

POS NOTATION MAPPING

(BETWEEN CLAWS2 USED BY RASP & PENN USED BY STANFORD ON DEPENDENCY LEVEL)

Amar Pratap Singh

amarsngh12@gmail.com

Banaras Hindu University

Abstract

At present Anusarka is using PENN, when once POS mapping is done between PENN(Stanford parser)and CLAWS2(rasp) than RASP will also come under same use.The present paper explores the problems during POS notation mappings and suggests how best can be done. Part-of-Speech (POS) tagging is the process of assigning the appropriate part of speech or lexical category to each word in a natural language sentence.This work ,I carried at IIIT Hyderabad.

Steps:-

- A) To downlowde both the parser RASP & STANFORD.
- B) Operate it on Linux system.
- C) Read the manuals of both the parsers(to understand parser's operation well)
- D) Uderstand the POS notation of PENN & CLAWS2,5,7 well.
- E) Make a definite choice among CLAWS 2,5,7 on the basis of RASP parser's output.
- F) Give the sentence input to receive parsed sentence (same sentence needs to be given in both the parser).
- G) Now start Mapping

LEXICAL ITEM	POS CLASS	CLAWS2	PENN
Is/sleeps	Verb	VBZ/VVZ	VBZ
Am/are/sleep	do	VBM/VBR/VVD	VBP
Was/were/slept	do	VBDZ/VBDR/VVD	VBD
Sleeping	do	VVG	VBG
Can/would/will/shall	modal	VM	MD
I,we,you,he, she ,it,they	Pronoun	PPIS1,PPIS2,PPY,PPHS,PPHS1,PPHS2	PRP
My,his,her,yours	Possessive,pre-nominal	APPS	PRP\$