Identifying Syntactic Categories

1 Syntactic categories are distributional not semantic

One of the fundamental problems with the way traditional parts of speech are defined is that they are often a mixture of semantic and syntactic features, and the definitions are not usually explicit enough to be useful. For example, nouns are often said to be a 'person, place, thing or idea', while verbs are said to be 'actions or states of being'. We can immediately see various problems with this kind of definition. For example, in (1a) fighting is a verb, while in (1b), fighting is a noun. But semantically, it seems that fighting in both words is describing an action. The reason we say that the two words are of different categories is because of their syntactic behaviour: in (1a) fighting combines with an auxiliary verbs might and be, while in (1b) it combines with a Determiner the.

(1) a. John might be fighting with Bill.
   b. The fighting was fierce.

We can see another problem with the semantic definition of verbs as states of being if we compare verbs and adjectives, as in (2). In (2a) we cannot say that fear is an action, so it must be a state of being. But then we can’t distinguish (2a) from afraid (2b) which is also a state of being. But afraid is an adjective and not a verb. We can tell this again, by distributional properties: in (2a) fear agrees with the subject fears vs. fear, but in (2b) afraid can’t stand by itself without is. This is why (2c/d) are ungrammatical: in (2c) fear as a verb can’t combine with is and afraid can’t behave like a verb by having agreement with the subject.

(2) a. John fears ghosts.
   b. John is afraid of ghosts.
   c. *John is fear ghosts.
   d. *John afraids of ghosts.

2 Ways of identifying syntactic categories

Instead of semantic properties to determine syntactic categories, we will use groups of distributional tests (not definitions) to determine the category of a word. Remember also, that words can fall into more than one category, so the category of a word is always relative to the sentence it is in.
2.1 NOUNS

- Traditionally: Person, place, or thing
- Distributionally:
  - modified by Adjectives
  - follow determiners (*the, a, this* etc. or the possessive *my, your* etc.)
  - can be singular or plural
  - can’t take an object
- Frame: *X is/are a pain in the neck.*

2.2 VERBS

- Traditionally: Action (sometimes state)
- Distributionally:
  - can combine with auxiliary verbs *can, will, might* etc. or *to.*
  - can agree with the subject
  - takes tense (-ed), aspect (-en) affixes
  - can take an object
- Frame: They can *X* or They *X-ed* the banana.

2.3 ADJECTIVES

- Traditionally: State (modifying), qualities, attributes
- Distributionally:
  - follows *very*
  - modifies noun (and follows determiner)
  - can’t take an object
  - takes derivational affixes like *-ish, -some*
- Frames: She is very *X*; or I want the *X* book.
2.4 ADVERBS

- Traditionally: Modifier of anything other than a noun.
- Distributionally:
  - takes -ly suffix
  - Cannot appear where adjectives often appear (e.g. between determiner and noun.)
  - Can appear between Subject and Aux.
- Frames: Bill treats Fred X; John can X do that.

2.5 PREPOSITIONS

- Traditionally: locations, space, directions
- Distributionally:
  - take no morphology
  - can take an object
  - can be modified by right
- Frame: He is right X. (doesn’t work for of, with and perhaps a few other Ps for semantic reasons.)

3 Lexical vs. Functional Categories

There are two kinds of lexical items (words). *lexical* categories are *open class*, in that there is an unlimited number of them, and new words are added all the time both to the language as a whole, and also to individual speaker’s mental lexicons. *Functional* categories, on the other hand are *closed class* in that they have a small and fixed number of items. These are also the words that traditional grammar usually doesn’t give semantic definitions for either. Since the number of closed class elements is small, it is usually easier to just list them. Here is a partial listing.

- CONJUNCTIONS (Conj): and, or, etc.
- DETERMINERS (D): This, that, the, a, my, your, our, his, her, their, each, every, some, etc.
- COMPLEMENTIZERS (C): that, if, for (all followed by a clause)
- AUXILIARIES/MODALS (T): is, have, can, must, should, would, etc.
4 Examples

4.1 Nouns

Here are some examples of the noun tests in action:

(3) a. Dogs/cats/people/tables/cars are a pain in the neck. (Frame test (also plural))
   b. The dog/my dog/Bill’s dog is a pain in the neck. (Determiner/possessive test)
   c. *John’s fight Bill (Can’t take an object (Bill))

4.2 Verbs

(4) a. John can/might/will leave/eat dinner. (Combines with Auxiliary verb)
   b. John wants to leave. (Combines with to)
   c. eat/ate/eaten; sing/sang/sung; walk/walked/walked (Inflects for tense or aspect)
   d. John ate dinner. (Can take an object).

4.3 Adjectives

(5) a. John is very happy/sad/unhelpful/busy. (Frame test)
   b. I want the big/red/first/top book. (Frame test).
   c. *I am proud Bill (Can’t take an object).

4.4 Adverbs

(6) a. *The quickly man (Can’t appear between Det and N)
   b. John treats Bill well/nicely/better (Frame test)
   c. John can usually/often/never do that. (Frame test)

4.5 Prepositions

(7) John is right outside/inside/under the table/behind me. (’right’ test; also under, behind take objects.)