

(1)

Lexicon Lookup

Input (word, PennPOS)

Direct
lexicon lookup within
entries of PennPOS

Success

Failure

Return lemma

Hierarchical
lexicon lookup within
entries of PennPOS

Success

Failure

Return lemma

Converting PennPOS to NUPOS

Direct
lexicon lookup within
entries of NUPOS

Success

Failure

Return lemma

Hierarchical
lexicon lookup within
entries of NUPOS

Success

Failure

Return lemma

Input (word, NUPOS)

Direct
lexicon lookup within
entries of NUPOS

Success

Failure

Return lemma

Hierarchical
lexicon lookup within
entries of NUPOS

Success

Failure

Return lemma

Input (word)

Lexicon lookup within
entries of all tagsets

Failure

Success

Return lemmas
of all possible
Part-of-Speech

Lexicon

Resource

(2)

Rule-based Lemmatization

Convert NUPOS
to word class

Success

Failure

Should the word class
be lemmatized ?

Yes

No

Return word

Is the word a
compound ?

No

Yes

Is word class available ?

Yes

No

Apply rules of
the word class

Apply all rules

Lexicon
validation

Return lemma

Lexicon
validation

Extract
individual words

Detachment
Rules

Resource

Should the word be
lemmatized ?

Yes

No

Return word