '(', '1993', ')']  
 ['Emiliano', 'Language', 'codes', 'Linguasphere', '51-AAA-oka', '.', '.',  
 '.', '-okh']  
 ['Romagnolo', 'Language', 'codes', 'Linguasphere', '51-AAA-oki', '.', '.',  
 '.', 'okl']  
 ['Gorgotoqui', 'Language', 'extinction', '17th', 'century']  
 ['Inga', 'Total', 'speakers', '25', ',', '000']

['Inga', 'Language', 'codes', 'ISO', '639-3', ':', 'inb', '\xe2', '\x80',  
 '\x93', 'Highland', 'Inga']  
 ['Gutian', 'Language', 'extinction', 'spoken', 'ca', '.', '2100', 'BCE']  
 ['Mbugu', 'Total', 'speakers', '7000', '(', '1997', ')']  
 ['Old', 'Occitan', ',', 'Old', 'Proven', '\xc3', '\xa7', 'al', 'Total',  
 'speakers', '\xe2', '\x80', '\x93']  
 ['Urapmin', 'Total', 'speakers', '366', '(', '2003', 'SIL', ')']  
 ['Urapmin', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Tifal,', 'Tifalmin', 'Total', 'speakers', '3', ',', '600', '(', '2003',  
 'SIL', ')']  
 ['Tifal,', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Quinigua', 'Language', 'codes', 'ISO', '639-3', 'The', 'location', 'of',  
 'Quinigua', 'relative', 'to', 'Tamaulipas', 'state']  
 ['Kh\xc3\xbcn', 'Writing', 'system', 'Tai', 'Tham', 'script', ',', 'Thai',  
 'script']  
 ['Kh\xc3\xbcn', 'Language', 'codes', 'ISO', '639-3', 'This', 'page',  
 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',  
 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']  
 ['Yemsa', 'Total', 'speakers', '92', ',', '177', '(', '2007', 'census',  
 ')']  
 ['Yemsa', 'Region', 'Oromia', '&', 'amp', ';', 'SNNPR']  
 ['Tangale', 'Total', 'speakers', '70', ',', '000', 'or', '130', ',', '000',  
 '(', '1995', ')']  
 ['Luwo', 'Total', 'speakers', '80', ',', '000', '(', '1983', 'census', ')']  
 ['Nukunu', 'Total', 'speakers', '&', 'lt', ';', '10']  
 ['Alorese', 'Total', 'speakers', '25', ',', '000']  
 ['Saliba', 'Total', 'speakers', '1', ',', '600', '(', '1993', ')']  
 ['Piaroa', 'Total', 'speakers', '15', ',', '000', 'Piaroa', ',', '2', ',',  
 '500', 'Maco']  
 ['Maco', 'Total', 'speakers', '2', ',', '500', '(', 'in', '2002', ')']  
 ['Fergana', 'Spoken', 'in', 'Central', 'Asia', '(', 'Fergana', 'Valley',  
 ')']  
 ['Hidatsa', 'Total', 'speakers', 'USA', ':', '508', '(', '2000', ')']  
 ['Olukonjo', 'Total', 'speakers', '361', ',', '709', '(', '1992', 'census',  
 ')']  
 ['Olukonjo', 'Region', 'East', 'Africa', '/', 'Central', 'Africa']  
 ['Quenya', 'Date', 'founded', 'c', '.', '1910', '\xe2', '\x80', '\x93',  
 '1973']  
 ['Quenya', 'Writing', 'system', 'Latin', 'alphabet', ',', 'Elvish',  
 'writing', 'systems', ':', '(', 'mainly', ')', 'tengwar', ',', 'and',  
 'sarati']  
 ['Quenya', 'Language', 'codes', 'ISO', '639-3', 'qya', 'Ah', '!', 'like',  
 'gold', 'fall', 'the', 'leaves', 'in', 'the', 'wind', ',', 'long', 'years',  
 'numberless', 'as', 'the', 'wings', 'of', 'trees', '!', 'The', 'beginning',  
 'of', 'the', 'Quenya', 'poem', 'Nam', '\xc3', '\xa1', 'ri', '\xc3', '\xab',

'written', 'in', 'tengwar', 'and', 'in', 'Latin', 'script']  
 ['Sindarin', 'Date', 'founded', 'c', '.', '1915', '\xe2', '\x80', '\x93',  
 '1973']  
 ['Sindarin', 'Category', '(', 'sources', ')', 'a', 'priori', 'language',  
 ', ', 'but', 'related', 'to', 'the', 'other', 'Elvish', 'languages', '.',  
 'Sindarin', 'was', 'influenced', 'primarily', 'by', 'Welsh']  
 ['Meadow', 'Spoken', 'in', 'Mariy', 'El', 'Autonomous', 'Republic', ', ',  
 'Russia']  
 ['Pannonian', 'Romance', 'Language', 'extinction', 'Early', 'Middle',  
 'Ages']  
 ['Taihu', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Jiangshan', 'dialect', 'Total', 'speakers', 'approx', '.', '500', ', ',  
 '000-600', ', ', '000']  
 ['Jiangshan', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Chuqu', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Kota', 'Total', 'speakers', '43', ', ', '497']  
 ['Usarufa', 'Spoken', 'in', 'Eastern', 'Highlands', 'Province']  
 ['Usarufa', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Usarufa', 'Total', 'speakers', '1', ', ', '300', '(', '1996', 'SIL', ')']  
 ['Sami', 'Total', 'speakers', 'Approximately', '20', ', ', '000', '-', '30',  
 ', ', '000']  
 ['Sami', 'Language', 'codes', 'ISO', '639-1', 'se', '(', 'Northern',  
 'Sami', ')']  
 ['Bolak', 'Category', '(', 'purpose', ')', 'International', 'auxiliary',  
 'language']  
 ['Bolak', 'Category', '(', 'sources', ')', 'Mixed', 'a', 'priori', 'and',  
 'a', 'posteriori', 'elements', '; ', 'vocabulary', 'from', 'Romance', 'and',  
 'Germanic', 'languages', '; ', 'phonology', 'from', 'French', 'language']  
 ['Bolak', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Idu', 'Spoken', 'in', 'India', ': ', 'Assam', '; ', 'Arunachal', 'Pradesh',  
 ': ', 'Dibang', 'Valley', 'District', '; ', 'West', 'Bengal', '.', 'China',  
 ': ', 'Southeastern', 'Tibet', ': ', 'Nyingchi', 'Prefecture', ': ', 'Zayu',  
 'County', '; ', 'Western', 'Yunnan']  
 ['Idu', 'Total', 'speakers', 'Idu', 'Mishmi', 'people', '(', 'Both',  
 'Categorized', 'as', 'Lhoba', 'and', 'Mishmi', ')']  
 ['Nung', 'Total', 'speakers', '380', '(', 'China', '\xef', '\xbc', '\x8c',  
 '2000', ')', '; ', '400', '(', 'Myanmar', ')']  
 ['Sakha', 'Total', 'speakers', 'ca', '.', '360', ', ', '000', '(', '1990s',  
 'estimate', ')']  
 ['Sakha', 'Language', 'codes', 'ISO', '639-3', 'Map', 'showing',  
 'locations', 'of', 'Sakha', '(', 'dark', 'blue', ')', 'and', 'Dolgan', '(',  
 'blue', ')']  
 ['High', 'Category', '(', 'purpose', ')', 'High', 'Speech']  
 ['Aromanian', 'Total', 'speakers', '300', ', ', '000']  
 ['Aromanian', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Aromanian',

'variant', ')']  
 ['Aromanian', 'Official', 'language', 'in', 'Recognised', 'as', 'minority',  
 'language', 'in', 'parts', 'of']  
 ['Bafanji', 'Total', 'speakers', '17', ',', '000', '(', '2008', 'est', '.',  
 ')']  
 ['Zeem', 'Total', 'speakers', '400', '(', 'As', 'of', '2003', ')']  
 ['Mwani', 'Total', 'speakers', '1st', 'language', ':', '100', 'thousand',  
 ';', '2nd', 'language', ':', '20', 'thousand']  
 ['Kemiehua', 'Total', 'speakers', '1', ',', '000', '(', '1991', ')']  
 ['Chilisso', 'Total', 'speakers', 'ca', '.', '2', ',', '300', '(', '1992',  
 ')']  
 ['Takuu', 'Total', 'speakers', '1', ',', '750', '(', '2003', 'SIL', ')']  
 ['Takuu', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Takuu', 'Region', 'Takuu', '(', 'Mortlock', 'atoll', ')']  
 ['Abip\xc3\xb3n', 'Total', 'speakers', 'Extinct', 'since', '1770s']  
 ['Cia-Cia', 'Total', 'speakers', '80', ',', '000', '(', 'as', 'of', '2005',  
 ')']  
 ['Cia-Cia', 'Writing', 'system', 'Latin', 'Alphabet', '(', 'also', 'used',  
 ')']  
 ['Lakon', 'Total', 'speakers', 'c', '.', '700']  
 ['Pnar', 'Total', 'speakers', '84', '000', 'in', 'India', '(', '1991', ')',  
 ',', '4000', 'in', 'Bangladesh', '(', '2002', ')']  
 ['Highland', 'Total', 'speakers', 'ca', '20', '.', '000']  
 ['Temoaya', 'Total', 'speakers', 'ca', '37', '.', '000']  
 ['Yodz', 'Total', 'speakers', '2', ',', '000']  
 ['Anuki', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Anuki', 'Region', 'Milne', 'Bay', 'Province', ',', 'tip', 'of', 'Cape',  
 'Vogel']  
 ['Doga', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Doga', 'Region', 'Milne', 'Bay', 'Province', ',', 'tip', 'of', 'Cape',  
 'Vogel']  
 ['Fuliiru', 'Total', 'speakers', '300', ',', '000', '(', '1999', 'SIL',  
 ')']  
 ['Fuliiru', 'Region', 'East', 'Africa', '/', 'Central', 'Africa']  
 ['Sampang', 'Total', 'speakers', '10', '.', '000']  
 ['Ga'dang', 'Total', 'speakers', '6', ',', '000']  
 ['Pyu', 'Total', 'speakers', '100', '(', '2000', ')']  
 ['Pyu', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Tinigua', 'Region', 'Colombia', ';', 'Sierra', 'de', 'la', 'Macarena']  
 ['Narom', 'Total', 'speakers', '2', ',', '420']  
 ['Jinxiang', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Kurt\xc3\xb6p', 'Total', 'speakers', '10', ',', '000']  
 ['Lou', 'Total', 'speakers', '~', '1', ',', '000']  
 ['Lou', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Sinauna', 'language', 'Total', 'speakers', '2530']  
 ['Jek', 'language', 'Total', 'speakers', '1', ',', '500', '+']

['Colville-Okanagan', 'Region', 'Southern', 'Interior', 'of', 'British',  
 'Columbia', ',', 'Central-northern', 'State', 'of', 'Washington']  
 ['Alekanó', 'Spoken', 'in', 'Eastern', 'Highlands', 'Province']  
 ['Alekanó', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Alekanó', 'Total', 'speakers', '25', ',', '000', '(', '1999', 'SIL', ')']  
 ['Fwe', 'Total', 'speakers', '10', ',', '000']  
 ['Kuhane', 'Total', 'speakers', '35', ',', '000']  
 ['Kalto', 'Total', 'speakers', '15', ',', '000', '(', '2003', ')']  
 ['Nihali', 'Total', 'speakers', '2', ',', '000']  
 ['Fonz', 'Dialects', 'Low', 'Ribagor', '\xc3', '\xa7', 'an', 'Aragonese']  
 ['Folkspraak', 'Regulated', 'by', 'The', 'Folkspraak', 'Institute']  
 ['Folkspraak', 'Category', '(', 'purpose', ')', 'International',  
 'auxiliary', 'language']  
 ['Folkspraak', 'Setting', 'and', 'usage', 'International', 'auxiliary',  
 'language']  
 ['Mongolian', 'Language', 'codes', 'ISO', '639-3', 'msr', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Cilentan', 'Total', 'speakers', '~', '250', ',', '000']  
 ['Cilentan', 'Region', 'Cilento', '(', 'Southern', 'Campania', ')']  
 ['Waxianghua', 'Total', 'speakers', '300', ',', '000']  
 ['Ngwe', 'Total', 'speakers', '73', ',', '200SIL', ',', '2001']  
 ['Venezuelan', 'Language', 'codes', 'ISO', '639-3', 'vsl', 'sign',  
 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages',  
 '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Bakumpai', 'Total', 'speakers', '100', ',', '000']  
 ["Ma'anyan", 'Total', 'speakers', '150', ',', '000']  
 ['Lawangan', 'Total', 'speakers', '100', ',', '000']  
 ['Benga', 'Total', 'speakers', '4', ',', '100']  
 ['Kombe', 'Total', 'speakers', '40', ',', '000']  
 ['Seki', 'Total', 'speakers', '15', ',', '000']  
 ['Gongduk', 'Total', 'speakers', 'First', 'language', ':', '~', '1', ',', '000']  
 ['Central', 'Pomo', 'Total', 'speakers', '8']  
 ['Northeastern', 'Pomo', 'Total', 'speakers', '0']  
 ['Southeastern', 'Pomo', 'Total', 'speakers', '5']  
 ['Kyakhta', 'Language', 'extinction', 'early', '20th', 'century']  
 ['Kyakhta', 'Spoken', 'in', 'Russian', '\xe2', '\x80', '\x93', 'Chinese',  
 'border']  
 ['Botlikh', 'Total', 'speakers', '5500', '(', '2006', ')']  
 ['Chamalal', 'Total', 'speakers', '5000', '(', '1990', ')']  
 ['Karata', 'Total', 'speakers', '6400', '(', '2006', ')']  
 ['Varhadi', 'Total', 'speakers', 'over', '21', '.', '5', 'million']  
 ['Varhadi', 'Region', '1', '.', 'Widely', 'spoken', 'in', 'Vidarbha',  
 'region', 'of', 'the', 'state', 'of', 'Maharashtra', 'in', 'India']  
 ['Varhadi', 'Region', '2', '.', 'Also', 'spoken', 'in', 'the', 'Western',

'and', 'Southern', 'parts', 'of', 'Chhattisgarh', ',', '']  
 ['Varhadi', 'Region', '3', '.', 'South-western', 'parts', 'of', 'Madhya', 'Pradesh']  
 ['Varhadi', 'Region', 'and', '4', '.', 'Northern', 'part', 'of', 'Andhra', 'Pradesh']  
 ['Varhadi', 'Official', 'language', 'in', 'not', 'an', 'official', 'language']  
 ['Varhadi', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Kryts', 'Total', 'speakers', '6000', '(', '1975', ')']  
 ['Budukh', 'Total', 'speakers', '1000', '(', '1990', ')']  
 ['Itsari', 'Total', 'speakers', '2', ',', '000']  
 ['Kanikkaran', 'Total', 'speakers', '20', ',', '000']  
 ['Kanikkaran', 'Region', 'Tamil', 'Nadu', ',', 'Kanyakumari']  
 ['Musasa', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Himariote', 'Total', 'speakers', 'at', 'least', '8', ',', '000']  
 ['Gwahatike', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Nkore-Kiga', 'Total', 'speakers', '2', ',', '330', ',', '000']  
 ['Plains', 'Total', 'speakers', '1', '(', '1962', 'H', '.', 'Landar', 'in', 'Sebeok', '1977', ')']  
 ['Kasem', 'Total', 'speakers', '250', ',', '000']  
 ['Cahuilla', 'Total', 'speakers', '15', '(', 'as', 'of', '2000', ')']  
 ['Siyi', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Siyi', 'Language', 'codes', 'ISO', '639-3', 'Siyi', 'is', 'the', 'yellow', 'area', 'on', 'the', 'coast']  
 ['Vao', 'Spoken', 'in', 'Vao', 'Island', 'and', 'north', 'Malekula', ',', 'Vanuatu']  
 ['Vao', 'Total', 'speakers', '1', ',', '900', '(', '2001', ')']  
 ['Siamou', 'Total', 'speakers', '40', ',', '000', '(', '1999', ')']  
 ['Mian', 'Total', 'speakers', '3', ',', '500', '(', '2007', ')']  
 ['Mian', 'Total', 'speakers', '1', ',', '400', '(', '2000', ')']  
 ['Mian', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Intemelio', 'Total', 'speakers', 'approximately', '10', ',', '000']  
 ['Intemelio', 'Language', 'codes', 'Linguasphere', '51-AAA-ohb', '(', 'Imperia', ')']  
 ['Kuuk', 'Thaayorre', 'Total', 'speakers', '250', '(', '2006', ')']  
 ['Kuuk', 'Total', 'speakers', '150', '(', '1991', ')']  
 ['Kuuk', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Yirrk-Thangalkl', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Dhivehi', 'Total', 'speakers', '350000', '+']



['Dhivehi', 'Regulated', 'by', 'Dhivehi', 'Language', 'Committee', ',', 'Maldives', 'College', 'of', 'Higher', 'Education']  
 ['Benasquese', 'Dialects', 'Catalan', 'dialect', 'transition']  
 ['Kuuk', 'Yak', 'Total', 'speakers', '0']  
 ['Kuuk', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland']  
 ['Bankon', 'Total', 'speakers', '12', ',', '000', '(', '2001', ',', 'SIL', ')']  
 ['Lombi', 'Total', 'speakers', '3', ',', '000', '(', '2001', ',', 'SIL', ')']  
 ['U', 'Total', 'speakers', '40', ',', '000', '(', '1992', 'Ethnologue', ')']  
 ['Algerian', 'Language', 'codes', 'ISO', '639-3', 'asp', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Sursurunga', 'Total', 'speakers', '3', ',', '000']  
 ['IsiNgqumo', 'Spoken', 'in', 'South', 'Africa', ',', 'Zimbabwe']  
 ['Koraga', 'Total', 'speakers', '14', ',', '000', '(', 'SIL', '1981', ')']  
 ['Karamojong', 'Total', 'speakers', '370', ',', '000']  
 ['Duruwa', 'Total', 'speakers', '80', ',', '000', '(', 'SIL', '2002', ')']  
 ['Globish', 'Category', '(', 'sources', ')', 'vocabulary', 'from', 'English', ',', 'simplified', 'spelling', 'and', 'pronunciation', '(', 'related', ',', 'but', 'independent', 'from', 'standard', 'English', ')']  
 ['Forest', 'Total', 'speakers', '41', ',', '302']  
 ['Tundra', 'Total', 'speakers', '41', ',', '302']  
 ['Globish', 'Category', '(', 'sources', ')', 'vocabulary', 'from', 'a', 'list', 'of', '1500', 'English', 'words', ',', 'and', 'grammar', 'based', 'on', 'a', 'subset', 'of', 'standard', 'English', 'grammar']  
 ['Margi', 'Total', 'speakers', '160', ',', '000', '(', '2006', ')']  
 ['Konkani', 'Total', 'speakers', '3', '.', '6', 'million']  
 ['Konkani', 'Region', 'Konkani', 'is', 'also', 'spoken', 'in', 'Kenya']  
 ['Konkani', 'Language', 'codes', 'ISO', '639-3', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Isconahua', 'Total', 'speakers', '82', '(', '2000', ',', 'Ethnologue', ')']  
 ['Tanzanian', 'Language', 'codes', 'ISO', '639-3', 'tza', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Koyraboro', 'Senni', 'Total', 'speakers', '&', 'gt', ';', '400', ',', '000', '(', 'as', 'of', '2007', ')']  
 ['Koyraboro', 'Region', 'East', 'of', 'Timbuktu', ',', 'Gao']  
 ['(Batak)', 'Mandailing', 'Total', 'speakers', '1', ',', '100', ',', '000', '(', '2000', 'census', ')']  
 ['Itsekiri', 'Total', 'speakers', '510', ',', '000', '(', '1991', 'as', 'cited', 'in', 'Ethnologue', ')']

['Croatian', 'Total', 'speakers', '5', '.', '5', 'million', '(', '2001', ')']

['Croatian', 'Region', 'Central', 'Europe', ',', 'Southern', 'Europe']

['Croatian', 'Spoken', 'in', 'Romania', '(', 'Cara', '\xc5', '\x9f', '- Severin', 'County', ')']

['Croatian', 'Standard', 'forms', 'Standard', 'Croatian']

['Croatian', 'Standard', 'forms', 'Burgenland', 'Croatian']

['Croatian', 'Regulated', 'by', 'Institute', 'of', 'Croatian', 'Language', 'and', 'Linguistics', '(', 'Council', 'for', 'Standard', 'Croatian', 'Language', 'Norm', ')']

['Croatian', 'Language', 'codes', 'Linguasphere', 'part', 'of', '53-AAA-g']

['Mandalorian', 'Date', 'founded', 'started', 'in', '2005']

['Lombardic', 'Language', 'extinction', 'Middle', 'ages']

['Orsam\xc3\xa5l', 'Total', 'speakers', 'about', '500']

['Ngwo', 'Total', 'speakers', '50', ',', '547', '(', '2000', 'WCD', ')']

['North', 'Upper', 'Saxon', 'Total', 'speakers', '\xe2', '\x80', '\x93']

['Pokomo', 'Total', 'speakers', '63', ',', '000', '(', '2007', ')']

['Cal\xc3\xa3o', 'Total', 'speakers', '5000', '(', 'in', '2000', ')']

['French', 'Total', 'speakers', '356', ',', '950']

['Cayuvava', 'Language', 'extinction', 'No', 'known', 'speakers', 'but', 'with', 'a', 'remaining', 'ethnic', 'population', 'of', '794', 'inhabitants']

['Wiru', 'Spoken', 'in', 'Southern', 'Highlands', 'Province', ',', 'Papua', 'New', 'Guinea']

['Wiru', 'Total', 'speakers', '15', ',', '300', '(', '1981', ')']

['Greek', 'Total', 'signers', '42', ',', '600', '(', '1986', ')']

['Greek', 'Language', 'codes', 'ISO', '639-3', 'gss', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Galician', 'Total', 'speakers', '3', '\xe2', '\x80', '\x93', '4', 'million', '(', '500', ',', '000', 'emigrants', 'throughout', 'Ibero-America', 'and', 'Europe', ')']

['Galician', 'Writing', 'system', 'Latin', 'alphabet', '(', 'Galician', 'variant', ')']

['Galician', 'Official', 'language', 'in', 'Galicia', ',', 'Spain', ';', 'accepted', 'orally', 'as', 'Portuguese', 'by', 'the', 'European', 'Union', 'Parliament']

['Zanaki', 'Total', 'speakers', '100', ',', '000']

['Abaknon', 'Total', 'speakers', '21', ',', '400']

['Zialo', 'Total', 'speakers', '25', ',', '000']

['Chokwe', 'Total', 'speakers', '1', ',', '000', ',', '000', '+']

['Balboa', 'Creole', 'French', 'Total', 'speakers', 'virtually', 'extinct', ';', 'a', 'few', 'families', 'are', 'bilingual', 'in', 'either', 'English', ',', 'or', 'rarely', 'in', 'French']

['Balboa', 'Region', 'limited', 'to', 'quarters', 'of', 'Balboa', 'Island']

['Old', 'Gallo', 'Romance', 'Language', 'extinction', 'evolved', 'into',

'Old', 'French', 'and', 'other', 'Gallo-Romance', 'Languages', 'by', 'the', '10th', 'century']

['Kwara'ae', 'Total', 'speakers', '32', ',', '000']

['Jewish', 'Palestinian', 'Aramaic', 'Language', 'extinction', '7', 'century', 'CE']

['Saint-Barth\xc3\xa9lemy', 'Spoken', 'in', 'U', '.', 'S', '.', 'Virgin', 'Islands']

['Saint-Barth\xc3\xa9lemy', 'Total', 'speakers', '&', 'lt', ';', '1000']

['Guro', 'Spoken', 'in', 'C', '\xc3', '\xb4', 'te', 'd'Ivoire']

['Guro', 'Total', 'speakers', '330', ',', '000', '(', '1993', ')']

['Bundjalung', 'Total', 'speakers', '10', '(', '1983', ')']

['Mamuju', 'Spoken', 'in', '4', '\xc2', '\xb0', 'S', '120', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '4', '\xc2', '\xb0', 'S', '120', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-4', ';', '120']

['Mamuju', 'Total', 'speakers', '50', ',', '000', '\xe2', '\x80', '\x93', '160', ',', '000']

['Ha', 'Total', 'speakers', '990', ',', '000', '(', '2001', ')']

['Nyakyusa', 'Total', 'speakers', '~', '1', '.', '2', 'million', '(', '800', ',', '000', 'in', 'Tanzania', 'in', '2006', ',', '300', ',', '000', 'in', 'Malawi', 'in', '1993', ')']

['Hehe', 'Total', 'speakers', '800', ',', '000', '(', '2006', ')']

['Bena', 'Total', 'speakers', '670', ',', '000', '(', '2001', ')']

['Kansa', 'Language', 'extinction', 'c', '.', '1983']

['Ohlone', 'Language', 'codes', 'ISO', '639-3', ':', 'cst', '\xe2', '\x80', '\x93', 'Northern', 'Ohlone']

['Ohlone', 'Language', 'codes', 'ISO', '639-3', 'css', '\xe2', '\x80', '\x93', 'Southern', 'Ohlone']

['Ohlone', 'Language', 'codes', 'ISO', '639-3', 'Ohlone', 'languages', '.', 'These', 'boundaries', 'are', 'approximate']

['Tapachultec', 'Language', 'codes', 'ISO', '639-3', 'qcs', '\xe2', '\x80', '\x93', ']', '']

['Kaikadi', 'Total', 'speakers', '23', ',', '694']

['Larantuka', 'Total', 'speakers', '20', ',', '000']

['Obispe\xc3\xb1o', 'Spoken', 'in', 'Southern', 'Californian', 'coastal', 'areas']

['Songe', 'Total', 'speakers', '1', ',', '000', ',', '000', '(', '1991', ')']

['Nande', 'Total', 'speakers', '900', ',', '000', '(', '1991', ')']

['Chuwabu', 'Total', 'speakers', '950', ',', '000', '(', '2006', ')']

['Chopi', 'Total', 'speakers', '960', ',', '000', '(', '2006', ')']

['Yaka', 'Total', 'speakers', '900', ',', '000', '(', '2000', ')']

['Aka', 'Spoken', 'in', 'Central', 'African', 'Republic', ',', 'Republic', 'of', 'Congo']

['Aka', 'Total', 'speakers', '30', ',', '000']

['Mbatì', 'Total', 'speakers', '60', ',', '000']

['Gyele', 'Total', 'speakers', '4', ',', '300']  
 ['Baka', 'Total', 'speakers', '43', ',', '000']  
 ['Basa', 'Total', 'speakers', '100', ',', '000', '(', '1973', ')']  
 ['Luo', 'Total', 'speakers', '1', '(', '1995', ')']  
 ['Asoa', 'Total', 'speakers', '26', ',', '000', '(', '2000', ')']  
 ['Kresh', 'Total', 'speakers', '16', ',', '000', '(', '1987', ')']  
 ['Kango', 'Total', 'speakers', '2', ',', '000', '(', '1998', ')']  
 ['Kara', 'Spoken', 'in', 'Central', 'African', 'Republic', 'and', 'Sudan']  
 ['Kara', 'Total', 'speakers', '14', ',', '000', '(', '1996', ')']  
 ['Doba', 'Total', 'speakers', '200', ',', '000', '?', '(', '1969-2006', ')']  
 ['Moru', 'Total', 'speakers', '70', ',', '000', 'in', 'Sudan', '(', '1982', ')']  
 ['Jur', 'Total', 'speakers', '100', ',', '000', '(', '2004', ')']  
 ['Bongo', 'Total', 'speakers', '10', ',', '000', '(', '2000', ')']  
 ['Bongo', 'Total', 'speakers', '12', ',', '000', '(', '2000', ')']  
 ['Gaam', 'Total', 'speakers', '67', ',', '000', '(', '2000', ')']  
 ['Nyala', 'Total', 'speakers', '80', ',', '000', '(', '2007', ')']  
 ['Sila', 'Daju', 'Total', 'speakers', '63', ',', '000', 'in', 'Chad', '(', '2000', ')']  
 ['Daju', 'Mongo', 'Total', 'speakers', '34', ',', '000', '(', '1993', 'census', ')']  
 ['Mamaind\xc3\xaa', 'Total', 'speakers', '250', '(', '2009', ')']  
 ['Mamaind\xc3\xaa', 'Total', 'speakers', '330', '(', '2007', ')']  
 ['Ofay\xc3\xa9', 'Total', 'speakers', '15', '(', 'in', '2005', ')']  
 ['Nahuatl', 'Spoken', 'in', '(', 'Mexico', 'State', ',', 'Distrito', 'Federal', ',', 'Puebla', ',', 'Veracruz', ',', 'Hidalgo', ',', 'Guerrero', ',', 'Morelos', ',', '(', 'Tlaxcala', ')', 'Oaxaca', ',', 'Michoac', '\xc3', '\xa1', 'n', 'and', 'Durango', ')', ',', '']  
 ['Nahuatl', 'Total', 'speakers', '1', '.', '45', 'million', '(', '2000', ')', 'Nahua', 'peoples']  
 ['Nahuatl', 'Language', 'codes', 'ISO', '639-3', 'nci', '\xe2', '\x80', '\x93', 'Classical', 'Nahuatl']  
 ['Ngbaka', 'Total', 'speakers', 'ca', '.', '200', ',', '000']  
 ['Krongo', 'Total', 'speakers', '22', ',', '000', '(', '1984', ')']  
 ['Kadugli', 'Total', 'speakers', '75', ',', '000', '(', '2004', ')']  
 ['Dengebu', 'Total', 'speakers', '38', ',', '000', 'including', 'Ngile', '(', '1982', ')']  
 ['Lumun', 'Spoken', 'in', 'Moro', 'Hills', ',', 'Sudan']  
 ['Lumun', 'Total', 'speakers', '45', ',', '000', '(', '1980', ')']  
 ['Tegali', 'Total', 'speakers', '36', ',', '000', '(', '1984', ')']  
 ['Tingal', 'Total', 'speakers', '8', ',', '000', '(', '1982', ')']  
 ['Mayogo', 'Total', 'speakers', '100', ',', '000', '(', '1991', ')']  
 ['Mba', 'Total', 'speakers', '36', ',', '000', '(', '2000', ')']  
 ['Wahgi', 'Spoken', 'in', 'Western', 'Highlands', 'Province', ',', 'Papua', 'New', 'Guinea']

['Wahgi', 'Total', 'speakers', '39', ',', ',', '000', '(', '1999', ')']  
 ['Chrau', 'Total', 'speakers', '22', ',', ',', '000']  
 ['Amba', 'Total', 'speakers', '40', ',', ',', '100']  
 ['Deori', 'Total', 'speakers', '29', ',', ',', '600', '(', '2000', ')']  
 ['Nyala', 'Total', 'speakers', '35', ',', ',', '000']  
 ['Nyole', 'Total', 'speakers', '120', ',', ',', '000']  
 ['Noy', 'Total', 'speakers', '36', '(', '1993', 'Census', ')']  
 ['Noone', 'Total', 'speakers', '25', ',', ',', '000']  
 ['Noone', 'Writing', 'system', 'Latin', 'alphabet', ',', ',', 'literacy',  
 'rate', 'in', 'L2', ':', '25', '%', '\xe2', '\x80', '\x93', '50', '%']  
 ['Bidyara', 'Total', 'speakers', '20', '(', 'Wurm', 'and', 'Hattori',  
 '1981', ')']  
 ['Kurmali', 'Total', 'speakers', '37', ',', ',', '000']  
 ['Frainc-Comtou', 'Total', 'speakers', 'ca', '.', '1', ',', ',', '000']  
 ['Frainc-Comtou', 'Language', 'codes', 'Linguasphere', '51-AAA-ja', '&',  
 'amp', ';', '51-AAA-hc']  
 ['Sichuanese', 'Standard', 'Mandarin', 'Total', 'speakers', '\xe2', '\x80',  
 '\x93']  
 ['Kongu', 'Spoken', 'in', 'Tamil', 'Nadu', '-', 'Kongu']  
 ['Kongu', 'Official', 'language', 'in', 'India', '(', 'Tamil', 'Nadu', ')']  
 ['Kongu', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of',  
 'native', 'Tamil', 'speakers', 'in', 'India', 'and', 'Sri', 'Lanka',  
 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering',  
 'support', 'you', 'may', 'see', 'irregular', 'vowel', 'positioning', 'and',  
 'a', 'lack', 'of', 'conjuncts', '.', 'More']  
 ['Yangbye', 'Total', 'speakers', '1', ',', ',', '143', ',', ',', '000']  
 ['Dothraki', 'Date', 'founded', 'from', '2009']  
 ['Kwambi', 'Total', 'speakers', '32', ',', ',', '700']  
 ['Embu', 'Total', 'speakers', '429', ',', ',', '000', '(', '1994', 'census',  
 ')']  
 ['Embu', 'Region', 'Eastern', 'Province', ',', ',', 'Embu', 'District']  
 ['Gurindji', 'Spoken', 'in', 'Kalkaringi', '-', 'Northern', 'Territory']  
 ['Gurindji', 'Total', 'speakers', '540', '(', '1996', ')']  
 ['Seimat', 'Total', 'speakers', '1', ',', ',', '000', '(', '1992', ',', ',', 'SIL',  
 ')']  
 ['Seimat', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kaniet', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kaniet', 'Region', 'Kaniet', 'and', 'western', 'Anchorite', 'island',  
 'groups', ',', ',', 'Manus', 'Province']  
 ['Bipi', 'Total', 'speakers', '1', ',', ',', '200']  
 ['Bipi', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Bipi', 'Region', 'Bipi', 'and', 'Sisi', 'Islands', ',', ',', 'Manus',  
 'Province']  
 ['Nyindrou', 'Total', 'speakers', '4', ',', ',', '200']  
 ['Nyindrou', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Likum', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Hermit', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Sori-Harengan', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Khehek', 'Total', 'speakers', '1', ',', '600']  
 ['Khehek', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Tulu-Bohuai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Tangale', 'Total', 'speakers', '10', ',', '000', '(', '1990', ')']  
 ['Yat\xc3\xaa-Fulni\xc3\xb4', 'Total', 'speakers', '3000', '(', '1999', ')']  
 ['Western', 'Dialects', 'A', '\xc3', '\xad', 'sa', 'Aragonese', 'and', 'Jaca', 'Aragonese']  
 ['Western', 'Region', 'Jaca', 'and', 'Sabi', '\xc3', '\xb1', '\xc3', '\xa1', 'nigo']  
 ['Cheso', 'Aragonese', '/', 'Hecho', 'Aragonese', 'Total', 'speakers', '658']  
 ['Cheso', 'Language', 'codes', 'ISO', '639-2']  
 ['Rakhine', 'Total', 'speakers', '3', ',', '346', ',', '000', '(', '3', ',', '107', ',', '000', 'in', 'Burma', ';', '207', ',', '000', 'in', 'Bangladesh', ';', 'and', '32', ',', '000', 'in', 'India', ')']  
 ['Rakhine', 'Region', 'Rakhine', 'State', 'of', 'western', 'Burma']  
 ['Southern', 'Malaysia', 'Hokkien', 'Total', 'speakers', 'approximately', '1', ',', '500', ',', '000']  
 ['Kei', 'Total', 'speakers', '85', ',', '000', '(', 'in', '2000', ')']  
 ['Damana', 'Region', 'Sierra', 'Nevada', 'de', 'Santa', 'Marta']  
 ['Yugoslav', 'Total', 'signers', '30', ',', '000', '(', 'Van', 'Cleve', '1986', ')']  
 ['Yugoslav', 'Language', 'codes', 'ISO', '639-3', ':', 'csq', '\xe2', '\x80', '\x93', 'Croatian', 'SL']  
 ['Yugoslav', 'Language', 'codes', 'ISO', '639-3', 'ysl', '\xe2', '\x80', '\x93', 'Yugoslav', 'SL', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Kenati', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kenati', 'Total', 'speakers', '950', '(', '1990', ')']  
 ['Ngombe', 'Total', 'speakers', '1', ',', '400', '(', '1996', ')']  
 ['Ngombe', 'Total', 'speakers', '150', ',', '000', '(', '1971', ')']  
 ['Missouri', 'French', 'Language', 'extinction', 'Nearly', 'extinct']  
 ['Missouri', 'Region', 'State', 'of', 'Missouri', 'and', 'elsewhere', 'along', 'the', 'Mississippi', 'River', 'valley']  
 ['Marra', 'Total', 'speakers', '15', '(', '1991', ')']  
 ['Sangu', 'Total', 'speakers', '20', ',', '900', '(', '2000', ')']  
 ['Sangu', 'Total', 'speakers', '75', ',', '000', '(', '1987', ')']  
 ['Bwisi', 'Total', 'speakers', '4', ',', '250', '(', '2000', ')']  
 ['Luvale', 'Total', 'speakers', '632', ',', '000']  
 ['Yaruro', 'Total', 'speakers', '5', ',', '840', '(', '2001', ')']  
 ['Ayamaru', 'Total', 'speakers', '20', ',', '000', '(', '1987', ')']  
 ['Bangka', 'language', 'Total', 'speakers', '340', ',', '000']

['Jakun', 'language', 'Total', 'speakers', '27', ',', ',', '448']  
 ['Eastern', 'Total', 'speakers', '2', ',', ',', '031', ',', ',', '800']  
 ['Mandara', 'Total', 'speakers', '2', ',', ',', '500', '(', '1985', 'SIL', ')']  
 ['Mandara', 'Region', 'Tabar', 'Group', ',', ',', 'New', 'Ireland', 'Province']  
 ['Mandara', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Bariba', 'Total', 'speakers', '560', ',', ',', '000']  
 ['Talinga', 'Total', 'speakers', '99', ',', ',', '400', '(', '2002', ')']  
 ['South', 'Oran', 'and', 'Figuig', 'Berber', 'Total', 'speakers']  
 ['Taushiro', 'Total', 'speakers', '1', '(', '2000', ')']  
 ['Jarawa', 'Total', 'speakers', '150', ',', ',', '000']  
 ['Nsenga', 'Total', 'speakers', '752', ',', ',', '100']  
 ['Ternate', 'Total', 'speakers', '42', ',', ',', '000', '(', '1981', 'Wurm',  
 'and', 'Hattori', ')']  
 ['Ternate', 'Region', 'North', 'Maluku', ',', ',', 'islands', 'of', 'Ternate',  
 'and', 'some', 'areas', 'of', 'Halmahera', '.', 'Coordinates']  
 ['Ternate', 'Region', '0', '\xc2', '\xb0', '1', '\xe2', '\x80', '\xb2',  
 'N', '127', '\xc2', '\xb0', '44', '\xe2', '\x80', '\xb2', 'E', '\xef',  
 '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '0', '.', '017', '\xc2',  
 '\xb0', 'N', '127', '.', '733', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
 '\xbf', '/', '0', '.', '017', ';', '127', '.', '733']  
 ['Paya', 'Total', 'speakers', '990', '(', '1993', ')']  
 ['Paya', 'Region', 'North', 'central', 'coast', '(', 'Olancho',  
 'Department', ')']  
 ['Ombamba', 'Total', 'speakers', '24', ',', ',', '800']  
 ['Mondropolon', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['British', 'Language', 'extinction', 'Developed', 'into', 'Old', 'Welsh']  
 ['British', 'Language', 'extinction', 'Cornish', 'and', 'Breton', 'by',  
 '600', 'AD']  
 ['Bandi', 'Total', 'speakers', '107', ',', ',', '000']  
 ['Mann', 'Total', 'speakers', '259', ',', ',', '000']  
 ['Oued', 'Righ', 'Berber', 'Total', 'speakers', '6', ',', ',', '000', '(',  
 '1995', ')', '(', 'source', ':', 'ethnologue', '.', 'com', ')']  
 ['Oued', 'Region', 'Oued', 'Righ', '(', 'wilaya', 'of', 'Ouargla', ')']  
 ['Oued', 'Language', 'codes', 'ISO', '639-3', 'and', 'Oued', 'Righ']  
 ['Bayot', 'Total', 'speakers', '18', ',', ',', '790']  
 ['Kukuya', 'Spoken', 'in', 'Plateau', 'Province', ',', ',', 'Congo']  
 ['Kukuya', 'Total', 'speakers', '40', ',', ',', '000']  
 ['Koro', 'Total', 'speakers', '800', '\xe2', '\x80', '\x93', '1', ',', ',',  
 '200', '(', '2010', 'est', '.', ')']  
 ['Koro', 'Language', 'codes', 'ISO', '639-3', '1ei', '\xe2', '\x80',  
 '\x93']  
 ['Awtuw', 'Spoken', 'in', 'Sandaun', 'Province', ',', ',', 'Papua', 'New',  
 'Guinea']  
 ['Awtuw', 'Total', 'speakers', '510', '(', '2003', 'SIL', ')']  
 ['Lhokpu', 'Total', 'speakers', '2', ',', ',', '500', '(', '1993', ')']  
 ['Numanggang', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Numanggang', 'Total', 'speakers', '2', ',', '260', '(', '2000', 'census', ')']  
 ['Mwerlap', 'Total', 'speakers', 'c', '.', '1100']  
 ['Taznatit', 'Total', 'speakers', '80', ',', '000', '(', '1995', ')', '(', 'source', ':', 'ethnologue', '.', 'com', ')']  
 ['Ouargli', 'Total', 'speakers', '5', ',', '000', '(', '1995', ')', '(', 'source', ':', 'ethnologue', '.', 'com', ')']  
 ['Ouargli', 'Language', 'codes', 'ISO', '639-3', 'and', 'Oued', 'Righ']  
 ['Macedonian', 'Total', 'signers', '~', '6', '.', '000']  
 ['Macedonian', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Matmata', 'Total', 'speakers', '3', ',', '726', '(', '1975', ')', '(', 'source', ':', 'Sahli', '1983', ')']  
 ['Ong', 'Be', 'Total', 'speakers', '600', '000', '(', 'as', 'of', '2000', ')']  
 ['Ong', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Asu', 'Total', 'speakers', '5', ',', '000', '(', '1998', ',', 'SIL', ')']  
 ['Asu', 'Region', 'Niger', 'State', ',', 'Western', 'Nigeria']  
 ['Ilongot', 'Spoken', 'in', 'Eastern', 'Nueva', 'Vizcaya', 'and', 'Western', 'Quirino', ',', 'Philippines']  
 ['Ilongot', 'Total', 'speakers', '50', ',', '800']  
 ['Ambelau', 'Total', 'speakers', '5', ',', '700', '(', '1989', ')']  
 ['Buru', 'Total', 'speakers', '33', ',', '000', '(', '1989', ')']  
 ['Kayeli', 'Total', 'speakers', '3', '(', '1995', ')']  
 ['Lisela', 'Total', 'speakers', '11', ',', '900', '(', '1989', ')']  
 ['Cab\xc3\xa9car', 'Total', 'speakers', '8', ',', '840', '(', '2000', ')']  
 ['Amuzgo', 'Guerrero', 'Total', 'speakers', '23', ',', '000']  
 ['Kensiu', 'Region', 'near', 'Thai', 'border', '.', 'Overlaps', 'into', 'south', 'Yala', 'Province']  
 ['Laha', 'Total', 'speakers', '1', ',', '394', '(', '1989', 'Vietnam', 'census', ')']  
 ['Lachi', 'Total', 'speakers', '10', ',', '300']  
 ['Kannada', 'Total', 'speakers', '180', ',', '000', '(', '2000', ')']  
 ['Koro', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Maonan', 'Total', 'speakers', '30', ',', '000', '(', '2005', ')', ';;', 'ethnic', 'population', ':', '100', ',', '000']  
 ['Chadong', 'Total', 'speakers', '20', ',', '000', '(', '2007', ')']  
 ['Then', 'Total', 'speakers', '15', ',', '000', '(', '1999', ')']  
 ['Mulam', 'Total', 'speakers', '86', ',', '000', '(', '2005', ')', ';;', '&', 'lt', ';;', '10', ',', '000', 'monolinguals']  
 ['Mak', 'Total', 'speakers', '10', ',', '000', '(', '1982', ')']  
 ['Lakkia', 'Total', 'speakers', '12', ',', '000', '(', '1999', ')', ';;', '&', 'lt', ';;', '4', ',', '000', 'monolinguals']  
 ['Lakkia', 'Region', 'Jinxu', 'Yao', 'Autonomous', 'County', ',', 'Guangxi']



['Biao', 'Total', 'speakers', '80', ',', '000', '(', '1985', ',', 'Liang', 'Min', ')']  
 ['Biao', 'Region', 'Southwestern', 'Huaiji', 'County', ',', 'Guangdong']  
 ['ut-Ma\xe2\x80\x99in', 'Total', 'speakers', '36', ',', '000', '(', '1992', 'SIL', ')']  
 ['Warrwa', 'Region', 'West', 'Kimberley', ',', 'Derby', 'region', 'of', 'Western', 'Australia']  
 ['Mamanwa', 'Total', 'speakers', '5', ',', '150', '(', '1990', 'census', ')']  
 ['Mamanwa', 'Regulated', 'by', 'Commission', 'on', 'the', 'Filipino', 'Language']  
 ['Fipa', 'Total', 'speakers', '195', ',', '000']  
 ['Kwama', 'Total', 'speakers', '319', '(', '2007', 'census', ')']  
 ['Kwegu', 'Total', 'speakers', '4', ',', '924', '(', '2007', 'census', ')']  
 ['Mesqan', 'Total', 'speakers', '&', 'gt', ';;', '25', ',', '000', '(', 'Ethnologue', ')']  
 ['Shekkacho', 'Total', 'speakers', '80', ',', '162', '(', '2007', 'census', ')']  
 ['Zayse-Zergulla', 'Total', 'speakers', '18', ',', '512', '(', '2007', 'census', ')']  
 ['Zayse-Zergulla', 'Region', 'North', 'Omo', 'River', 'region', ',', 'west', 'of', 'Lake', 'Chamo']  
 ['Koorete', 'Total', 'speakers', '156', ',', '749', 'of', 'Koregna', ',', '2', ',', '473', 'of', 'Koyrigna', '(', '2007', 'census', ')']  
 ['Koorete', 'Region', 'East', 'of', 'Lake', 'Chamo']  
 ['Ait', 'Region', 'Central', 'Morocco', '\xe2', '\x80', '\x93', 'Middle', 'Atlas']  
 ['Makuv'a', 'Total', 'speakers', '&', 'lt', ';;', '50']  
 ['Fiuman', 'Total', 'speakers', 'between', '5000', 'and', '40000']  
 ['Proto-Kra', 'Spoken', 'in', 'Southern', 'China', ',', ]  
 ['Teanu', 'Total', 'speakers', '520', '(', '2007', ')']  
 ['Teanu', 'Region', 'Vanikoro', 'island', ',', 'Eastern', 'Solomons', 'Coordinates']  
 ['Teanu', 'Region', '11', '\xc2', '\xb0', '39', '\xe2', '\x80', '\xb2', 'S', '166', '\xc2', '\xb0', '54', '\xe2', '\x80', '\xb2', 'E', '\xef', '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '11', '.', '65', '\xc2', '\xb0', 'S', '166', '.', '9', '\xc2', '\xb0', 'E', '\xef', '\xbb', '\xbf', '/', '-11', '.', '65', ';;', '166', '.', '9']  
 ['Khuzdul', 'Date', 'founded', 'c', '.', '1935']  
 ['Khuzdul', 'Category', '(', 'sources', ')', 'influenced', 'by', 'Hebrew', 'in', 'phonology', 'and', 'morphology']  
 ['Balanta', 'Spoken', 'in', '(', 'Balanta-Kentohe', ')', 'Guinea-Bissau', ',', 'The', 'Gambia', '(', 'Balanta-Ganja', ')', 'Senegal']  
 ['Enggano', 'Region', 'Enggano', 'Island', ',', 'off', 'Sumatra']  
 ['Roglai', 'Total', 'speakers', '30', ',', '000', '(', 'Cobbey', '1977', ')']

['Krenak', 'Total', 'speakers', '150', '(', 'in', '1997', ')']  
 ['Aghwan', 'Language', 'codes', 'ISO', '639-3', 'A', '7th-century',  
 'column', 'capital', 'with', 'Caucasian', 'Albanian', 'text']  
 ['Kui', 'Spoken', 'in', 'East', 'Nusa', 'Tenggara']  
 ['Kui', 'Total', 'speakers', '4', ',', '240', '(', '2000', ')']  
 ['Pakanha', 'Total', 'speakers', '10', '(', '1981', ')']  
 ['Pakanha', 'Region', 'Cape', 'York', 'Peninsula', ',', 'Queensland',  
 'Coordinates']  
 ['Pakanha', 'Region', '14', '\xc2', '\xb0', '30', '\xe2', '\x80', '\xb2',  
 'S', '142', '\xc2', '\xb0', '25', '\xe2', '\x80', '\xb2', 'E', '\xef',  
 '\xbb', '\xbf', '/', '\xef', '\xbb', '\xbf', '14', '.', '5', '\xc2',  
 '\xb0', 'S', '142', '.', '417', '\xc2', '\xb0', 'E', '\xef', '\xbb',  
 '\xbf', '/', '-14', '.', '5', ';;', '142', '.', '417']  
 ['Phom', 'Total', 'speakers', '115', ',', '389', '(', '2001', ')']  
 ['Mushunguli', 'Total', 'speakers', '23', ',', '000', '(', '2006', ')']  
 ['Kposo', 'Total', 'speakers', '160', ',', '000', '(', '2002', ')']  
 ['Akebu', 'Total', 'speakers', '56', ',', '000']  
 ['Bunu', 'Total', 'speakers', '390', ',', '141']  
 ['Jiongnai', 'Total', 'speakers', '1', ',', '000']  
 ['Younuo', 'Total', 'speakers', '9', ',', '720']  
 ['Bunu', 'Total', 'speakers', 'about', '400', ',', '000']  
 ['Dzala', 'Total', 'speakers', '68', ',', '400', '(', '2006', ')']  
 ['Ba-Shu', 'Chinese', 'Language', 'extinction', 'Mostly', 'extinct',  
 'during', 'the', 'Ming', 'Dynasty', '.', 'Some', 'remnants', 'are',  
 'still', 'present', 'in', 'Sichuan', 'Mandarin']  
 ['Akose', 'Total', 'speakers', '200', ',', '000', '(', '2010', ')']  
 ['Olekha', 'Total', 'speakers', '1', ',', '000', '(', '1993', ')']  
 ['Moroccan', 'Arabic', 'Total', 'speakers', '21', 'million', '(', '1995',  
 ')']  
 ['Pijin', 'Total', 'speakers', '24', ',', '390', '(', '1999', 'SIL', ',',  
 'first', 'language', ')', ',', 'understood', 'by', 'many', 'more']  
 ['Banyumasan', 'Spoken', 'in', 'Western', 'Part', 'of', 'Central', 'Java',  
 '(', 'Indonesia', ')']  
 ['Min', 'Bei', 'Total', 'speakers', '10', '.', '3', 'million']  
 ['Min', 'Spoken', 'in', 'Southern', 'China', ',', 'United', 'States', '(',  
 'mostly', 'California', ')']  
 ['Min', 'Region', 'northwestern', '&', 'amp', ';;', 'central', 'Fujian',  
 ';;', 'Nanping']  
 ['Chocangacakha', 'Total', 'speakers', '20', ',', '000', '(', 'as', 'of',  
 '1993', ')']  
 ['Lakha', 'Total', 'speakers', '8', ',', '000', '(', 'as', 'of', '1993',  
 ')']  
 ['Brokkat', 'Total', 'speakers', '300', '(', 'as', 'of', '1993', ')']  
 ['Brokpa', 'Total', 'speakers', '5', ',', '000', '(', 'as', 'of', '2006',  
 ')']  
 ['Bumthang', 'Total', 'speakers', '36', ',', '500', '(', 'as', 'of',

'1993', ')']  
 ['Dakpa', 'Total', 'speakers', '1', ',', '000', '(', 'as', 'of', '1993', ')']  
 ['Chali', 'Total', 'speakers', '8', ',', '200', '(', 'as', 'of', '2006', ')']  
 ['Nyenkha', 'Total', 'speakers', '10', ',', '000', '(', 'as', 'of', '1993', ')']  
 ['Russian', 'Language', 'codes', 'ISO', '639-3', 'rsl', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Dargwa', 'Total', 'speakers', '439', ',', '000']  
 ['Gaulish', 'Language', 'extinction', 'After', 'the', '6th', 'century', 'contributed', 'to', 'Old', 'French']  
 ['Gaulish', 'Language', 'codes', 'ISO', '639-3', 'xcg', '\xe2', '\x80', '\x93', 'Cisalpine', 'Gaulish']  
 ['Terengganuan', 'Malay', 'Total', 'speakers', '1', '.', '1', 'million', '(', '2010', 'est', '.', ')']  
 ['Tepes', 'Total', 'speakers', '(', '5', ',', '000', 'in', '1972', ')']  
 ['Ik', 'Total', 'speakers', '10', ',', '000']  
 ['Lan\xe2\x80\x93Yin', 'Language', 'codes', 'ISO', '639-3', 'Lan-Yin', 'Mandarin', 'is', 'the', 'light', 'yellow']  
 ['Apinay\xc3\xa9', 'Total', 'speakers', '1', ',', '260']  
 ['Cubeo', 'Total', 'speakers', '4', ',', '632']  
 ['Lyons', 'Language', 'codes', 'ISO', '639-3', 'lsg', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Armenian', 'Language', 'codes', 'ISO', '639-3', 'aen', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Ecuadorian', 'Total', 'signers', '188', ',', '000', '(', 'Van', 'Cleve', '1986', ')']  
 ['Ecuadorian', 'Language', 'codes', 'ISO', '639-3', 'ecs', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Nepali', 'Total', 'signers', '5', ',', '743', '(', '2001', 'census', ')']  
 ['Nepali', 'Language', 'codes', 'ISO', '639-3', 'nsp', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Rennellese', 'Language', 'codes', 'ISO', '639-3', 'rsi', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Saudi', 'Signed', 'in', 'Saudi', 'Arabia']  
 ['Saudi', 'Language', 'codes', 'ISO', '639-3', 'sdl', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']  
 ['Sri', 'Signed', 'in', 'Sri', 'Lanka']

['Sri', 'Total', 'signers', 'unknown', 'number', 'of', '12', ',', '800', 'deaf', 'persons', '(', 'Van', 'Cleve', '1986', ')']

['Sri', 'Language', 'codes', 'ISO', '639-3', 'sq', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Peruvian', 'Language', 'codes', 'ISO', '639-3', 'prl', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Tijuana', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Croatian', 'Total', 'signers', '(', '30', ',', '000', 'all', 'dialects', 'of', 'YSL', ')']

['Croatian', 'Language', 'codes', 'ISO', '639-3', 'csq', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Swiss', 'Total', 'signers', '7', ',', '000', '(', '1986', ')']

['Swiss', 'Language', 'codes', 'ISO', '639-3', ':', 'sgg', '\xe2', '\x80', '\x93', 'Swiss', 'German', 'SL']

['Swiss', 'Language', 'codes', 'ISO', '639-3', 'ssr', '\xe2', '\x80', '\x93', 'Swiss', 'French', 'SL']

['Swiss', 'Language', 'codes', 'ISO', '639-3', 'slf', '\xe2', '\x80', '\x93', 'Swiss', 'Italian', 'SL', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Bugar', 'Total', 'speakers', '10', ',', '000', '(', '1993', ')']

['Plateau', 'Sign', 'Language', 'Language', 'extinction', '18th', 'century']

['Plateau', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Inuit', 'Language', 'codes', 'ISO', '639-3', 'None', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Mehek', 'Total', 'speakers', '6', ',', '300']

['Mehek', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Ugandan', 'Language', 'codes', 'ISO', '639-3', 'ugn', 'sign', 'language', '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80', '\x94', 'legal', 'recognition']

['Canara', 'Regulated', 'by', 'Karnataka', 'Konkani', 'Sahitya', 'Academy', ',', 'Kerala', 'Konkani', 'Academy']

['Canara', 'Language', 'codes', 'ISO', '639-3', 'knn', '\xe2', '\x80', '\x93', 'Konkani', '(', 'Independent', 'Language', ')']

['Canara', 'Language', 'codes', 'ISO', '639-3', 'Distribution', 'of', 'native', 'Canarese', 'Konkani', 'speakers', 'in', 'India', 'This', 'page', 'contains', 'Indic', 'text', '.', 'Without', 'rendering', 'support', 'you',

'may', 'see', 'irregular', 'vowel', 'positioning', 'and', 'a', 'lack',  
 'of', 'conjuncts', '.', 'More']

['Laua', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Laua', 'Total', 'speakers', '1', '(', '1987', 'SIL', ')']

['Nunivak', 'Cup'ig", 'language', 'Total', 'speakers', '\xe2', '\x80',  
 '\x93']

['Nunivak', 'Region', 'Central', 'Alaska', ', ', 'Nunivak', 'Island']

['Chevak', 'Region', 'Central', 'Alaska', '(', 'Chevak', ')']

['Aghem', 'Total', 'speakers', '27', ', ', '000']

['Vaccarizzo', 'Albanian', 'Total', 'speakers', 'c', '.', '3', ', ', '000']

['Vaccarizzo', 'Region', 'Calabria', '(', 'Vaccarizzo', 'Albanese', 'and',  
 'San', 'Giorgio', 'Albanese', ')']

['Kalix', 'Language', 'Total', 'speakers', '5', '000', '-', '10', '000']

['Kalix', 'Spoken', 'in', 'Kalix', 'Municipality', '(', 'Northern',  
 'Europe', ')']

['Kalix', 'Writing', 'system', 'Latin', 'Alphabet', '(', 'Kalix',  
 'variant', ')', ', ', 'IPA', '(', 'kj', '\xc5', '\x93', '\xc9', '\xad',  
 '\xc9', '\xaa', 's', 'variant', ')']

['If\xc3\xa8', 'Total', 'speakers', '182', ', ', '000', '(', '2002', ')']

['Tokano', 'Total', 'speakers', '6', ', ', '000', '(', '1982', ')']

['Tokano', 'Region', 'Eastern', 'Highlands', 'Province', ', ', 'Goroka',  
 'District']

['Tokano', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Zambian', 'Language', 'codes', 'ISO', '639-3', 'zsl', 'sign', 'language',  
 '\xe2', '\x80', '\x94', 'list', 'of', 'sign', 'languages', '\xe2', '\x80',  
 '\x94', 'legal', 'recognition']

['Zangwal', 'Total', 'speakers', '100', '(', '1993', ')']

['Old', 'Anatolian', 'Turkish', 'Language', 'extinction', 'Developed',  
 'into', 'the', 'Early', 'Ottoman', 'Turkish']

['Old', 'Writing', 'system', 'Arabic', 'alphabet', '(', 'Perso-Arabic',  
 'script', ')']

['Old', 'Language', 'codes', 'Linguasphere', 'part', 'of', '44-AAB-a']

['Burduna', 'Total', 'speakers', '3', '(', 'Wurm', 'and', 'Hattori',  
 '1981', ')']

['Nakkara', 'Total', 'speakers', '75', 'in', '1983']

['Andegerebinha', 'Total', 'speakers', '10', '(', 'Wurm', 'and', 'Hattori',  
 '1981', ')']

['Alumu-Tesu', 'Total', 'speakers', '7000', '(', '1999', ')']

['Aranadan', 'Total', 'speakers', '200', '(', '2001', ')']

['Ambrak', 'Total', 'speakers', '290', '(', '2000', ')']

['Ambrak', 'Spoken', 'in', 'Papua', 'New', 'Guinea']

['Aramanik', 'Total', 'speakers', '3000', '(', '2002', ')']

['Pamona', 'Total', 'speakers', '106', ', ', '000']

['Papantla', 'Total', 'speakers', '80', ', ', '000', '(', '1982', 'SIL',  
 ')']

['Mambila', 'Total', 'speakers', '130', ', ', '000', '(', '1993', ')']

['Mambila', 'Language', 'codes', 'ISO', '639-3', ':', 'mzk', '\xe2',  
 '\x80', '\x93', 'Nigerian', 'Mambila']  
 ['Western', 'Total', 'speakers', '180', ',', '000', '(', '1993', ')']  
 ['Terei', 'Total', 'speakers', '26', ',', '500', '(', '2003', ')']  
 ['Galela', 'Spoken', 'in', 'North', 'Halmahera', ',', 'Indonesia']  
 ['Galela', 'Total', 'speakers', '79', '\xe2', '\x80', '\x93', '94', ',', '000', '(', '1990', ')']  
 ['Tehit', 'Total', 'speakers', '10', ',', '000', '(', '2000', ')']  
 ['Amberbaken', 'Spoken', 'in', 'North', 'coast', 'of', 'Bird's', 'Head']  
 ['Hattam', 'Spoken', 'in', 'Eastern', 'Bird's', 'Head']  
 ['Hattam', 'Total', 'speakers', '16', ',', '000', '(', '1993', ')']  
 ['Abun', 'Spoken', 'in', 'Bird's', 'Head', 'Peninsula']  
 ['Abun', 'Total', 'speakers', '3', ',', '000', '(', '1995', ')']  
 ['Abun', 'Dialects', 'Abun', 'Tat', '(', 'Karon', 'Pantai', ')']  
 ['Abun', 'Dialects', 'Abun', 'Ji', '(', 'Madik', ')']  
 ['West', 'Makian', 'Total', 'speakers', '12', ',', '000', '(', '1977', ')']  
 ['West', 'Region', 'North', 'Maluku', ',', 'islands', 'of', 'Ternate',  
 'and', 'some', 'areas', 'of', 'Halmahera']  
 ['Morori', 'Total', 'speakers', '50', 'out', 'of', 'an', 'ethnic',  
 'population', 'of', '250']  
 ['Amung', 'Total', 'speakers', '14', ',', '000']  
 ['Dem', 'Total', 'speakers', '1000', '(', 'in', '1987', ')']  
 ['Tanah', 'Total', 'speakers', '500', '(', 'in', '1978', ')']  
 ['Karas', 'Total', 'speakers', '240', '(', 'in', '1983', ')']  
 ['Pawaia', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Pawaia', 'Total', 'speakers', '4000', '(', 'in', '1991', ')']  
 ['Kamula', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kamula', 'Total', 'speakers', '800', '(', 'in', '1998', ')']  
 ['Guhu-Samane', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Guhu-Samane', 'Total', 'speakers', '12', ',', '800']  
 ['Wersing', 'Total', 'speakers', '3', ',', '700', '(', '1997', ')']  
 ['Wersing', 'Language', 'codes', 'ISO', '639-3', 'kvw', 'Languages', 'of',  
 'the', 'Alor', 'Islands', '.', 'Wersing', 'is', 'shown', 'at', 'four',  
 'locations', 'around', 'the', 'east', 'coast']  
 ['Lele', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Lele', 'Total', 'speakers', '23', ',', '000']  
 ['Paynamar', 'Total', 'speakers', '150', '(', 'in', '1975', ')']  
 ['Massep', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Massep', 'Total', 'speakers', '&', 'lt', ';', '50']  
 ['Molof', 'Total', 'speakers', '200', '(', '1978', ')']  
 ['Samarokena', 'Total', 'speakers', '400', '(', '1982', ')']  
 ['Manggarainese', 'Total', 'speakers', '730', ',', '000']  
 ['Hagahai', 'Total', 'speakers', '600', '(', '1997', ')']  
 ['Hagahai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Hagahai', 'Region', 'Enga', '\xe2', '\x80', '\x93', 'Madang', 'border']  
 ['Tor', 'Total', 'speakers', '50', ',', '000', '(', '1998', ')']

['Lomwe', 'Total', 'speakers', '1', '.', '5', 'million', '(', '2006', ')']  
 ['Meax', 'Spoken', 'in', 'Bird's', 'Head', 'Peninsula']  
 ['Meax', 'Total', 'speakers', '15', ',', '000']  
 ['Rasawa', 'Total', 'speakers', '200', '(', '1987', ')']  
 ['Saponi', 'Total', 'speakers', '4', '(', '2000', ')']  
 ['Nakwi', 'Total', 'speakers', '280', '(', '2003', ')']  
 ['Nakwi', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Bo', 'Total', 'speakers', '85', '(', '1998', ')']  
 ['Bo', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Ama', 'Total', 'speakers', '480', '(', '1990', ')']  
 ['Ama', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Nimo', 'Total', 'speakers', '350', '(', '1998', ')']  
 ['Nimo', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Owiniga', 'Total', 'speakers', '330', '(', '1998', ')']  
 ['Owiniga', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Fas', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Fas', 'Total', 'speakers', '2', ',', '500']  
 ['Baibai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kwomtari', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kwomtari', 'Total', 'speakers', '600', '(', '1998', ')']  
 ['Guriaso', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Guriaso', 'Total', 'speakers', '160', '(', 'SIL', '2003', ')']  
 ['Kiwai', 'Total', 'speakers', '13', ',', '000', '(', '1975', '\xe2',  
 '\x80', '\x93', '1986', ')']  
 ['Kiwai', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Kiwai', 'Region', 'southern', 'Papua', 'New', 'Guinea']  
 ['Kiwai', 'Dialects', 'Southern', 'Coast', 'Kiwai']  
 ['Kiwai', 'Language', 'codes', 'ISO', '639-3', 'kjd', '\xe2', '\x80',  
 '\x93', 'Southern', 'Kiwai']  
 ['Abelam', 'Spoken', 'in', 'Sepik', 'River', 'basin', ',', 'Papua', 'New',  
 'Guinea']  
 ['Abelam', 'Total', 'speakers', '44', ',', '000', '(', '1991', ')']  
 ['Iatm\xc3\xbccl', 'Spoken', 'in', 'Sepik', 'River', 'basin', ',', 'Papua',  
 'New', 'Guinea']  
 ['Iatm\xc3\xbccl', 'Total', 'speakers', '8', ',', '400', '(', '2003', ')']  
 ['Boiken', 'Spoken', 'in', 'Sepik', 'River', 'basin', ',', 'Papua', 'New',  
 'Guinea']  
 ['Boiken', 'Total', 'speakers', '31', ',', '000', '(', '2003', ')']  
 ['Hawu', 'Total', 'speakers', '110', ',', '000', '(', '1997', ')']  
 ['Dhao', 'Total', 'speakers', '5', ',', '000', '(', '1997', ')']  
 ['Havu', 'Total', 'speakers', '500', ',', '000']  
 ['Jino', 'Spoken', 'in', 'southern', 'Yunnan', '(', 'People's', 'Republic',  
 'of', 'China', ')', 'Ethnicity']  
 ['Jino', 'Total', 'speakers', '14', ',', '000']  
 ['Tegem', 'Total', 'speakers', '600', '(', 'in', '2000', ')']  
 ['Baa', 'Total', 'speakers', '7', ',', '000']

['Tyefo', 'Total', 'speakers', '1', ',', '000', 'out', 'of', 'an',  
 'ethnic', 'population', 'of', '12', '\xe2', '\x80', '\x93', '15', ',', '000', 'in', '1995']  
 ['Vyemo', 'Total', 'speakers', '8', ',', '000', '(', '1995', ')']  
 ['Betoi', 'Language', 'extinction', '18th', 'century']  
 ['Huamo\xc3\xa9', 'Total', 'speakers', '0', 'out', 'of', 'an', 'ethnic',  
 'population', 'of', '2', ',', '500']  
 ['Namlá', 'Total', 'speakers', '30', 'in', '2005']  
 ['Maratino', 'Language', 'codes', 'ISO', '639-3', 'The', 'location', 'of',  
 'Maratino', 'in', 'Tamaulipas', 'state']  
 ['Naolan', 'Language', 'extinction', 'ca', '.', '1950']  
 ['Naolan', 'Language', 'codes', 'ISO', '639-3', 'The', 'location', 'of',  
 'Naolan', 'in', 'Tamaulipas', 'state']  
 ['Tux\xc3\xa1', 'Language', 'extinction', 'end', '19th', 'century']  
 ['Taruma', 'Language', 'extinction', 'a', 'few']  
 ['Huarpe', 'Language', 'extinction', 'ca', '.', '1630']  
 ['Chimane', 'Total', 'speakers', '5', ',', '320', '(', 'Adelaar', '2000',  
 ')']  
 ['Chimane', 'Dialects', 'Tsiman', '\xc3', '\xa9', '(', '90', '%', ')']  
 ['Chimane', 'Dialects', 'Santa', 'Ana', 'Moset', '\xc3', '\xa9', 'n']  
 ['Guaycura', 'Language', 'extinction', 'before', '1800']  
 ['Guaycura', 'Language', 'codes', 'ISO', '639-3', 'The', 'location', 'of',  
 'Guaycura', '.', 'Monqui', 'and', 'Peric', '\xc3', '\xba', 'are',  
 'essentially', 'unattested', ';', 'Cochim', '\xc3', '\xad']  
 ['Pankarar\xc3\xba', 'Language', 'extinction', 'ca', '.', 'early', '20th',  
 'century']  
 ['Alag\xc3\xbcilac', 'Language', 'codes', 'ISO', '639-3', '0im', '\xe2',  
 '\x80', '\x93']  
 ['Chol\xc3\xb3n', 'Language', 'extinction', 'ca', '.', '2000']  
 ['Hibito', 'Language', 'extinction', '?', '(', '500', 'in', '1850', ')']  
 ['Yar\xc3\xad', 'Spoken', 'in', 'Caquet', '\xc3', '\xa1', 'Department',  
 ',', 'Yar', '\xc3', '\xad', 'River', '&', 'amp', ';', 'upper', 'Vaup',  
 '\xc3', '\xa9', 's', 'River']  
 ['Nat\xc3\xba', 'Language', 'extinction', 'ca', '.', '1900', '?']  
 ['Nat\xc3\xba', 'Language', 'codes', 'ISO', '639-3', 'qjm', '\xe2', '\x80',  
 '\x93', '[', 'http', ':', '/', '/', 'multitree', '.', 'org', '/', 'codes',  
 '/', 'qjm', '.', 'html']  
 ['Tarairi\xc3\xba', 'Language', 'codes', 'ISO', '639-3', '0ix', '\xe2',  
 '\x80', '\x93']  
 ['Jingulu', 'Total', 'speakers', '10', '(', '1997', ')']  
 ['Jingulu', 'Total', 'speakers', '12-15', '(', '1997', ')']  
 ['Jingulu', 'Region', 'Barkly', 'Tableland', ',', 'Northern', 'Territory']  
 ['Nung', 'Total', 'speakers', '856', ',', '000', '(', '1999', 'census',  
 ')']  
 ['Guamo', 'Language', 'codes', 'ISO', '639-3', '0in', '\xe2', '\x80',  
 '\x93']



['Westrobothnian', 'Total', 'speakers', '5', '000', '-', '100', '000']  
 ['Westrobothnian', 'Spoken', 'in', 'Westrobothnia', '(', 'Northern',  
 'Europe', ')']  
 ['Westrobothnian', 'Writing', 'system', 'Latin', 'Alphabet', '(',  
 'Westrobothnian', 'variant', ')', ',', 'IPA']  
 ['Westrobothnian', 'Pronunciation', "'m", '\xc9', '\x94', '\xcb', '\x90',  
 '\xc9', '\xad', '\xc9', '\x9b']  
 ['Panzaleo', 'Language', 'extinction', '17th', 'century']  
 ['Ngarnka', 'Total', 'speakers', 'Extinct', 'as', 'of', '1998']  
 ['Ngarnka', 'Region', 'Barkly', 'Tableland', ',', 'Northern', 'Territory']  
 ['Kambiw\xc3\xa1', 'Language', 'extinction', 'early', '20th', 'century']  
 ['Biatah', 'Total', 'speakers', '72', ',', '000']  
 ['Kaimb\xc3\xa9', 'Language', 'extinction', 'early', '20th', 'century']  
 ['Natagaimas', 'Language', 'extinction', 'not', 'spoken', 'for', 'several',  
 'generations']  
 ['Shenenawa', 'Total', 'speakers', '360', 'in', '2002']  
 ['Aranama', 'Language', 'extinction', 'late', '19th', 'century']  
 ['Lepki', 'Total', 'speakers', '530', 'in', '1991']  
 ['Lepki', 'Spoken', 'in', 'Western', 'New', 'Guinea']  
 ['Murkim', 'Total', 'speakers', '290', 'in', '2004']  
 ['Murkim', 'Spoken', 'in', 'Western', 'New', 'Guinea']  
 ['Doso', 'Spoken', 'in', 'Papua', 'New', 'Guinea']  
 ['Doso', 'Total', 'speakers', '700', 'in', '1973']  
 ['Kehu', 'Total', 'speakers', '25', 'in', '2002']  
 ['Sorothaptic', 'Language', 'extinction', 'ca', '.', '200', 'CE']  
 ['Sorothaptic', 'Spoken', 'in', 'Iberian', 'peninsula', 'Ethnicity']  
 ['Kaskian', 'Language', 'extinction', 'Bronze', 'Age']  
 ['Buol', 'Total', 'speakers', '75', ',', '000']  
 ['Paleo-Sardinian', 'Language', 'extinction', 'ca', '.', '2nd', 'century']  
 ['Ta\xc3\xadno', 'Language', 'extinction', '(', 'survives', 'in', 'Gar',  
 '\xc3', '\xad', 'funa', ')']  
 ['Ta\xc3\xadno', 'Dialects', 'Descendent', ':', 'Garifuna', '(', 'modern',  
 'Central', 'America', ')']  
 ['Mara', 'Total', 'speakers', '15', '(', '1991', ')']  
 ['Palta', 'Language', 'extinction', '(', 'colonial', 'era', ')']  
 ['Kukur\xc3\xa1', 'Language', 'codes', 'ISO', '639-3', '0iv', '\xe2',  
 '\x80', '\x93']  
 ['Jeik\xc3\xb3', 'Language', 'extinction', '19th', 'century']  
 ['Jeik\xc3\xb3', 'Language', 'codes', 'ISO', '639-3', '0it', '\xe2',  
 '\x80', '\x93']  
 ['Tawasa', 'Language', 'extinction', '18th', 'century']  
 ['Tawasa', 'Language', 'codes', 'ISO', '639-3', 'tjm', '\xe2', '\x80',  
 '\x93', '(', 'Timucua', ')']  
 ['Tawasa', 'Language', 'codes', 'ISO', '639-3', 'Pre-contact',  
 'distribution', 'of', 'the', 'Timucua', 'language', '(', 'Florida', ')',  
 'and', 'Tawasa']

['Berik', 'Spoken', 'in', 'Tor', 'Atas', 'district', ',', 'Sarmi',  
 'Regency']  
 ['Cabiari', 'Total', 'speakers', '50', '(', 'dated', ')', '(', '1976',  
 ')']  
 ['Timote', 'Language', 'codes', 'ISO', '639-3', 'apj', '(', 'Timote', ')']  
 ['Jebero', 'Total', 'speakers', '2', ',', '500', '(', '2006', ')']  
 ['Sanapan\xc3\xa1', 'Total', 'speakers', '6', ',', '800']  
 ['Enxet', 'Total', 'speakers', '15', ',', '000']  
 ['Mawe', 'Total', 'speakers', '7', ',', '000']  
 ['Molala', 'Region', 'Central', 'Oregon', 'and', 'Washington', 'Ethnicity']  
 ['Cochim\xc3\xad', 'Language', 'extinction', 'Beginning', 'of', '20th',  
 'century']  
 ['Gaya', 'Language', 'extinction', '7th', 'century', '?']  
 ['Gaya', 'Language', 'codes', 'ISO', '639-3', 'The', 'Three', 'Kingdoms',  
 'of', 'Korea', ',', 'with', 'Gaya', '(', '"Kaya"', ')', 'in', 'pink']  
 ['Guambiano', 'Total', 'speakers', '24', ',', '000']  
 ['Guambiano', 'Region', 'Guambiano', '(', 'Misak', ')']  
 ['Mijikenda', 'Total', 'speakers', 'estimated', '360', ',', '000']  
 ['Lembena', 'Spoken', 'in', 'East', 'Sepik', 'Province']  
 ['Lembena', 'Total', 'speakers', '1', ',', '760', '(', '2000', ')']  
 ['Misima-Paneati', 'Spoken', 'in', 'Milne', 'Bay', 'Province', ',',  
 'Papua', 'New', 'Guinea']  
 ['Misima-Paneati', 'Total', 'speakers', '18', ',', '000', '(', '2002',  
 'SIL', ')', '.', '4', ',', '000', 'monolinguals']  
 ['Danezaa', 'Total', 'speakers', '300', 'in', '1991']  
 ['Taizhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Wuzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Xuanzhou', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Old', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Old', 'Language', 'codes', 'ISO', '639-3', 'Old', 'Xiang', 'is', 'in',  
 'yellow', '.', 'It', 'is', 'in', 'contact', 'with', 'New', 'Xiang', '(',  
 'orange', ')', ',', 'Ji-Xu', 'Xiang', '(', 'red', ')', ',', 'Southwestern',  
 'Mandarin', '(', 'blue', 'and', 'green', ',', 'north', ')', ',', 'Gan',  
 '(', 'purple', ')', ',', 'and', 'Hunan', 'Tuhua', '(', 'dark', 'green',  
 ',', 'south', ')']  
 ['Ji-Xu', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Ji-Xu', 'Language', 'codes', 'ISO', '639-3', 'Ji-Xu', 'Xiang', 'is',  
 'in', 'red', '.', 'It', 'is', 'in', 'contact', 'with', 'Old', 'Xiang', '(',  
 'yellow', ')', ',', 'Southwestern', 'Mandarin', '(', 'medium', 'and',  
 'light', 'blue', ')', ',', 'Gan', '(', 'purple', ')', ',', 'and',  
 'Waxianghua', '(', 'dark', 'blue', ')']  
 ['Yong-Xun', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Qin-Lian', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Yuehai', 'Spoken', 'in', 'People's', 'Republic', 'of', 'China']  
 ['Yuehai', 'Language', 'codes', 'ISO', '639-3', 'The', 'four', 'Yuehai',  
 'dialects']

['Yuehai', 'Language', 'codes', 'ISO', '639-3', 'are', 'marked', 'with',  
'a', "'Y'"]  
['Goulou', 'Spoken', 'in', "People's", 'Republic', 'of', 'China']  
['Luo-Guang', 'Spoken', 'in', "People's", 'Republic', 'of', 'China']  
['Luo-Guang', 'Language', 'codes', 'ISO', '639-3', 'Luo-Guang', 'is', 'at',  
'the', 'top', ',', 'in', 'blue']  
['Wu-Hua', 'Spoken', 'in', "People's", 'Republic', 'of', 'China']  
['Hainan', 'Spoken', 'in', "People's", 'Republic', 'of', 'China']  
['Hainan', 'Language', 'codes', 'ISO', '639-3', 'On', 'this', 'map', ',',  
'the', 'dialects', 'of', 'Hainan', 'are', 'classified', 'as', 'Yuehai']  
['Bima', 'Total', 'speakers', '500', ',', '000', '(', '1989', ')']  
['Sumbawa', 'Total', 'speakers', '300', ',', '000', '(', '1989', ')']