POS

PRP
VBZ
NNS
IN
DT
NN
She
sells
seashells
on
the
seashore
Part Of Speech Tagging

Hannanum	Kkma	Komoran	Mecab	Twitter
아버지가방에 들어가 / N	아버지 / NNG	아버지가방에 들어가신다 / NNP	아버지 / NNG	아버지 / Noun
이 / J	가방 / NNG		가 / JKS	가방 / Noun
시ㄴ다 / E	에 / JKM		방 / NNG	에 / Josa
	들어가 / VV		에 / JKB	들어가신 / Verb
	시 / EPH		들어가 / VV	다 / Eomi
	ㄴ다 / EFN		신다 / EP+EC	

Hannanum	Kkma	Komoran	Mecab	Twitter
하늘 / N	하늘 / NNG	하늘 / NNG	하늘 / NNG	하늘 / Noun
을 / J	을 / JKO	을 / JKO	을 / JKO	을 / Josa
나 / N	날 / VV	나 / NP	나 / NP	나 / Noun
는 / J	는 / ETD	는 / JX	는 / JX	는 / Josa
자동차 / N	자동차 / NNG	자동차 / NNG	자동차 / NNG	자동차 / Noun

Collocations

Collocations

WHAT A sequence of words or terms

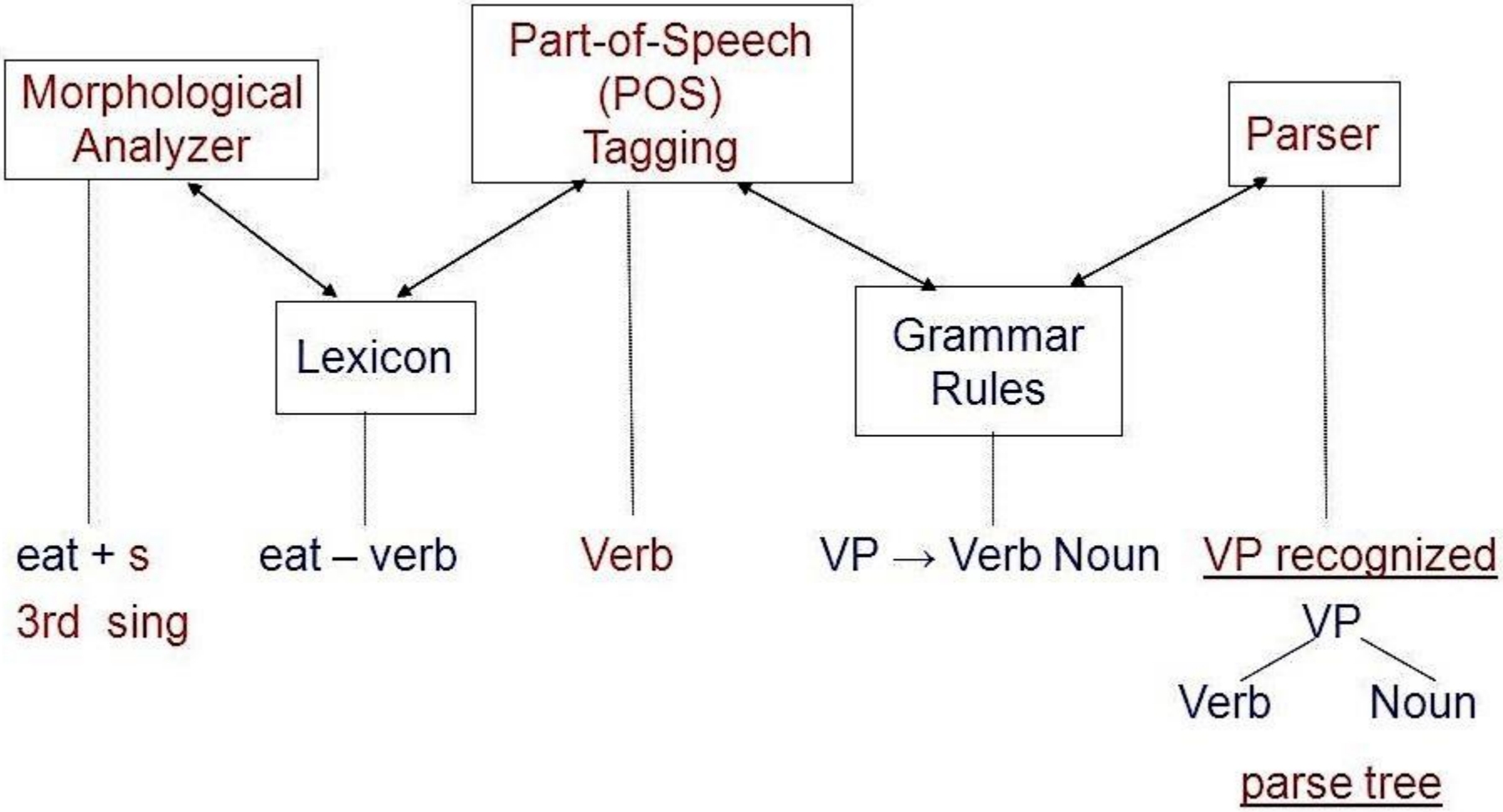
co-occur more often than would be expected by chance

frequency counting, ***Pointwise Mutual Information****(PMI), hypothesis testing(t-test and* ***chi-square****), …*

Collocations	
Correct	**Incorrect**
• High temperature • Have an experience • Heavy rain	• Tall temperature • Make an experience • Thick rain

Phrase

- Longer phrases may be used rather than a single word, but fulfilling the same role in a sentence.
  - ***Noun phrases*** refer to objects: *four fried chickens*.
  - ***Verb phrases*** state what the noun phrase does: *kicks the dog*.
  - ***Adjective phrases*** describe/qualify an object: *sickly sweet*.
  - ***Adverbial phrases*** describe how it is done: *very carefully*.
  - ***prepositional phrases***: add information to a verb phrase: *on the table*


Morphological Analyzer
Part-of-Speech (POS) Tagging
Parser
Lexicon
Grammar Rules
eat + s
3rd sing
eat – verb
Verb
VP → Verb Noun
VP recognized
VP
Verb
Noun
parse tree

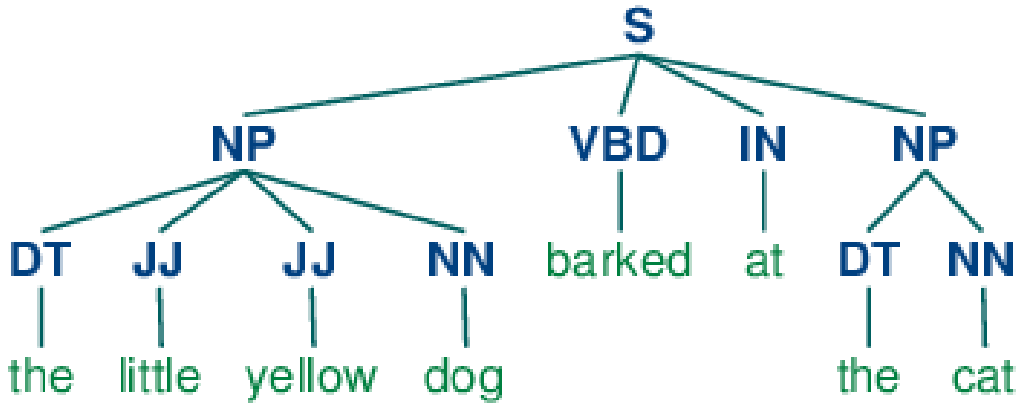
class `nltk.chunk.regexp.`**`RegexpParser`**(*grammar*, *root_label='S'*, *loop=1*, *trace=0*) [source]

Bases: `nltk.chunk.api.ChunkParserI`

A grammar based chunk parser. `chunk.RegexpParser` uses a set of regular expression patterns to specify the behavior of the parser. The chunking of the text is encoded using a `ChunkString`, and each rule acts by modifying the chunking in the `ChunkString`. The rules are all implemented using regular expression matching and substitution.

A grammar contains one or more clauses in the following form:

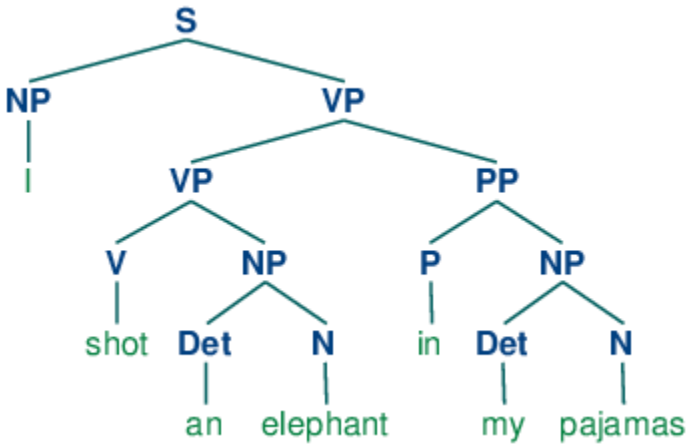
```
NP:
  {<DT|JJ>}          # chunk determiners and adjectives
  }<[\.VI].*>+{      # chink any tag beginning with V, I, or .
  <.*>}{<DT>         # split a chunk at a determiner
  <DT|JJ>{}<NN.*>    # merge chunk ending with det/adj
                     # with one starting with a noun
```

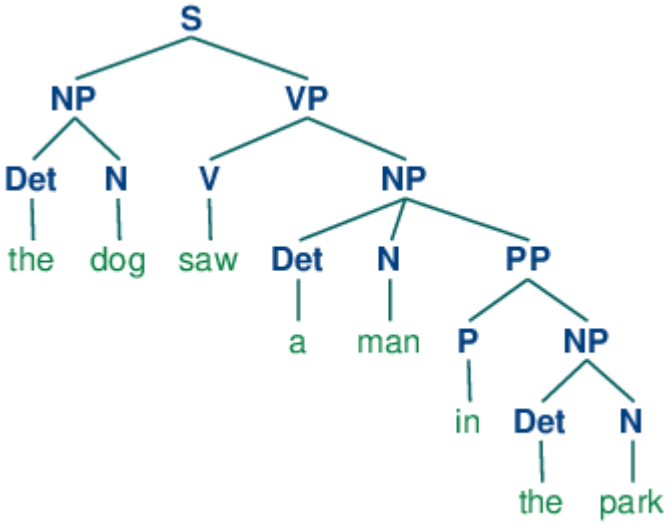


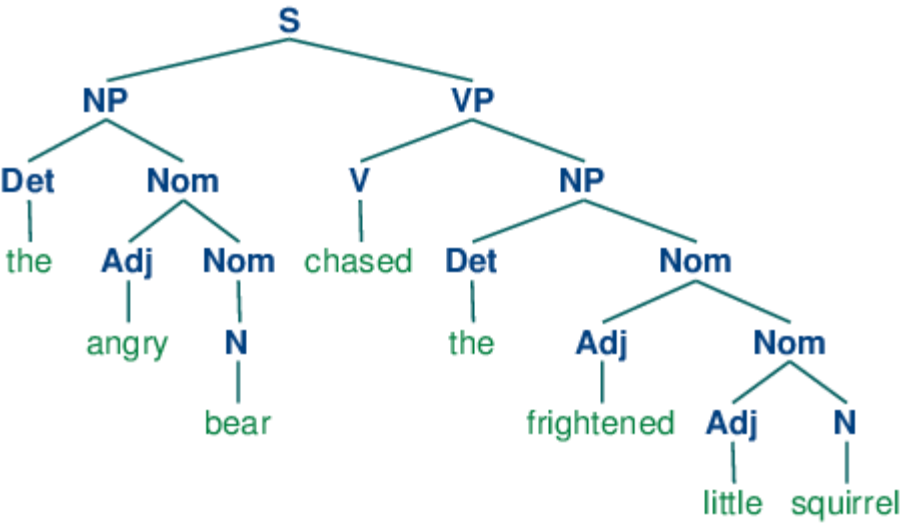
내 친구가 잠을 많이 잔다.

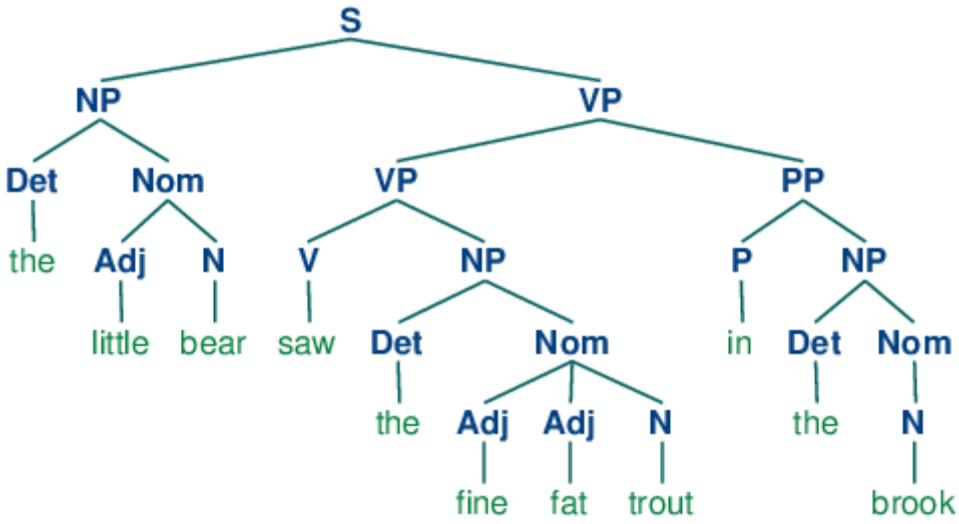
명사구 동사구

Exercises


S
NP
VP
I
VP
PP
V
NP
P
NP
shot
Det
N
in
Det
N
an
elephant
my
pajamas


S
NP
VP
Det
N
V
NP
the
dog
saw
Det
N
PP
a
man
P
NP
in
Det
N
the
park


S
NP
VP
Det
Nom
V
NP
the
Adj
Nom
chased
Det
Nom
angry
N
the
Adj
Nom
bear
frightened
Adj
N
little
squirrel


S
NP
VP
Det
Nom
VP
PP
the
Adj
N
V
NP
P
NP
little
bear
saw
Det
Nom
in
Det
Nom
the
Adj
Adj
N
the
N
fine
fat
trout
brook