

Word categories and lexical representation

September 24, 2015

Feature-based category system

Instead of positing a list of categories arbitrarily, we need a more structured system for two reasons:

- a) to minimise storage requirements: compose categories of smaller “building blocks”
- b) to provide an *exhaustive* list of categories: we want to limit the number of possible options

With the use of features with binary values (+ or -), we achieve both goals n features define 2^n categories.

Use 3 features to define the following 8 syntactic categories F (functional) N (nominal), and V (verbal):

		+N	-N
-F	+V	A (adjective/adverb)	V (verb)
	-V	N (noun)	P (preposition)
+F	+V	Deg (degree adverb)	I (inflection)
	-V	D (determiner)	C (complementiser)

Underspecification

We also have the option not to supply a value for a feature. This can be used to describe the categories that seem to be “halfway between” functional and thematic categories. To mark that the feature [F] is underspecified, $[\pm F]$ is used. With $[\pm F]$, four additional categories are added.

$\pm F$	+N	-N
+V	post-determiner	aspectual auxiliary
-V	measure/group noun	non-thematic prepositions

Syntactic information in the lexicon

The lexicon stores all idiosyncratic information about all kinds of morphemes. Apart from phonological and semantic content it must store syntactic characteristics, as well. Lexical entries must show:

- a) syntactic category (using feature values)
- b) thematic grid
- c) subcategorisation frame

The latter two concerns the number and type of arguments a lexical item might have. The arguments are the individuals that are related by a predicate. Based on the number of arguments (*arity*), we can differentiate between one-, two-, and three-place predicates.

What are some of the tests that can determine what belongs in the same category?

The same method of building more complex entities from simple features might be familiar from phonology, too.

What happens to the number of categories if we add an extra feature?

Notice that there are no separate adjectives in this system; only a common category *A*. What could motivate merging these two categories?

Notes about $[\pm F]$ categories:

- a) post-determiners are the elements that can follow determiners and precede nouns (and thematic adjectives) e.g. *the several/many/few red balloons*
- b) non-thematic prepositions are *of* (used in a non-possessive sense) and *by* (as in passive sentences)

Informally put, predicates have “slots” that need to be filled phrases referring to individuals. So if we plug the phrases *John* and *Mary* into the predicate $know(x,y)$, we get a statement (*John knows Mary*) that asserts that the two individuals are in a *know*-relation. This statement, then, can be evaluated as either true or false.

Are there any zero-place predicates?

Thematic roles

In the following example the two arguments play quite different roles.

- (1) a. The horse kicked John.
b. The horse frightened John.

The role that an argument plays in the event described by its predicate is called a *thematic role* (Θ -role). A short list of theta roles:

- *Agent*: the participant deliberately initiating the action
- *Theme*: the participant moved by the action
- *Patient*: the affected participant undergoing the action
- *Experiencer*: the participant experiencing some (psychological, emotional etc.) state
- *Goal*: the participant towards which the action is directed

The ability to assign Θ -roles is one of the main differences between [-F] and [+F] categories.

We are not going to compile a definitive list of Θ -roles as there are many possible versions used by different authors. E.g. many collapse the categories Theme and Patient.

What Θ -role does each argument have in (1)?

Subcategorisation

Lexical elements can also determine the category of their complements, which also has to be included in the lexicon. Hence subcategorisation frames show the category associated with each complement.

Possible lexical entries:

put category: [-F, -N, +V]

Θ -grid: <agent, theme, goal>

subcat: [nominal, prepositional]

think category: [-F, -N, +V]

Θ -grid: <experiencer, proposition>

subcat: [sentential]

Exercises

1. Identify the arguments in the sentences below, and give the lexical entry of the predicates.

- | | |
|---|---|
| a) The train arrived from Paris. | h) The audience was bored of the third act. |
| b) Bill received a letter yesterday. | i) Laura gave everybody a gift. |
| c) The children ate their dinner on the porch. | j) The policemen ordered everybody out of the room. |
| d) I think that all my friends have seen this film already. | k) Sarah seems to have left. |
| e) The armchair sat in the corner. | l) Paul considers crime stories boring. |
| f) Mary sat in the armchair. | m) It has been raining all day. |
| g) Peter and Ann met in the city last week. | |

Reading: Newson et al. *BESE* Ch 1.3 "A Typology of word categories" (pp.10-51)