

In [1]: `import polyglot`

```
In [2]: from polyglot.downloader import downloader  
print(downloader.supported_languages_table("morph2"))
```

1. Kapampangan	2. Italian	3. Upper Sorbian
4. Sakha	5. Hindi	6. French
7. Spanish; Castilian	8. Vietnamese	9. Arabic
10. Macedonian	11. Pashto, Pushto	12. Bosnian-Croatian-Serbian
13. Egyptian Arabic	14. Norwegian Nynorsk	15. Sundanese
16. Sicilian	17. Azerbaijani	18. Bulgarian
19. Yoruba	20. Tajik	21. Georgian
22. Tatar	23. Galician	24. Malagasy
25. Uighur, Uyghur	26. Amharic	27. Venetian
28. Yiddish	29. Norwegian	30. Alemannic
31. Estonian	32. West Flemish	33. Divehi; Dhivehi; Maldive
34. Japanese	35. Ilokano	36. Haitian; Haitian Creole
37. Belarusian	38. Greek, Modern	39. Ossetian, Ossetic
40. Welsh	41. Malayalam	42. Albanian
43. Marathi (Marāṭhī)	44. Armenian	45. Slovene
46. Korean	47. Breton	48. Irish
49. Luxembourgish, Letzebuergesch	50. Bengali	51. Serbian
52. Fiji Hindi	53. Javanese	54. Finnish
55. Gan Chinese	56. Kirghiz, Kyrgyz	57. Catalan; Valencian
58. Quechua	59. Croatian	60. Dutch
61. Swedish	62. Ido	63. Tagalog
64. Sanskrit (Saṃskṛta)	65. Piedmontese language	66. Asturian
67. Danish	68. Cebuano	69. Western Frisian
70. Kannada	71. Scots	72. Maltese
73. Swahili	74. Limburgish, Limburgan	75. Waray-Waray
76. Lombard language	77. Uzbek	78. Kurdish
79. Latvian	80. Burmese	81. Aragonese
82. Volapük	83. Northern Sami	84. Faroese
85. Kazakh	86. Telugu	87. Ukrainian
88. Assamese	89. Chuvash	90. Silesian
91. Turkmen	92. Romanian, Moldavian, Moldovan	93. Persian
94. Tibetan Standard, Tibetan	95. Latin	96. Slovak
97. Sinhala, Sinhalese	98. Bavarian	99. Icelandic
100. Mongolian	101. Walloon	102. Portuguese
103. Urdu	104. Gujarati	105. Manx
106. Tamil	107. Khmer	108. English
109. Malay	110. Chechen	111. Bishnupriya Manipuri
112. Afrikaans	113. Basque	114. Polish
115. German	116. Esperanto	117. Indonesian
118. Occitan	119. Chinese	120. Czech
121. Hebrew (modern)	122. Romansh	123. Lithuanian
124. Turkish	125. Nepali	126. Bosnian
127. Interlingua	128. Zazaki	129. Oriya
130. Hungarian	131. Scottish Gaelic; Gaelic	132. Bashkir
133. Thai	134. Panjabi, Punjabi	135. Russian

```
In [6]: from polyglot.text import Text, Word
tokens=["unconditional", "precooked", "impossible", "painful", "entered"]
for s in tokens:
    s=Word(s, language="en")
    print("{:<20}{ }".format(s, s.morphemes))
```

```
unconditional      ['un', 'condition', 'al']
precooked           ['pre', 'cook', 'ed']
impossible          ['im', 'poss', 'ible']
painful             ['pain', 'ful']
entered             ['enter', 'ed']
```

```
In [15]: sent="Ihopeyoufindthebookinteresting"
para=Text(sent)
para.language="en"
para.morphemes
```

```
Out[15]: WordList(['I', 'hope', 'y', 'ou', 'find', 'the', 'book', 'interest', 'ing'])
```

```
In [16]: from polyglot.detect import Detector
arabic_text = u"""
أفاد مصدر امني في قيادة عمليات صلاح الدين في العراق بأن " القوات الامنية تتوقف لليوم
الثالث على التوالي عن التقدم الى داخل مدينة تكريت بسبب
انتشار قناصي التنظيم الذي يطلق على نفسه اسم "الدولة الاسلامية" والعبوات الناسفة
. والمنازل المفخخة والانتحاريين، فضلا عن ان القوات الامنية تنتظر وصول تعزيزات اضافية
"""
```

```
In [17]: detector = Detector(arabic_text)
print(detector.language)
```

```
name: Arabic      code: ar      confidence: 99.0 read bytes: 907
```

```
In [19]: telugu_text = "వార్తలు"
```

```
In [20]: detector = Detector(telugu_text)
print(detector.language)
```

```
name: Telugu      code: te      confidence: 100.0 read bytes: 1024
```

```
In [27]: sent="ఇక అనారోగ్యంతో బాధపడుతున్న పోసానిని వైఎస్సా కాంగ్రె పార్టీ ప్రధాన కార్యదర్శి సజ్జ
ల రామకృష్ణారెడ్డి ఆదివారం పరామర్శించారు. ఆయనతో పాటు వైఎస్సార్ రాష్ట్ర కార్యదర్శి, క
మెడియన్ పృథ్వీ పరామర్శించిన వారిలో ఉన్నారు. "
para=Text(sent)
para.language="te"
para.morphemes
```

```
Out[27]: WordList(['ఇక అనారోగ్యంతో ', 'బాధ', 'పడుతున్న', ' పోసానిని వైఎస్సా కాంగ్రె ',
'పార్టీ', ' ప్రధాన కార్యదర్శి సజ్జల ', 'రామ', 'కృష్ణ', 'ారెడ్డి', ' ఆదివారం పరామర్శ',
'ించారు', ' . ఆయనతో పాటు వైఎస్సార్\u200c ', 'రాష్ట్ర', ' కార్యదర్శి, కమెడియన్ ',
'పృథ్', 'వీ', ' పరామర్శించిన వారిలో ఉన్నారు. '])
```