

# Appendix D

## Part-of-Speech Tags used in the Hepple Tagger

CC - coordinating conjunction: "and", "but", "nor", "or", "yet", plus, minus, less, times (multiplication), over (division). Also "for" (because) and "so" (i.e., "so that").

CD - cardinal number

DT - determiner: Articles including "a", "an", "every", "no", "the", "another", "any", "some", "those".

EX - existential there: Unstressed "there" that triggers inversion of the inflected verb and the logical subject; "There was a party in progress".

FW - foreign word

IN - preposition or subordinating conjunction

JJ - adjective: Hyphenated compounds that are used as modifiers; happy-go-lucky.

JJR - adjective - comparative: Adjectives with the comparative ending "-er" and a comparative meaning. Sometimes "more" and "less".

JJS - adjective - superlative: Adjectives with the superlative ending "-est" (and "worst"). Sometimes "most" and "least".

JJSS - -unknown-, but probably a variant of JJS

-LRB- - -unknown-

LS - list item marker: Numbers and letters used as identifiers of items in a list.

MD - modal: All verbs that don't take an "-s" ending in the third person singular present: "can", "could", "dare", "may", "might", "must", "ought", "shall", "should", "will", "would".

NN - noun - singular or mass

NNP - proper noun - singular: All words in names usually are capitalized but titles might not be.

NNPS - proper noun - plural: All words in names usually are capitalized but titles might not be.

NNS - noun - plural

NP - proper noun - singular

NPS - proper noun - plural  
 PDT - predeterminer: Determinerlike elements preceding an article or possessive pronoun; "all/PDT his marbles", "quite/PDT a mess".  
 POS - possessive ending: Nouns ending in "'s" or "'".  
 PP - personal pronoun  
 PRPR\$ - unknown-, but probably possessive pronoun  
 PRP - unknown-, but probably possessive pronoun  
 PRP\$ - unknown, but probably possessive pronoun, such as "my", "your", "his", "his", "its", "one's", "our", and "their".  
 RB - adverb: most words ending in "-ly". Also "quite", "too", "very", "enough", "indeed", "not", "n't", and "never".  
 RBR - adverb - comparative: adverbs ending with "-er" with a comparative meaning.  
 RBS - adverb - superlative  
 RP - particle: Mostly monosyllabic words that also double as directional adverbs.  
 STAART - start state marker (used internally)  
 SYM - symbol: technical symbols or expressions that aren't English words.  
 TO - literal to  
 UH - interjection: Such as "my", "oh", "please", "uh", "well", "yes".  
 VBD - verb - past tense: includes conditional form of the verb "to be"; "If I were/VBD rich...".  
 VBG - verb - gerund or present participle  
 VBN - verb - past participle  
 VBP - verb - non-3rd person singular present  
 VB - verb - base form: subsumes imperatives, infinitives and subjunctives.  
 VBZ - verb - 3rd person singular present  
 WDT - wh-determiner  
 WP\$ - possessive wh-pronoun: includes "whose"  
 WP - wh-pronoun: includes "what", "who", and "whom".  
 WRB - wh-adverb: includes "how", "where", "why". Includes "when" when used in a temporal sense.

:: - literal colon  
 , - literal comma  
 \$ - literal dollar sign  
 - - - literal double-dash  
 - " - literal double quotes  
 - ` - literal grave  
 ( - literal left parenthesis  
 . - literal period  
 # - literal pound sign  
 ) - literal right parenthesis  
 - ' - literal single quote or apostrophe