

WORDNET ANNOTATION FOR NOUNS

Head node: <n>

Attribute: <sense>

Possible values for the attribute <sense> correspond to the encoding of the synset for each lexical item: two numbers indicating WordNet version (currently 1.6 –encoded as 16), followed by a dot and 8 more numbers (the offset of the synset), and, finally, a letter which indicates the type of PoS (*n* –noun–, *v* –verb–, *a* –adjective–, *r* –adverb–). For instance:

sense = “16.05443572n” corresponds to the nominal synset 05443572n in WordNet 1.6.

Apart from these specific labels, there are seven generic values for the attribute covering a variety of cases:

- **cs1**: “The word exists in dictionary, but not its sense”
- **cs2**: “The word does not exist in dictionary”
- **cs3**: “The word is part of a Multiword Lexical Unit or a lexicalized inflected form”
- **cs4**: “The word is part of a Named Entity”
- **cs5**: “The tagger is strongly uncertain”
- **cs6**: “The word was improperly lemmatized or PoS-Tagged”
- **cs7**: “The word is wrongly used: misspelling or a loanword”