Computational Linguistics, Linguistic Data and Theories

Parameswari Krishnamurthy

IIIT-Hyderabad

param.krishna@iiit.ac.in

January 3, 2025

Overview

Natural Language Processing Tasks

Core technologies

- Language modelling
- Part-of-speech tagging
- Syntactic parsing
- Named-entity recognition
- Coreference resolution
- Word sense disambiguation
- Semantic Role Labelling
- ...

Applications

- Machine Translation
- Information Retrieval
- Question Answering
- Dialogue Systems
- Information Extraction
- Summarization
- Sentiment Analysis
- ...

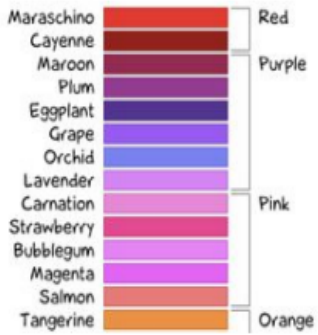
Why NLP is hard?

1. **Ambiguily** at many levels:
   - Word senses: bank (finance or river?)
   - Part of speech: chair (noun or verb?)
   - Syntactic structure: I saw a man with a telescope
   - Quantifier scope: Every child loves some movie
   - Multiple: I saw her duck

   $\Rightarrow$ NLP algorithms model ambiguity, and choose the correct analysis in context
2. **Linguistic diversity**

Linguistic Diversity: Semantics

Every language describes the world in a different way,
for example, it depends on culture or historical conditions.



- Russian has relatively few names for colors; Japanese has hundreds
- Multiword expressions, e.g. it's raining cats and dogs or wake up and metaphors, e.g. Love is a journey are very different across languages

Sapir-Whorf Hypothesis:
the language we speak both affects and reflects our view of the world

Linguistic Diversity: Language Families

www.ethnologue.com

1. Niger–Congo (1,538 languages) (20.6%)
2. Austronesian (1,257 languages) (16.8%)
3. Trans–New Guinea (480 languages) (6.4%)
4. Sino-Tibetan (457 languages) (6.1%)
5. Indo-European (444 languages) (5.9%)
6. Australian (378 languages) (5.1%)
7. Afro-Asiatic (375 languages) (5.0%)
8. Nilo-Saharan (205 languages) (2.7%)
9. Oto-Manguean (177 languages) (2.4%)
10. Austroasiatic (169 languages) (2.3%)
11. Volta Congo (108 languages) (1.5%)
12. Tai–Kadai (95 languages) (1.3%)
13. Dravidian (85 languages) (1.1%)
14. Tupian (76 languages) (1.0%)

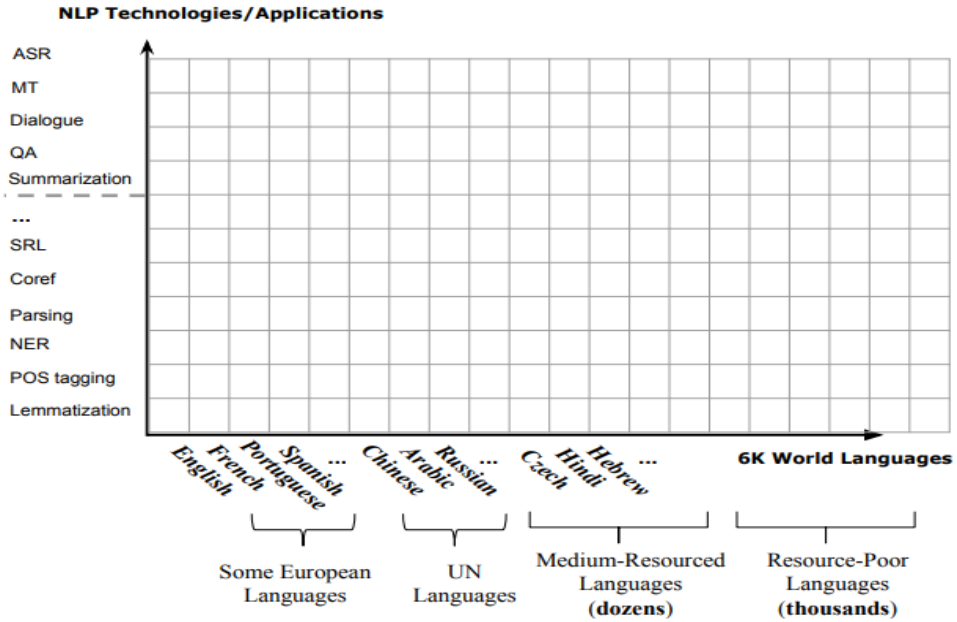
Why NLP is hard?

1. **Ambiguily** at many levels:
   - Word senses: bank (finance or river?)
   - Part of speech: blue (noun or verb?)
   - Syntactic structure: I saw a man with a telescope
   - Quantifier scope: Every child loves some movie
   - Multiple: I saw her duck

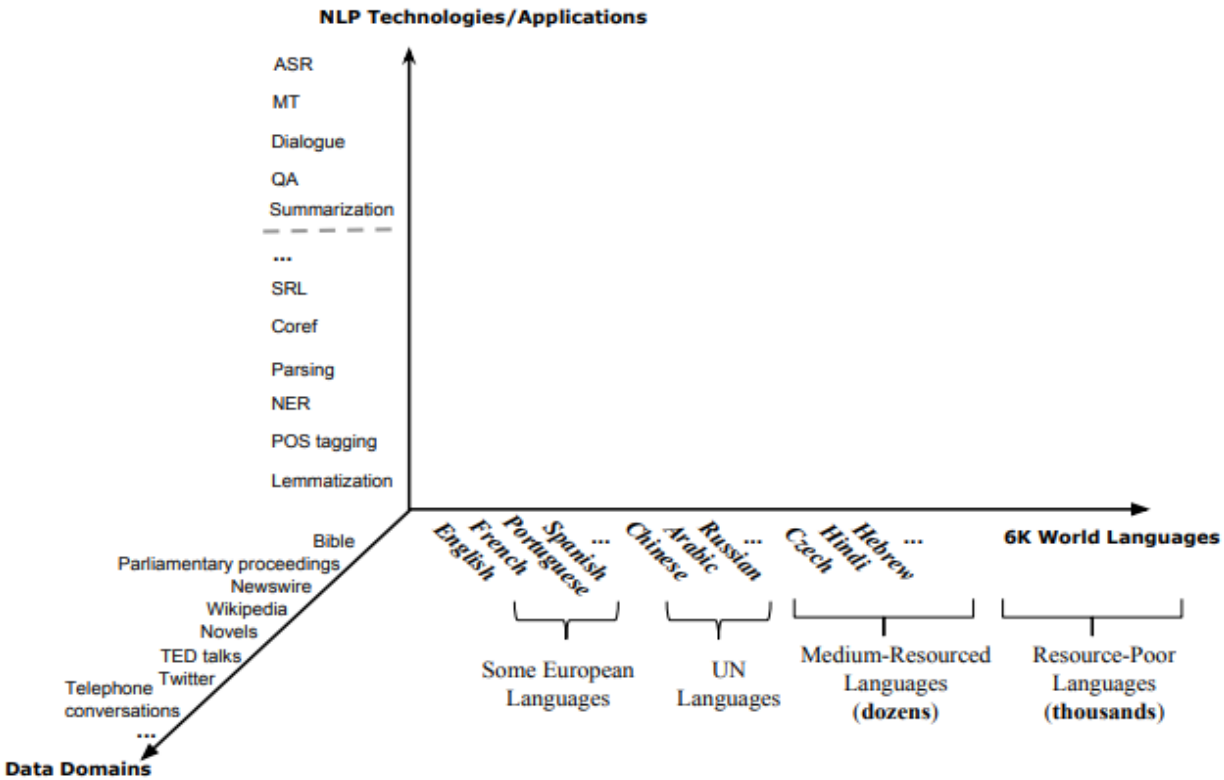
   $\Rightarrow$ NLP algorithms model ambiguity, and choose the correct analysis in context
2. **Linguistic diversity**
   - 6–7K languages in the world, $>$ 14 language families
   - Languages diverge across all levels of linguistic structure
     $\Rightarrow$ no generic solution for a particular NLP task
   - Most of the languages do not have sufficient resources to build statistical NLP models

Low-resource languages – languages lacking large monolingual or parallel corpora and/or manually crafted linguistic resources sufficient for building statistical NLP applications

What NLP Technologies are Resource-Rich?



Low-Resource NLP is Not Only About Multilinguality



- How do linguists, as language scientists, build up and formulate their theories when describing and explaining *What is human language and how language works according to the Scientific standards?*
- What kind of data would they examine?
- How do they analyse the data?
- What kind of methods and techniques of gathering data would they apply?
- What does it mean to "account for" the data with a theory

Language and Grammar

- Linguistic knowledge as in the speaker's mind is

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**
- Grammar includes:

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**
- Grammar includes:
  - phonetics and phonology (the sounds and the sound system or patterns)

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**
- Grammar includes:
  - phonetics and phonology (the sounds and the sound system or patterns)
  - morphology (the structure of words)

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**
- Grammar includes:
  - phonetics and phonology (the sounds and the sound system or patterns)
  - morphology (the structure of words) and lexicon (the words or vocabulary in the mental dictionary)

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**
- Grammar includes:
  - phonetics and phonology (the sounds and the sound system or patterns)
  - morphology (the structure of words) and lexicon (the words or vocabulary in the mental dictionary)
  - syntax (the structure of phrases and sentences, and the constraints on well-formedness of sentences)

Language and Grammar

- Linguistic knowledge as in the speaker's mind is **GRAMMAR**
- Need to understand the **mental grammar**
- Grammar includes:
  - phonetics and phonology (the sounds and the sound system or patterns)
  - morphology (the structure of words) and lexicon (the words or vocabulary in the mental dictionary)
  - syntax (the structure of phrases and sentences, and the constraints on well-formedness of sentences)
  - semantics (the meaning of words and sentences)

Grammar

- A tool to understand grammatical vs. ungrammatical sentences
  - well-formedness vs. ill-formedness

Grammar

- A tool to understand grammatical vs. ungrammatical sentences
  - well-formedness vs. ill-formedness

- **Grammatical sentences** accord with the rules and principles of the syntax of a particular language
  - Acceptable vs. Unacceptable

Grammar

- A tool to understand grammatical vs. ungrammatical sentences
  - well-formedness vs. ill-formedness

- **Grammatical sentences** accord with the rules and principles of the syntax of a particular language
  - Acceptable vs. Unacceptable
  - `I am happy` vs. `The pen is happy`
  - `My friend is pregnant` vs. `My friend's car is pregnant`
  - `Colourless green ideas sleep furiously`

- **Ungrammatical sentences** violate one or more syntactic rules or principles.
  - `*My students is angry`
  - `*There is water bottles on the table`
  - `*I saw girl the tall`

Methods: Inductive and Deductive

Different ways of building theories and use corresponding methods:

Methods: Inductive and Deductive

Different ways of building theories and use corresponding methods:

1. Inductive Method

Methods: Inductive and Deductive

Different ways of building theories and use corresponding methods:

1. Inductive Method

- begin by collecting observations or produce data by experiments

Methods: Inductive and Deductive

Different ways of building theories and use corresponding methods:

1. Inductive Method

- begin by collecting observations or produce data by experiments
- after a large and sufficient number of such observations or experiments, then proceeds to a generalization about these data

Methods: Inductive and Deductive

Different ways of building theories and use corresponding methods:

1. Inductive Method

- begin by collecting observations or produce data by experiments
- after a large and sufficient number of such observations or experiments, then proceeds to a generalization about these data
- After several attempts at generalizing, then may proceed to a new (modified) hypothesis by looking at new data
- start without any pre-knowledge or pre-conception about the linguistic object under examination

2. Deductive Method

2. Deductive Method

- The scientist has some ideas, some knowledge, or may be interested in a problem (including some knowledge) as the input to his/her theory–construction

2. Deductive Method

- The scientist has some ideas, some knowledge, or may be interested in a problem (including some knowledge) as the input to his/her theory–construction
- The scientist then formulates a first working hypothesis as a tentative answer to his problem

Methods: Inductive and Deductive

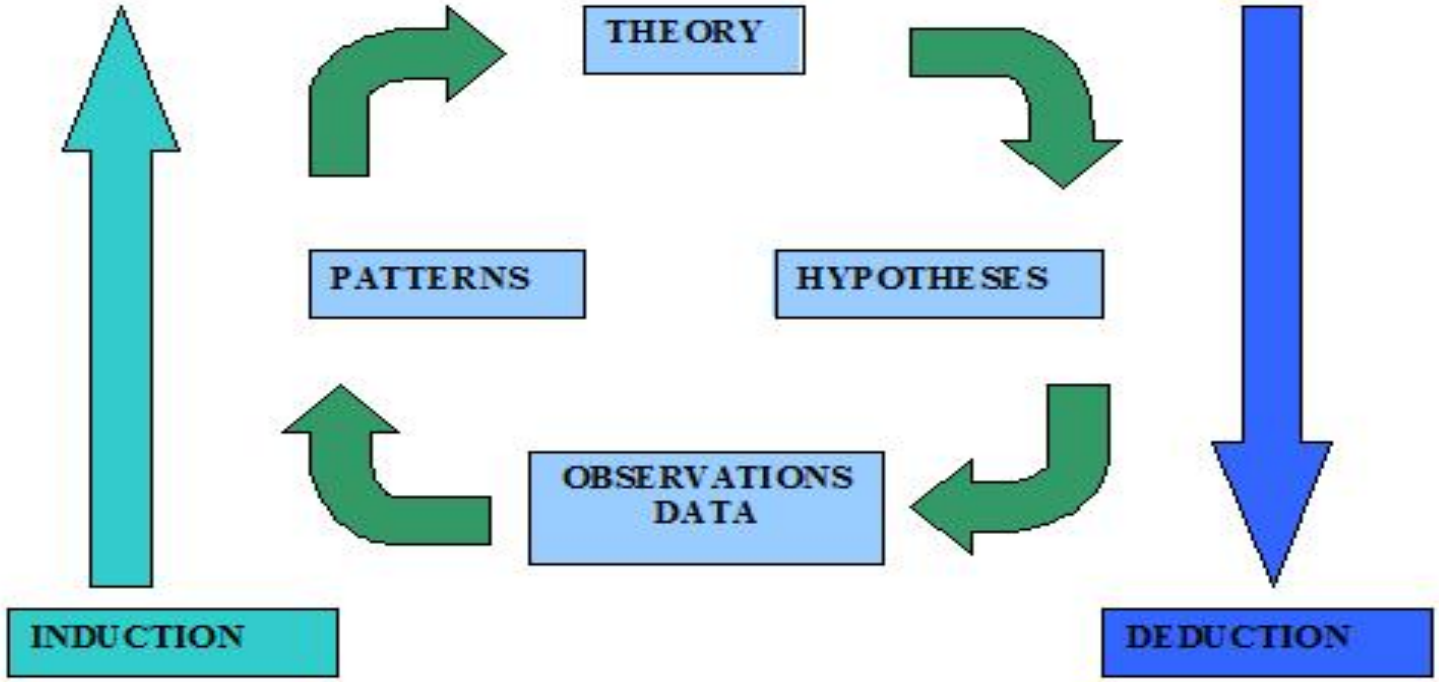
2. Deductive Method

- The scientist has some ideas, some knowledge, or may be interested in a problem (including some knowledge) as the input to his/her theory–construction
- The scientist then formulates a first working hypothesis as a tentative answer to his problem
- It is then tested against a collection of observations or experimental data and might be modified on this data

Methods: Inductive and Deductive

2. Deductive Method

- The scientist has some ideas, some knowledge, or may be interested in a problem (including some knowledge) as the input to his/her theory–construction
- The scientist then formulates a first working hypothesis as a tentative answer to his problem
- It is then tested against a collection of observations or experimental data and might be modified on this data
- it assumes that the hypothesis is derived (deduced) from already existent knowledge and then tested by empirical data


THEORY
PATTERNS
HYPOTHESES
OBSERVATIONS
DATA
INDUCTION
DEDUCTION

Corpus-Based Vs. Corpus Driven Methodology

1. Corpus-based Methodology

- corpus-based methodology essentially involves a deductive approach

Corpus-Based Vs. Corpus Driven Methodology

1. Corpus-based Methodology

- corpus-based methodology essentially involves a deductive approach
- here, a given corpus acts as a catalyst helping to confirm or refute a pre-existing theoretical construct

Corpus-Based Vs. Corpus Driven Methodology

1. Corpus-based Methodology

- corpus-based methodology essentially involves a deductive approach
- here, a given corpus acts as a catalyst helping to confirm or refute a pre-existing theoretical construct
- In other words, it tends to test the truth or falsity of a given hypothesis against a carefully chosen corpus data

2. Corpus-driven approach

- By contrast, the corpus-driven approach is inductive in that it chooses a targeted corpus and through cyclical analysis and interpretation of the corpus data

2. Corpus-driven approach

- By contrast, the corpus-driven approach is inductive in that it chooses a targeted corpus and through cyclical analysis and interpretation of the corpus data
- it tends to generalize a rule or some kind of theory

- Finally, when the research circumstances allow for both corpus based and corpus driven analyses
- the researcher may hybridize both approaches combining the merits of deduction and induction at the same time

Study on Syntax

- finding out the rule-governing system of sentence formation

Study on Syntax

- finding out the rule-governing system of sentence formation
- understanding how words and phrases are arranged

Study on Syntax

- finding out the rule-governing system of sentence formation
- understanding how words and phrases are arranged
- concerning the relationship between the finite and the infinite

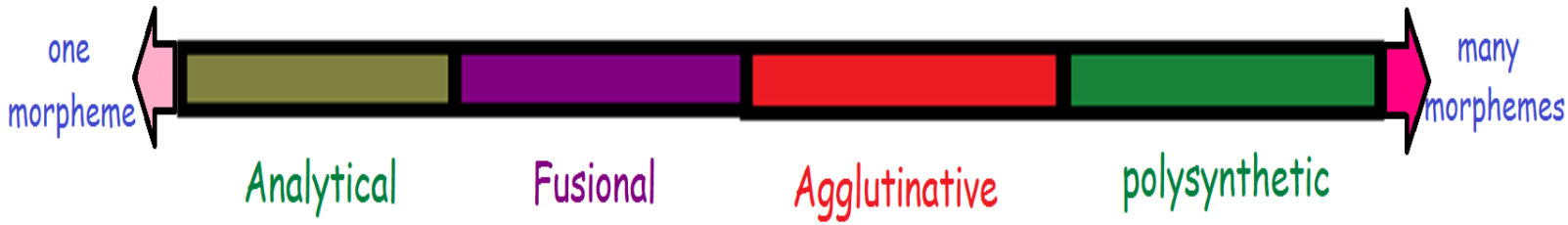
Analytic Vs.Synthetic Languages

Analytic languages

- language with a low morpheme-per-word ratio
- word order is significant
- E.g. Isolating languages

Synthetic languages

- language with high morpheme-per-word ratio
- word order is not significant and morphology is highly significant
- E.g. Agglutinating and fusional languages


one
morpheme
many
morphemes
Analytical
Fusional
Agglutinative
polysynthetic

Analytical languages are most common in Southeast Asia (Chinese, Vietnamese), but some such languages are also found among the Austronesian languages (Fijian, Tongan) and some Niger-Congo languages (Gbe, Yoruba).

Example from Vietnamese:

no se khong doc sach
he FUT NEG read book
'he will not read book'

Analytical languages are most common in Southeast Asia (Chinese, Vietnamese), but some such languages are also found among the Austronesian languages (Fijian, Tongan) and some Niger-Congo languages (Gbe, Yoruba).

Example from Vietnamese:

no se khong doc sach
he FUT NEG read book
'he will not read book'

Example from Chinese:

Ta ba shu mai le
He NOM book buy Asp
'He bought the book.'

Example for derivation from Telugu:
pagalagottiMcipettamananivvadalacukooleekapotunnaanu.
pagulu+a-kottu+iMcu+i-pettu+a-manu+a-ivvu+a-daluvu+i-konu+a-leeka-poo+
tunn+1,sg,any
break+inf-strike+cause+cpm-benefactive+inf-tell+inf-permit+inf-think+cpm-reflexive+inf-neg+go+prog+1,sg
'I could not think to permit someone to tell for my sake to break something'
(pc, Prof. G. Uma Maheshwar Rao)

Polysynthetic languages:

Inuktitut (Canada) for instance the word-phrase:
tavvakiqutiqarpiit
roughly translates to ´´Do you have any tobacco for sale?"

Yup'ik (Alaska):
angya-li-ciq- sugnar- quq-llu
boat- make-FUT- PROB- 3sg.NOM-also
'Also, he probably will make a boat'

Eskimo (West Greenlandic):
tusaa-nngit-su-usaar-tuaannar-sinnaa-nngi-vip-putit
'hear'-neg.-intrans.participle-'pretend'-'all the time'-'can'-neg.-
'really'-2nd.sng.indicative
Ýou simply cannot pretend not to be hearing all the time'

Syntactic Categories

Two types of syntactic categories:

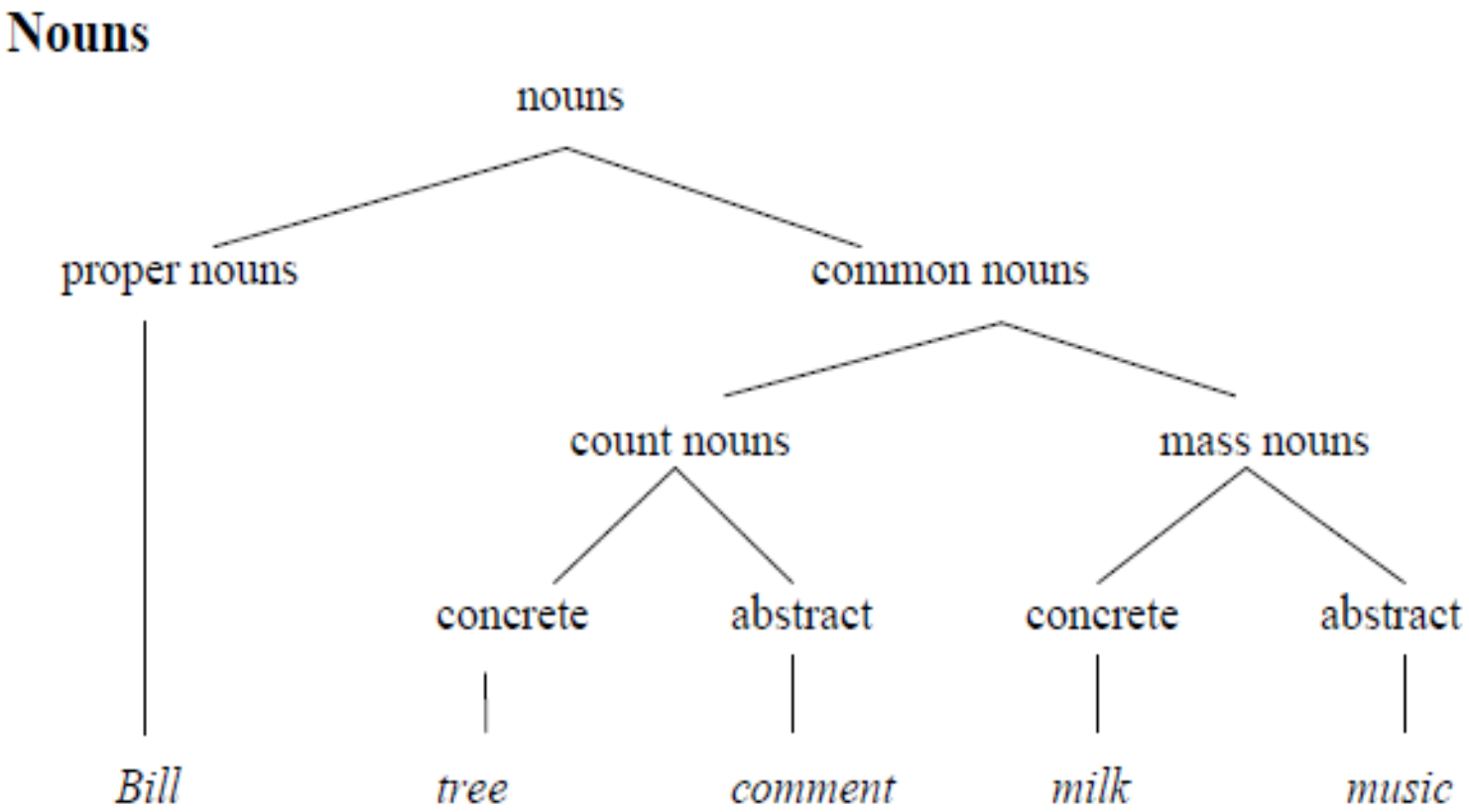
1. **Lexical:**
   - content words and receive **inflection**
   - open class : nouns,verbs, adjectives and adverbs
   - closed class: pronouns and number words

Syntactic Categories

Two types of syntactic categories:

1. **Lexical:**
   - content words and receive **inflection**
   - open class : nouns,verbs, adjectives and adverbs
   - closed class: pronouns and number words
2. **Functional:**
   - functional words do not receive inflection i.e. **indeclinables**
   - prepositions/postpositions, conjunctions, interjections, demonstratives, intensifiers, quotatives etc.

1. **Nouns** are words used to refer to people (boy), objects (backpack), creatures (dog), places (school), qualities (roughness), phenomena (earthquake) and abstract ideas (love) as if they were all "things."



- Proper names don't take a determiner
- Proper names and mass nouns don't pluralize
- Count nouns require a determiner
- Mass nouns do not take the indefinite determiner *a*

	Proper noun	Count noun	Mass noun
1.	*the Bill	The tree	The milk
2.	*three Bills	Three trees	*Three milks
3.	I met Bill	*I saw tree	I hate milk
4.	*a Bill	a tree	*a milk

Grammatical features of the noun:

Grammatical features of the noun:

1. number (boy-boys)

Grammatical features of the noun:

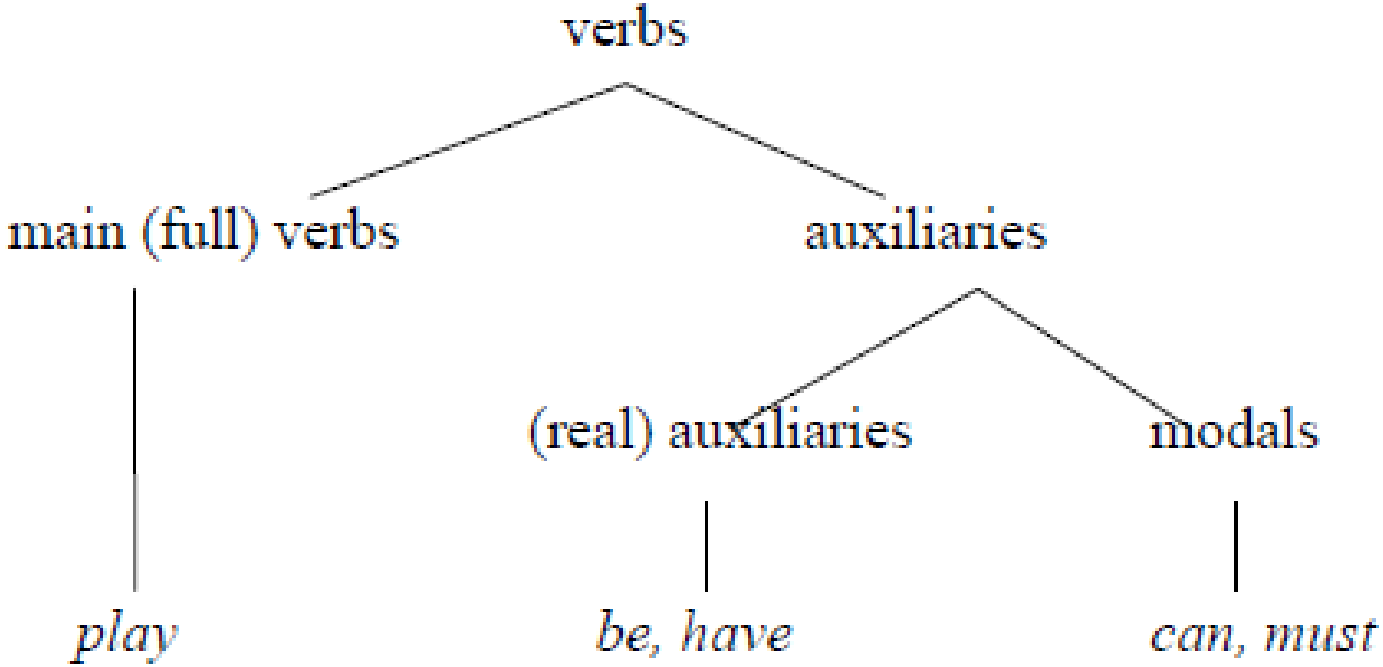
1. number (boy-boys)
2. gender (masculine (boy), feminine (girl), neuter(cat))

Grammatical features of the noun:

1. number (boy-boys)
2. gender (masculine (boy), feminine (girl), neuter(cat))
3. case (queen's)

2. **Verbs** are words used to refer to various kinds of actions (go, talk) and states (be, have) involving people and things in events.

Verbs



Grammatical features of the verb:

1. tense (walk vs. walked)

Grammatical features of the verb:

1. tense (walk vs. walked)
2. aspect (I have eaten)

Grammatical features of the verb:

1. tense (walk vs. walked)
2. aspect (I have eaten)
3. mood (come vs. don't come)

Grammatical features of the verb:

1. tense (walk vs. walked)
2. aspect (I have eaten)
3. mood (come vs. don't come)
4. agreement (gender-number-person)

Ta. *kutup ṣā kōlkoMṭā.v- aik kaṭṭ- iṉ- āṉ.*
Qutub Shah Golkonda- ACC build- PST- 3.SG.M.

Te. *kutub ṣā gōlkoMḍa- ni kāṭṭ- ā- ḍu*
Qutub Shah Golkonda- ACC build- PST- 3.SG.M.
'Qutub Shah built the Golkonda.'

3. Adjectives

Adjectives are words used, typically with nouns, to provide more information about the things referred to.

1. **attributive** : A blue ball

3. Adjectives

Adjectives are words used, typically with nouns, to provide more information about the things referred to.

1. **attributive** : A blue ball
2. **predicative** : The ball is blue.

4. Adverbs

Syntactically, adverbs function as modifiers of (i) clauses, (ii) verbs, (iii) adjectives.

3. Adjectives

Adjectives are words used, typically with nouns, to provide more information about the things referred to.

1. **attributive** : A blue ball
2. **predicative** : The ball is blue.

4. Adverbs

Syntactically, adverbs function as modifiers of (i) clauses, (ii) verbs, (iii) adjectives.

(1) **Fortunately**, we didn't miss the train.
(2) Peter walked **slowly** back to the car.
(3) That was **extremely** useful.

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)
2. Possessive pronouns (*my, our, his*)

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)
2. Possessive pronouns (*my, our, his*)
3. Reflexive pronouns (*myself, ourselves, himself*)

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)
2. Possessive pronouns (*my, our, his*)
3. Reflexive pronouns (*myself, ourselves, himself*)
4. Relative pronouns (*who, which, that*)

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)
2. Possessive pronouns (*my, our, his*)
3. Reflexive pronouns (*myself, ourselves, himself*)
4. Relative pronouns (*who, which, that*)
5. Interrogative pronouns (*who, whose, which one*)

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)
2. Possessive pronouns (*my, our, his*)
3. Reflexive pronouns (*myself, ourselves, himself*)
4. Relative pronouns (*who, which, that*)
5. Interrogative pronouns (*who, whose, which one*)
6. Demonstrative pronouns (*this, that, these*)

5. Pronouns

Pronouns are words (she, herself, they, it, you) used in place of noun phrases, typically referring to people and things already known.

1. Personal pronouns (subject: *I, we, he* ; object: *me, us, him*)
2. Possessive pronouns (*my, our, his*)
3. Reflexive pronouns (*myself, ourselves, himself*)
4. Relative pronouns (*who, which, that*)
5. Interrogative pronouns (*who, whose, which one*)
6. Demonstrative pronouns (*this, that, these*)
7. Indefinite pronouns (*everyone, all, none*)

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*
3. Cause: *because of, on account of, for*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*
3. Cause: *because of, on account of, for*
4. Instrument: *by, with*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*
3. Cause: *because of, on account of, for*
4. Instrument: *by, with*
5. Accompaniment: *with, without*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*
3. Cause: *because of, on account of, for*
4. Instrument: *by, with*
5. Accompaniment: *with, without*
6. Concession: *despite, in spite of*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*
3. Cause: *because of, on account of, for*
4. Instrument: *by, with*
5. Accompaniment: *with, without*
6. Concession: *despite, in spite of*
7. Exception: *except for, excluding, apart from*

6. Prepositions

Prepositions are words (at, in, on, near, with, without) used with nouns in phrases providing information about time (at five o'clock, in the morning), place (on the table, near the window) and other connections (with a knife, without a thought) involving actions and things.

1. Space: *in, on, at, above, under, through*
2. Time: *since, in, during, for, at, from .... to*
3. Cause: *because of, on account of, for*
4. Instrument: *by, with*
5. Accompaniment: *with, without*
6. Concession: *despite, in spite of*
7. Exception: *except for, excluding, apart from*
8. Addition: *in addition, besides, as well as*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*
3. Condition: *if, unless, in case that*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*
3. Condition: *if, unless, in case that*
4. Concession: *although*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*
3. Condition: *if, unless, in case that*
4. Concession: *although*
5. Result: *so that*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*
3. Condition: *if, unless, in case that*
4. Concession: *although*
5. Result: *so that*
6. Contrast: *but, whereas, while*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*
3. Condition: *if, unless, in case that*
4. Concession: *although*
5. Result: *so that*
6. Contrast: *but, whereas, while*
7. Purpose: *in order to*

7. Conjunctions

Conjunctions are words (and, but, because, when) used to make connections and indicate relationships between events.

E.g. `Sita's husband was so sweet and he helped her a lot because she couldn't do much when she was pregnant.`

1. Time: *when, since, after, before, until, as soon as*
2. Cause: *because, as, since*
3. Condition: *if, unless, in case that*
4. Concession: *although*
5. Result: *so that*
6. Contrast: *but, whereas, while*
7. Purpose: *in order to*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*
3. Demonstratives: *this, that*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*
3. Demonstratives: *this, that*
4. Possessives: *my, your, his, her, our, their*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*
3. Demonstratives: *this, that*
4. Possessives: *my, your, his, her, our, their*
5. WH determiner: *which, whose, whichever, whatever, whoever*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*
3. Demonstratives: *this, that*
4. Possessives: *my, your, his, her, our, their*
5. WH determiner: *which, whose, whichever, whatever, whoever*
6. Negative determiner: *no*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*
3. Demonstratives: *this, that*
4. Possessives: *my, your, his, her, our, their*
5. WH determiner: *which, whose, whichever, whatever, whoever*
6. Negative determiner: *no*
7. Quantifier: *each, either, neither, some, any, enough*

8. Determiners

Determiners are placed in front of a noun to specify quantity (e.g., one cat, many cats) or to clarify what the noun refers to (e.g., my cat, that cat, the cat)

S

1. Definite article: *the*
2. Indefinite article: *a, an*
3. Demonstratives: *this, that*
4. Possessives: *my, your, his, her, our, their*
5. WH determiner: *which, whose, whichever, whatever, whoever*
6. Negative determiner: *no*
7. Quantifier: *each, either, neither, some, any, enough*

Phrasal Categories

1. Noun Phrase

Functions as Subject, object, indirect object etc.,

- N = John
- DET N = the boy
- DET ADJ N= a little boy
- DET N PP= a boy in a bubble

Phrasal Categories

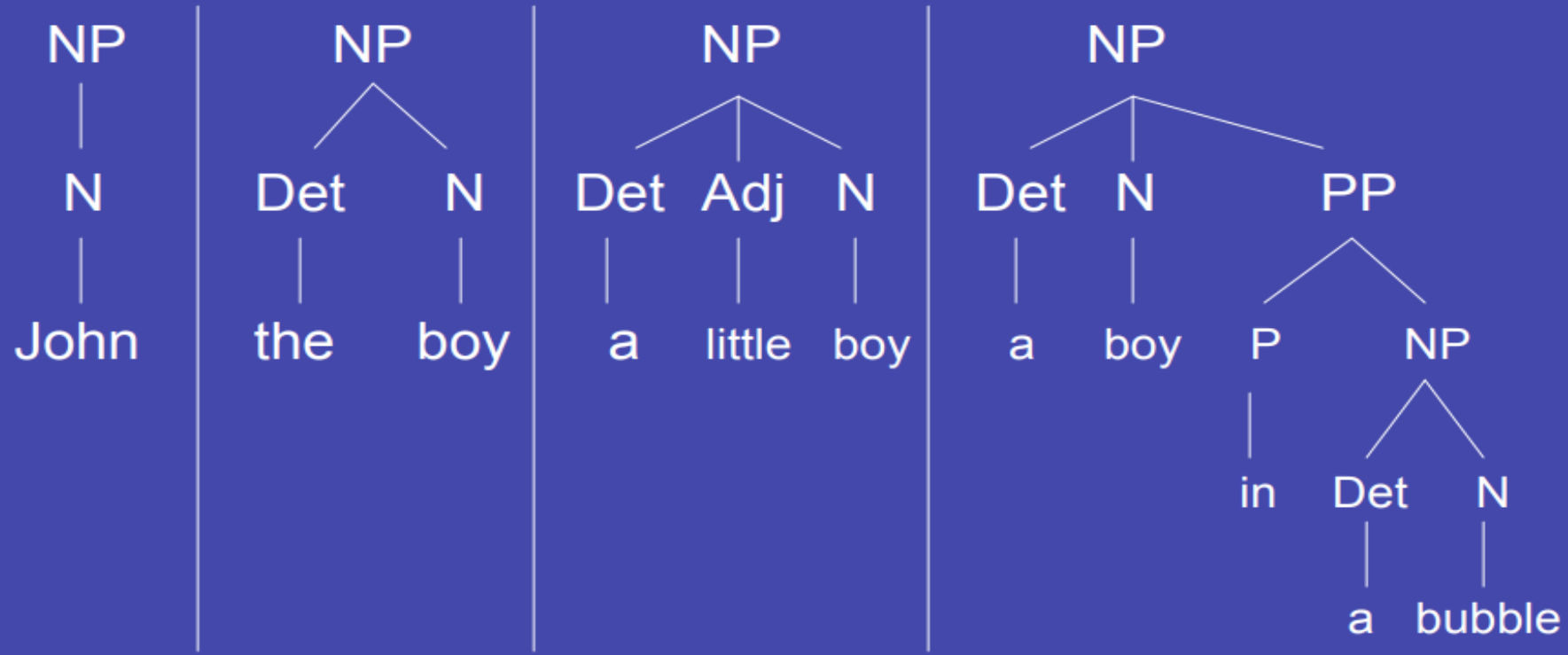
1. Noun Phrase

Functions as Subject, object, indirect object etc.,

- N = John
- DET N = the boy
- DET ADJ N= a little boy
- DET N PP= a boy in a bubble

Phrase structure rule for NPs:
NP = (Det) (Adj)* N (PP)

Phrase structure trees



2. Prepositional Phrase

PP has a preposition as its head and this is usually followed by a noun phrase .

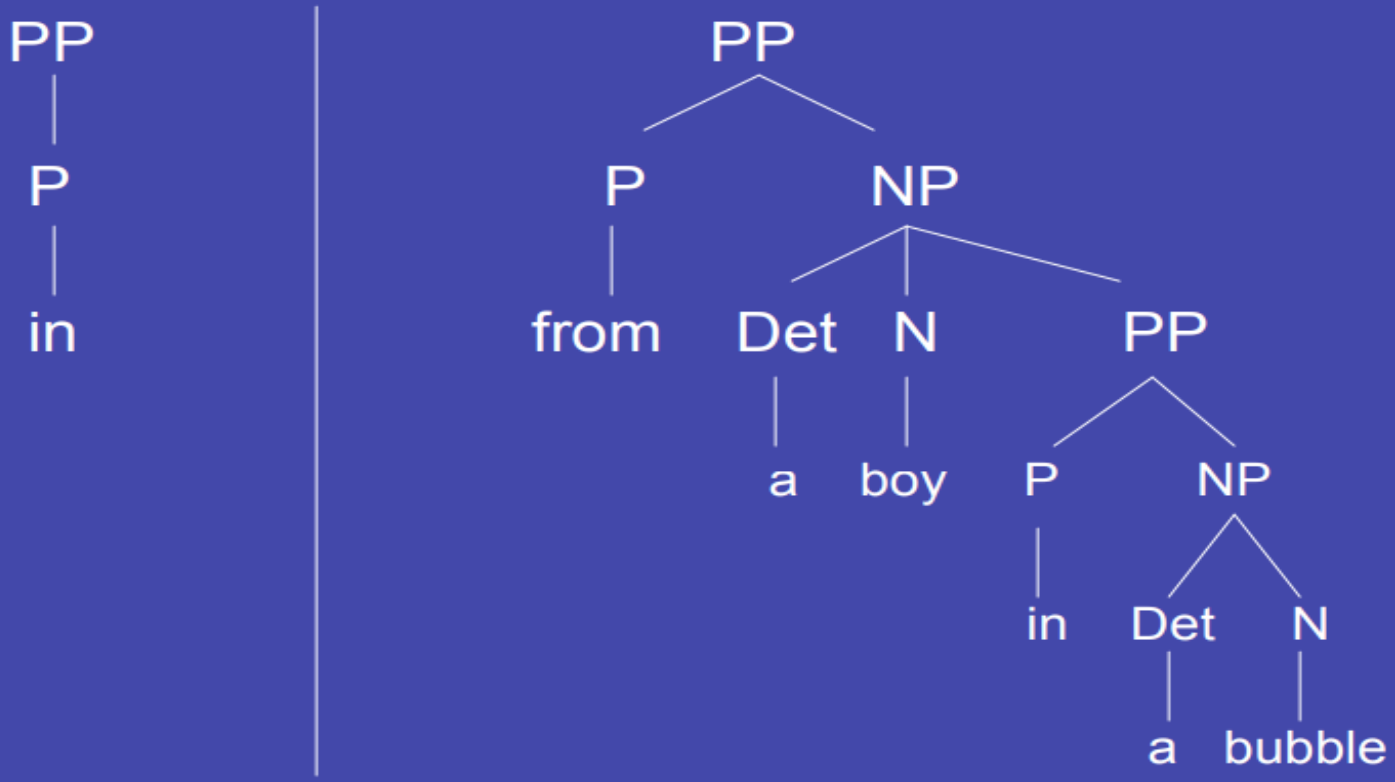
2. Prepositional Phrase

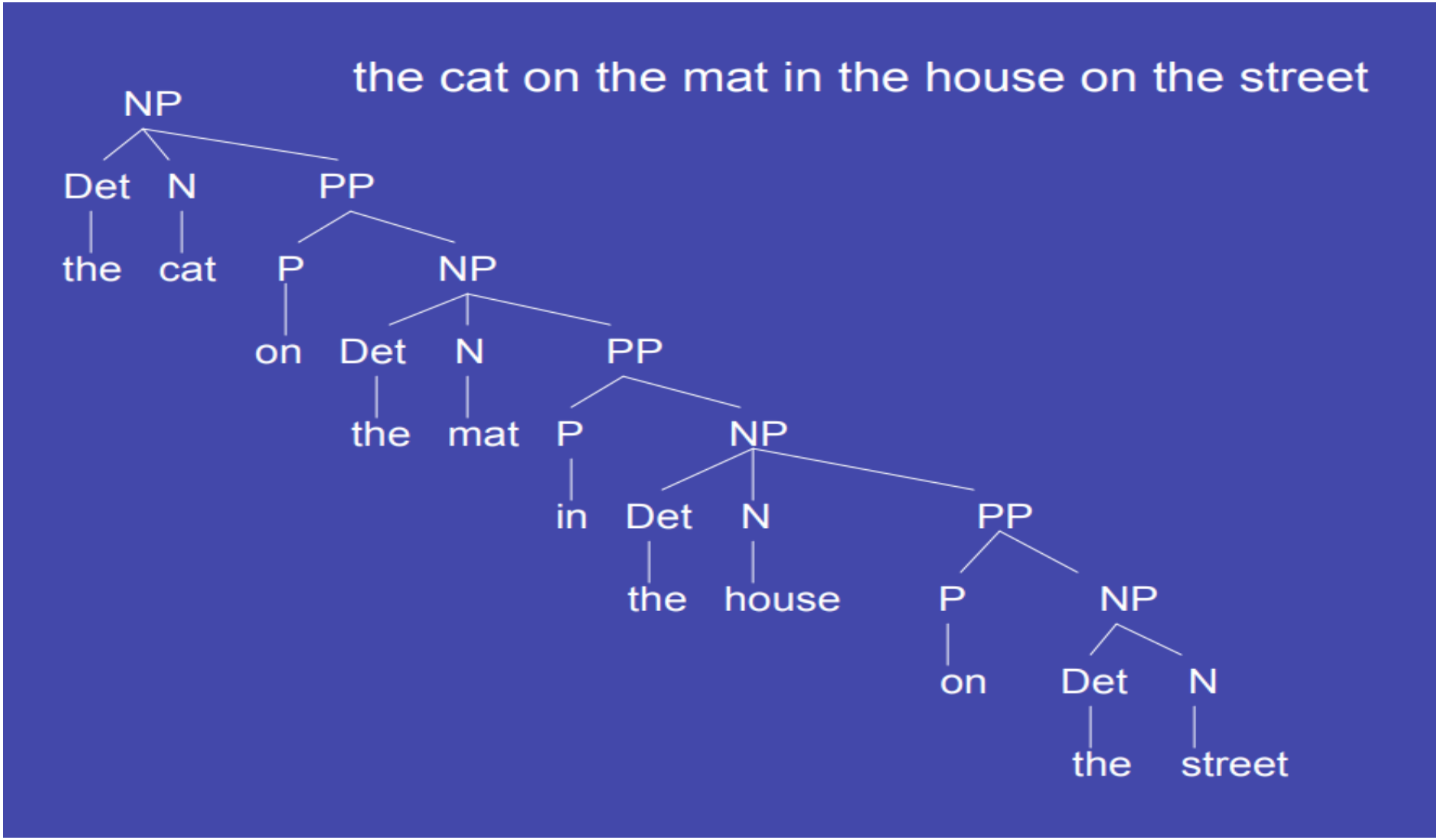
PP has a preposition as its head and this is usually followed by a noun phrase .

- P NP = With long hair

Phrase structure rule for PPs:
PP = P NP

Phrase structure trees




the cat on the mat in the house on the street
NP
Det
N
PP
the
cat
P
NP
on
Det
N
PP
the
mat
P
NP
in
Det
N
PP
the
house
P
NP
on
Det
N
the
street

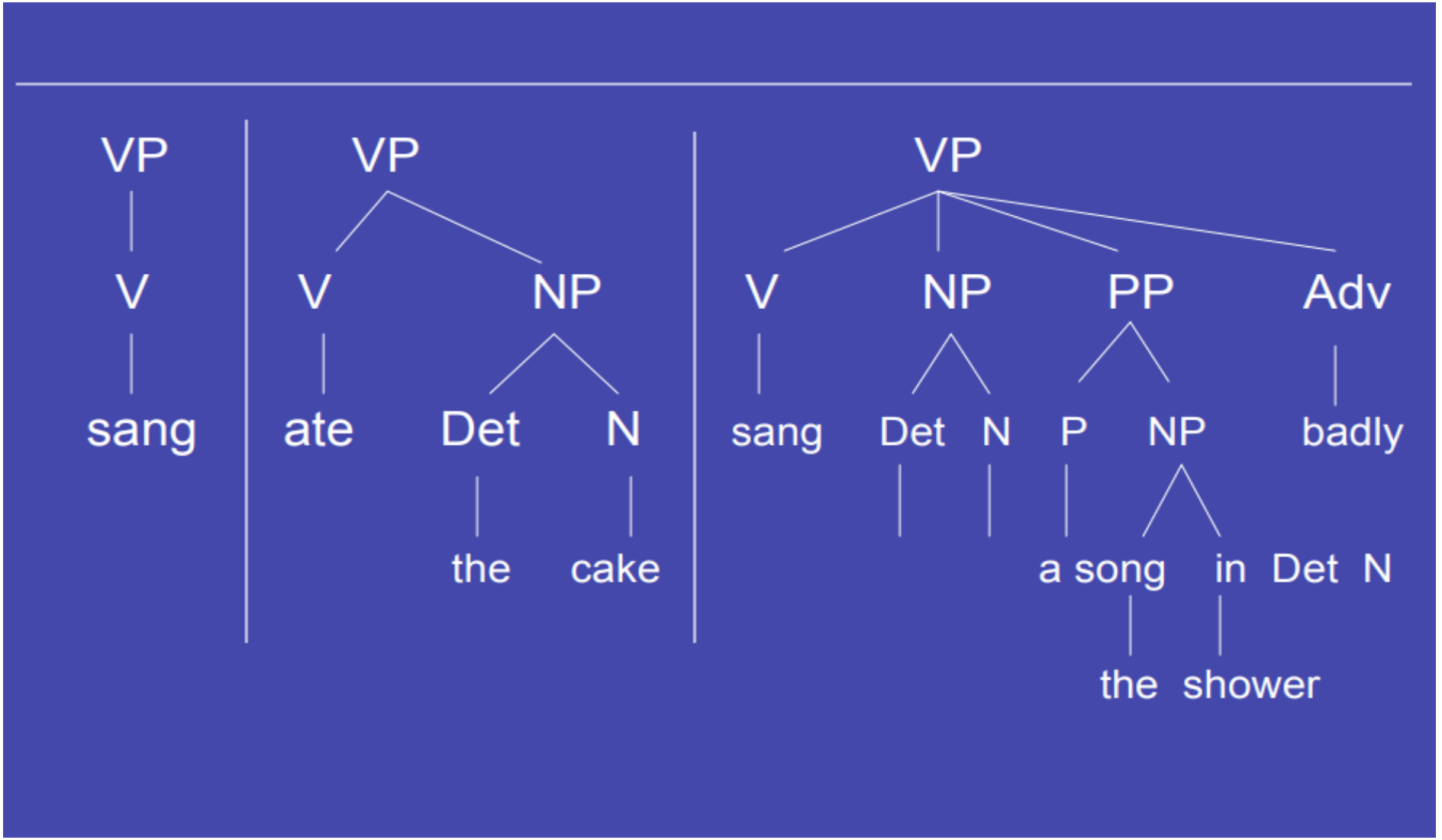
2. Verb phrase

- V = They **sleep, sit, work** (intransitive verbs)
- V NP = I **hit, kiss, see** him (transitive verbs)
- V NP NP = The girl **give, send, bring** a book to him (ditransitive verb)
- V NP Adv = Mary **ate a cake hungrily**
- V NP PP= John **sang a song in the shower**
- V PP Adv= I **fell into the pond slowly**

2. Verb phrase

- V = They **sleep, sit, work** (intransitive verbs)
- V NP = I **hit, kiss, see** him (transitive verbs)
- V NP NP = The girl **give, send, bring** a book to him (ditransitive verb)
- V NP Adv = Mary **ate a cake hungrily**
- V NP PP= John **sang a song in the shower**
- V PP Adv= I **fell into the pond slowly**

Phrase structure rule for VPs:
VP = V (NP)* (PP) (Adv)*


VP
V
sang
VP
V
NP
ate
Det
N
the
cake
VP
V
NP
PP
Adv
sang
Det
N
P
NP
badly
a song
in
Det
N
the
shower

Sentences

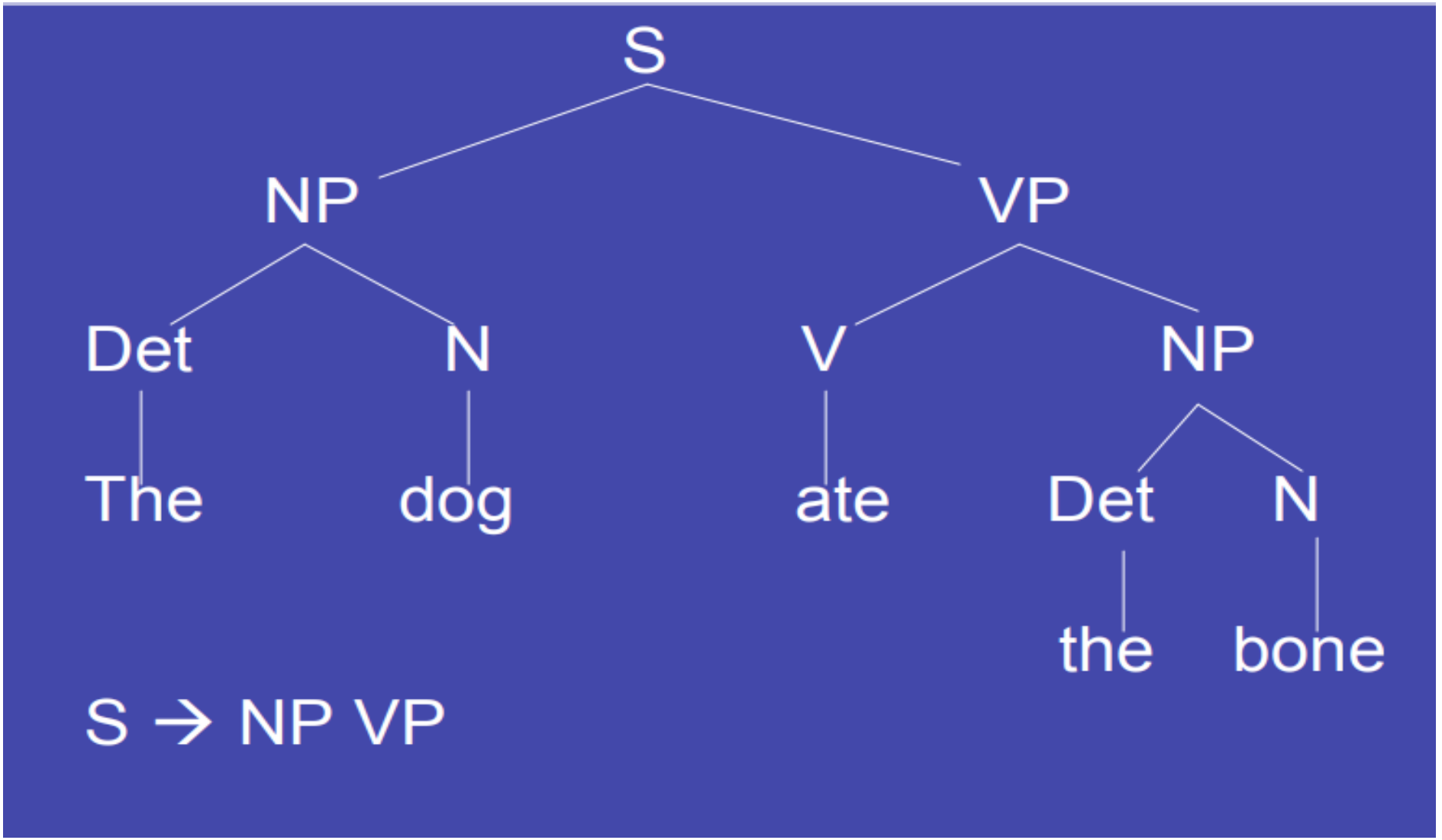
- Must contain an NP and a VP
- NP VP = The dog barked
- *The dog, *barked

Sentences

- Must contain an NP and a VP
- NP VP = The dog barked
- *The dog, *barked

Phrase structure rule for S:

S = NP VP


S
NP
VP
Det
N
V
NP
The
dog
ate
Det
N
the
bone
S → NP VP

Structure of Clauses

A clause is a group of related words containing a subject and a verb.

It is different from a phrase in that a phrase does not include a subject and a verb relationship.

- independent clause/ matrix clause
  - Contains a subject and a verb, expresses complete sense i.e. finite
  - e.g. I went to the market

- dependent clause/ subordinate clause
  - Contains a subject and a verb, does not express complete sense i.e. non-finite
  - e.g. **If I go out**, I usually take my umbrella

Independent Clauses

1. Coordinate clause

A coordinate clause is a clause belonging to a series of two or more clauses not syntactically dependent one on another

- He gets up early in the morning and goes out for a walk.
- He is poor but he is honest.
- You must work hard or you can't succeed.

Independent Clauses

2. Correlative Conjunctions

Correlative conjunctions are pairs of conjunctions which work together to coordinate two items. They always appear in pairs.

Independent Clauses

2. Correlative Conjunctions

Correlative conjunctions are pairs of conjunctions which work together to coordinate two items. They always appear in pairs.

- either/or - I want either the cheesecake or the chocolate cake.
- both/and - We'll have both the cheesecake and the chocolate cake.
- whether/or - I didn't know whether you'd want the cheesecake or the chocolate cake, so I got both.
- neither/nor - Oh, you want neither the cheesecake nor the chocolate cake? No problem.
- not only/but also - I'll eat them both - not only the cheesecake but also the chocolate cake.
- not/but - I see you're in the mood not for desserts but appetizers. I'll help you with those, too.

Independent Clauses

3. Complement clause:
A complement clause is a notional sentence or predication that is an argument of a predicate.

Independent Clauses

3. Complement clause:
A complement clause is a notional sentence or predication that is an argument of a predicate.

E.g.

- I heard the evidence that he did it.
- I am sure that he did it.
- I am not certain what we did.

Independent Clauses

3. Complement clause:
A complement clause is a notional sentence or predication that is an argument of a predicate.

E.g.

- I heard the evidence that he did it.
- I am sure that he did it.
- I am not certain what we did.

Relative clauses are not complement clauses.
Relative clauses modify a noun phrase, whereas complement clauses are arguments which are selected by a verb, noun, or adjective.

Dependent Clauses

1. Relative Clause
   - A relative clause begins with a relative pronoun and functions as an adjective.
   - introduced by a relative pronoun, *who, whom, whose, which, that*
   - E.g. I want a book in a language which I can read.

Dependent Clauses

2. Adverbial clause
   - introduced by a subordinating conjunction: because, as, if, since, when, while, whereas, although, etc.
   - typically function as adverbials
   - E.g. I'll tell you about that after we are married. (time adverbial)
   - E.g. Since they could see me, their presence in the hall disturbed me. (adverbial of reason).
   - E.g. She could do the work herself if it was necessary. (adverbial of condition).

Sentence Composition Types

1. Simple Sentences
   - one independent clause
   - E.g. I like Mysore

Sentence Composition Types

1. Simple Sentences
   - one independent clause
   - E.g. I like Mysore
2. Compound Sentences
   - two or more coordinate clause
   - E.g. I like Mysore and he likes CIIL

Sentence Composition Types

1. Simple Sentences
   - one independent clause
   - E.g. I like Mysore
2. Compound Sentences
   - two or more coordinate clause
   - E.g. I like Mysore and he likes CIIL
3. Complex Sentences
   - at least one independent clause, and at least one dependent clause
   - E.g. My Dad laughed when I told a joke.

Sentence Composition Types

1. Simple Sentences
   - one independent clause
   - E.g. I like Mysore
2. Compound Sentences
   - two or more coordinate clause
   - E.g. I like Mysore and he likes CIIL
3. Complex Sentences
   - at least one independent clause, and at least one dependent clause
   - E.g. My Dad laughed when I told a joke.
4. Compound-Complex Sentences
   - at least two independent clauses and one or more dependent clauses
   - E.g. Ponny forgot her friend's birthday, so she sent her a card when she finally remembered.

Complexities in Syntax

1. Ambiguity

- (i) Attachment ambiguity: (Structural Ambiguity)

- It occurs when the grammar can assign more than one parse to a sentence.
- e.g. `I saw a man with a binocular`
- It is ambiguous because the phrase 'with a binocular' can be part of the NP headed by 'binocular' or a part of the VP headed by 'saw'

Complexities in Syntax

1. Ambiguity

- (ii) Coordination ambiguity

- In coordination ambiguity, different sets of phrases can be conjoined by a conjunction like 'and'.
- the phrase `old men and women` can be bracketed as `[old [men and women]]` , referring to old men and old women , or as [old men] and [women] , in which case it is only the men who are old.
- the men and women or boys
- Depending upon how the scope of the conjunctions is interpreted, this phrase can mean 'either men and women, or men and boys', or 'either men and women, or boys'.

Complexities in Syntax

2. Garden-Path Sentences

- The government plans to raise taxes were defeated.
  first interpretation : The government is planning to raise taxes...
  final interpretation : The plans of the government to raise taxes were defeated.

- The old man the boat
  first interpretation : The man, who is old...
  final interpretation : The boat is manned by the old.

- The horse raced past the barn
  .....fell.

3. Recursiveness

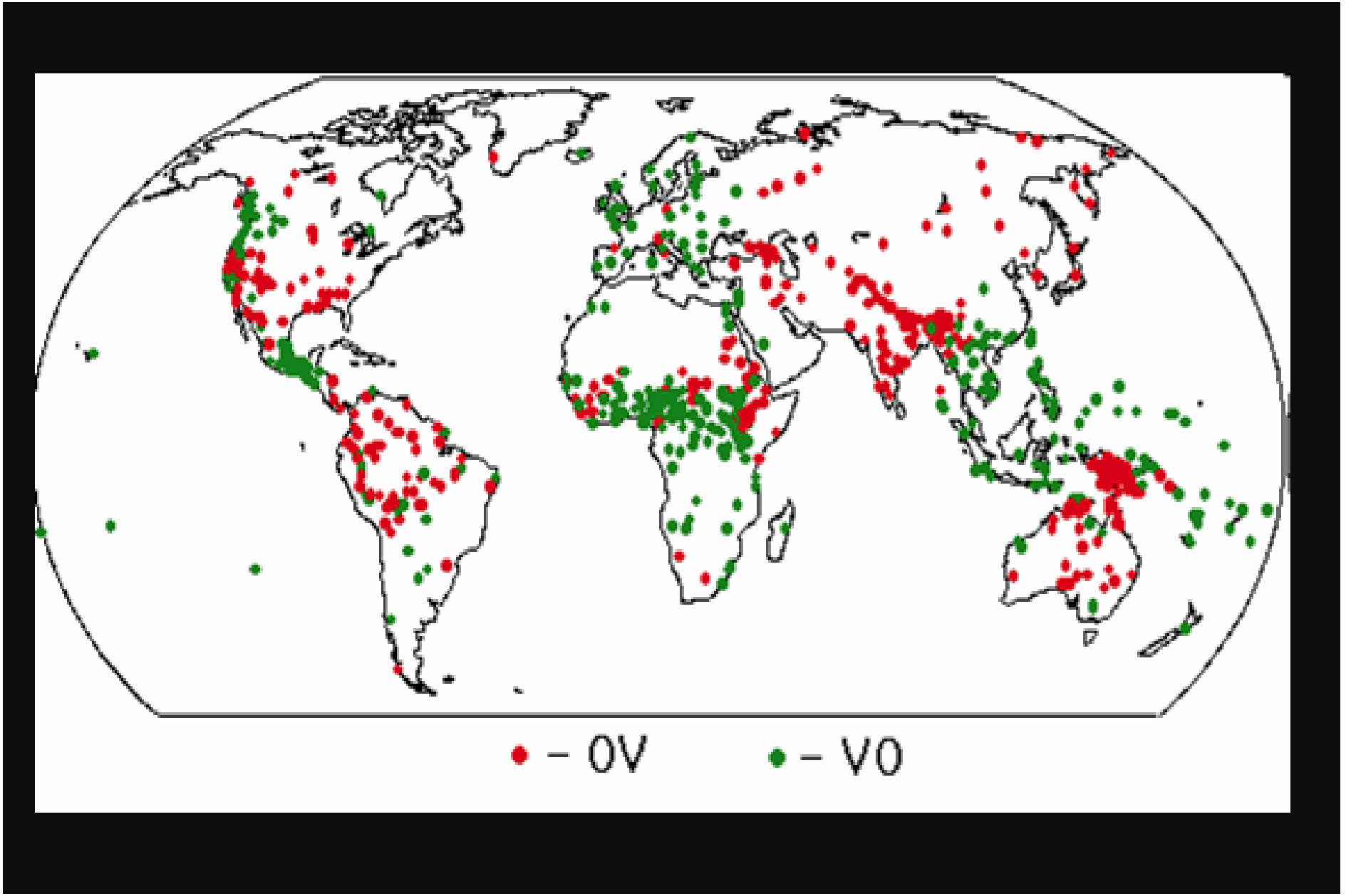
- This is the cat that bit the rat that ate the cheese that Jack bought
- They came for dinner and they stayed all night and they left right after breakfast but they were still late and were punished when they got to school

4. Ellipsis

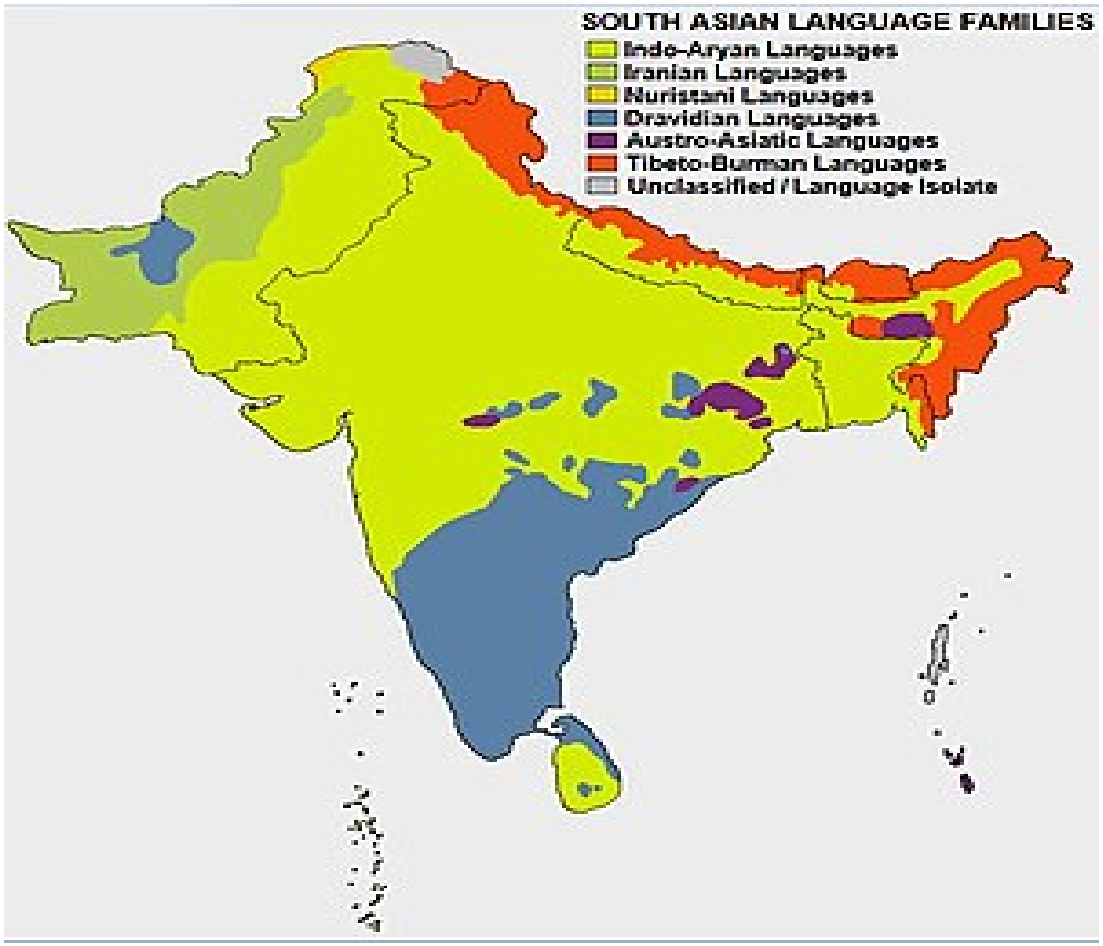
- She dances better than he does
- I did it; he didn't.
- Elliptical reference between sentences frequently relies on structural symmetry between the two related expressions.

Word Order Distribution of Languages		
Basic Word Order	Proportion of Languages	Examples
Subject-[Verb-Object]	42%	English, Indonesian
Subject-[Object-Verb]	45%	Japanese, Turkish
Verb-Subject-Object	9%	Welsh, Zapotec
[Verb-Object]-Subject	3%	Malagasy
[Object-Verb]-Subject	1%	
Object-Subject-Verb	0%	

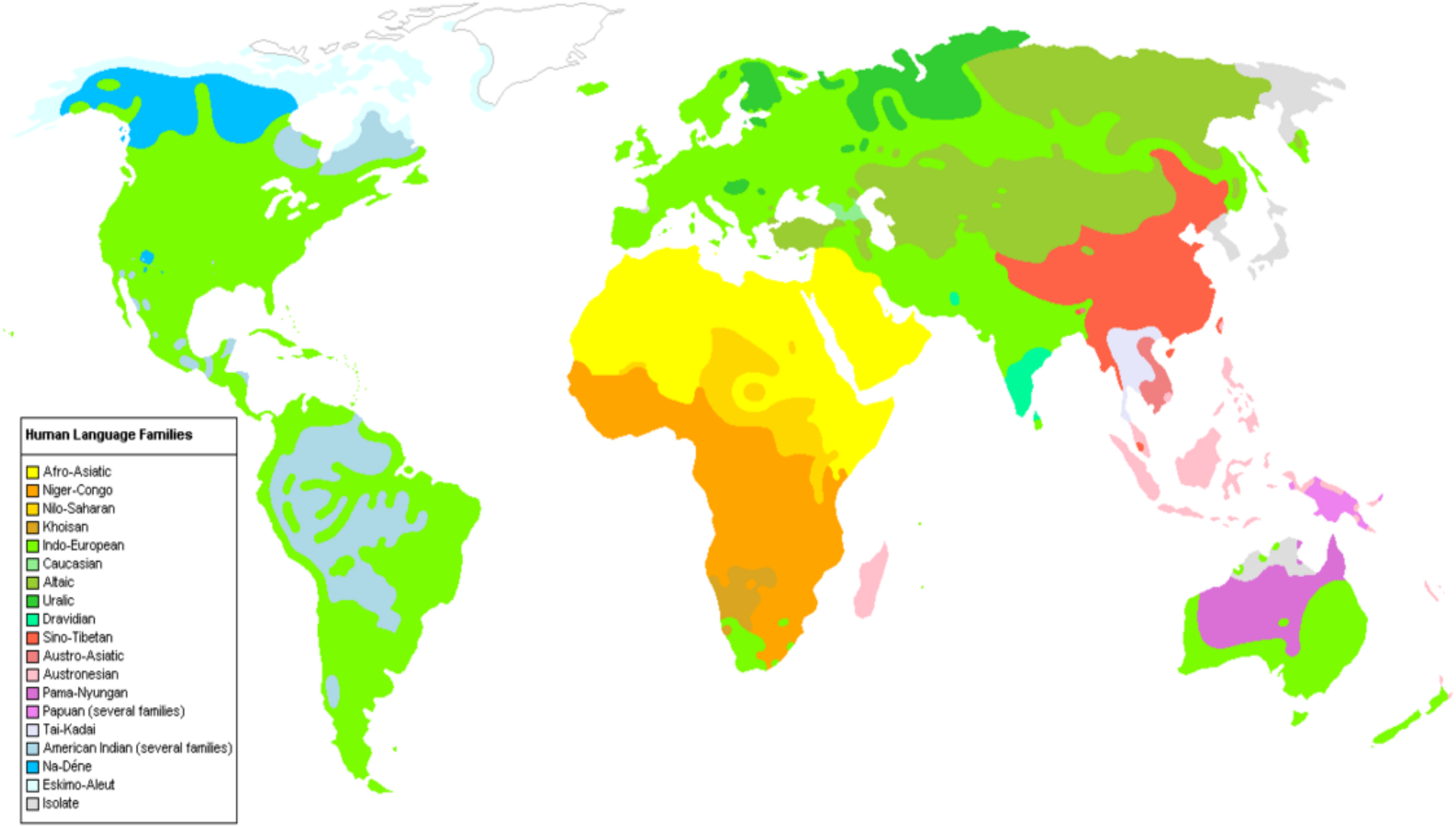
Russell Tomlin, *Basic Word Order: Functional Principles*, (Croom Helm, London, 1986) page 22



extracted from https://www.languagesoftheworld.info



extracted from https://www.wikipedia.com



extracted from https://www.wikipedia.com