

Appendix 2 CLAW7 and Penn Treebank tagsets

Appendix 2.1 UCREL CLAW7 Tagset

APPGE	possessive pronoun, pre-nominal (e.g. <i>my, your, our</i>)
AT	article (e.g. <i>the, no</i>)
AT1	singular article (e.g. <i>a, an, every</i>)
BCL	before-clause marker (e.g. <i>in order (that), in order (to)</i>)
CC	coordinating conjunction (e.g. <i>and, or</i>)
CCB	adversative coordinating conjunction (<i>but</i>)
CS	subordinating conjunction (e.g. <i>if, because, unless, so, for</i>)
CSA	<i>as</i> (as conjunction)
CSN	<i>than</i> (as conjunction)
CST	<i>that</i> (as conjunction)
CSW	<i>whether</i> (as conjunction)
DA	after-determiner or post-determiner capable of pronominal function (e.g. <i>such, former, same</i>)
DA1	singular after-determiner (e.g. <i>little, much</i>)
DA2	plural after-determiner (e.g. <i>few, several, many</i>)
DAR	comparative after-determiner (e.g. <i>more, less, fewer</i>)
DAT	superlative after-determiner (e.g. <i>most, least, fewest</i>)
DB	before-determiner or pre-determiner capable of pronominal function (<i>all, half</i>)
DB2	plural before-determiner (<i>both</i>)
DD	determiner (capable of pronominal function) (e.g. <i>any, some</i>)
DD1	singular determiner (e.g. <i>this, that, another</i>)
DD2	plural determiner (<i>these, those</i>)
DDQ	<i>wh</i> -determiner (<i>which, what</i>)
DDQGE	<i>wh</i> -determiner, genitive (<i>whose</i>)
DDQV	<i>wh</i> -ever determiner, (<i>whichever, whatever</i>)
EX	existential <i>there</i>
FO	formula
FU	unclassified word
FW	foreign word
GE	germanic genitive marker - (' or 's)
IF	<i>for</i> (as preposition)
II	general preposition
IO	<i>of</i> (as preposition)

IW	<i>with, without</i> (as prepositions)
JJ	general adjective
JJR	general comparative adjective (e.g. <i>older, better, stronger</i>)
JJT	general superlative adjective (e.g. <i>oldest, best, strongest</i>)
JK	catenative adjective (<i>able in be able to, willing in be willing to</i>)
MC	cardinal number, neutral for number (<i>two, three..</i>)
MC1	singular cardinal number (<i>one</i>)
MC2	plural cardinal number (e.g. <i>sixes, sevens</i>)
MCGE	genitive cardinal number, neutral for number (<i>two's, 100's</i>)
MCMC	hyphenated number (<i>40-50, 1770-1827</i>)
MD	ordinal number (e.g. <i>first, second, next, last</i>)
MF	fraction, neutral for number (e.g. <i>quarters, two-thirds</i>)
ND1	singular noun of direction (e.g. <i>north, southeast</i>)
NN	common noun, neutral for number (e.g. <i>sheep, cod, headquarters</i>)
NN1	singular common noun (e.g. <i>book, girl</i>)
NN2	plural common noun (e.g. <i>books, girls</i>)
NNA	following noun of title (e.g. <i>M.A.</i>)
NNB	preceding noun of title (e.g. <i>Mr., Prof.</i>)
NNL1	singular locative noun (e.g. <i>Island, Street</i>)
NNL2	plural locative noun (e.g. <i>Islands, Streets</i>)
NNO	numeral noun, neutral for number (e.g. <i>dozen, hundred</i>)
NNO2	numeral noun, plural (e.g. <i>hundreds, thousands</i>)
NNT1	temporal noun, singular (e.g. <i>day, week, year</i>)
NNT2	temporal noun, plural (e.g. <i>days, weeks, years</i>)
NNU	unit of measurement, neutral for number (e.g. <i>in, cc</i>)
NNU1	singular unit of measurement (e.g. <i>inch, centimetre</i>)
NNU2	plural unit of measurement (e.g. <i>ins., feet</i>)
NP	proper noun, neutral for number (e.g. <i>IBM, Andes</i>)
NP1	singular proper noun (e.g. <i>London, Jane, Frederick</i>)
NP2	plural proper noun (e.g. <i>Browns, Reagans, Koreas</i>)
NPD1	singular weekday noun (e.g. <i>Sunday</i>)
NPD2	plural weekday noun (e.g. <i>Sundays</i>)
NPM1	singular month noun (e.g. <i>October</i>)
NPM2	plural month noun (e.g. <i>Octobers</i>)
PN	indefinite pronoun, neutral for number (none)

PN1	indefinite pronoun, singular (e.g. <i>anyone, everything, nobody, one</i>)
PNQO	objective wh-pronoun (<i>whom</i>)
PNQS	subjective wh-pronoun (<i>who</i>)
PNQV	wh-ever pronoun (<i>whoever</i>)
PNX1	reflexive indefinite pronoun (<i>oneself</i>)
PPGE	nominal possessive personal pronoun (e.g. <i>mine, yours</i>)
PPH1	3rd person sing. neuter personal pronoun (<i>it</i>)
PPHO1	3rd person sing. objective personal pronoun (<i>him, her</i>)
PPHO2	3rd person plural objective personal pronoun (<i>them</i>)
PPHS1	3rd person sing. subjective personal pronoun (<i>he, she</i>)
PPHS2	3rd person plural subjective personal pronoun (<i>they</i>)
PPIO1	1st person sing. objective personal pronoun (<i>me</i>)
PPIO2	1st person plural objective personal pronoun (<i>us</i>)
PPIS1	1st person sing. subjective personal pronoun (<i>I</i>)
PPIS2	1st person plural subjective personal pronoun (<i>we</i>)
PPX1	singular reflexive personal pronoun (e.g. <i>yourself, itself</i>)
PPX2	plural reflexive personal pronoun (e.g. <i>yourselves, themselves</i>)
PPY	2nd person personal pronoun (<i>you</i>)
RA	adverb, after nominal head (e.g. <i>else, galore</i>)
REX	adverb introducing appositional constructions (<i>namely, e.g.</i>)
RG	degree adverb (<i>very, so, too</i>)
RGQ	wh- degree adverb (<i>how</i>)
RGQV	wh-ever degree adverb (<i>however</i>)
RGR	comparative degree adverb (<i>more, less</i>)
RGT	superlative degree adverb (<i>most, least</i>)
RL	locative adverb (e.g. <i>alongside, forward</i>)
RP	prep. adverb, particle (e.g. <i>about, in</i>)
RPK	prep. adv., catenative (<i>about in be about to</i>)
RR	general adverb
RRQ	wh- general adverb (<i>where, when, why, how</i>)
RRQV	wh-ever general adverb (<i>wherever, whenever</i>)
RRR	comparative general adverb (e.g. <i>better, longer</i>)
RRT	superlative general adverb (e.g. <i>best, longest</i>)
RT	quasi-nominal adverb of time (e.g. <i>now, tomorrow</i>)
TO	infinitive marker (<i>to</i>)

UH	interjection (e.g. <i>oh, yes, um</i>)
VB0	<i>be</i> , base form (finite i.e. imperative, subjunctive)
VBDR	<i>were</i>
VBDZ	<i>was</i>
VBG	<i>being</i>
VBI	<i>be</i> , infinitive (To be or not... It will be ..)
VBM	<i>am</i>
VCN	<i>been</i>
VBR	<i>are</i>
VBZ	<i>is</i>
VD0	<i>do</i> , base form (finite)
VDD	<i>did</i>
VDG	<i>doing</i>
VDI	<i>do</i> , infinitive (<i>I may do... To do...</i>)
VDN	<i>done</i>
VDZ	<i>does</i>
VH0	<i>have</i> , base form (finite)
VHD	<i>had</i> (past tense)
VHG	<i>having</i>
VHI	<i>have</i> , infinitive
VHN	<i>had</i> (past participle)
VHZ	<i>has</i>
VM	modal auxiliary (<i>can, will, would, etc.</i>)
VMK	modal catenative (<i>ought, used</i>)
VV0	base form of lexical verb (e.g. <i>give, work</i>)
VVD	past tense of lexical verb (e.g. <i>gave, worked</i>)
VVG	<i>-ing</i> participle of lexical verb (e.g. <i>giving, working</i>)
VVGK	<i>-ing</i> participle catenative (<i>going in be going to</i>)
VVI	infinitive (e.g. <i>to give... It will work...</i>)
VVN	past participle of lexical verb (e.g. <i>given, worked</i>)
VVNK	past participle catenative (e.g. <i>bound in be bound to</i>)
VVZ	<i>-s</i> form of lexical verb (e.g. <i>gives, works</i>)
XX	<i>not, n't</i>
ZZ1	singular letter of the alphabet (e.g. <i>A, b</i>)
ZZ2	plural letter of the alphabet (e.g. <i>A's, b's</i>)

Appendix 2.2 Penn Treebank for spaCy

Source: https://www.ling.upenn.edu/courses/Fall 2003/ling001/penn_treebank_pos.html

See further information on <https://spacy.io/usage/linguistic-features>.

CC	Coordinating conjunction	VBZ	Verb, 3rd person singular present
CD	Cardinal number	WDT	Wh-determiner
DT	Determiner	WP	Wh-pronoun
EX	Existential there	WP\$	Possessive wh-pronoun
FW	Foreign word	WRB	Wh-adverb
IN	Preposition or subordinating conjunction		
JJ	Adjective		
JJR	Adjective, comparative		
JJS	Adjective, superlative		
LS	List item marker		
MD	Modal		
NN	Noun, singular or mass		
NNS	Noun, plural		
NNP	Proper noun, singular		
NNPS	Proper noun, plural		
PDT	Predeterminer		
POS	Possessive ending		
PRP	Personal pronoun		
PRP\$	Possessive pronoun		
RB	Adverb		
RBR	Adverb, comparative		
RBS	Adverb, superlative		
RP	Particle		
SYM	Symbol		
TO	to		
UH	Interjection		
VB	Verb, base form		
VBD	Verb, past tense		
VBG	Verb, gerund or present participle		
VBN	Verb, past participle		
VBP	Verb, non-3rd person singular present		