1. Lexical Categories

If syntactic theory seeks to explain how words combine to create well-formed sentences, then we need a well-defined theory of *words*. You already have a theory of words from grade school: ***Parts of Speech***

However, everything you were told was a lie …

Parts of speech ('PoS'; a.k.a. syntactic categories, lexical categories, word classes)

- Words belong to various categories of different types, e.g., noun, verb, adjective, adverb, preposition, modal, determiner, etc.
- In grade school, you were simply told that there were different categories
- Two empirical methods for showing that different categories exist:

1. Traditional approach to PoS:

- A word's PoS is defined semantically, i.e., according to its meaning.
- Conventional basis for deciding that a group of words belong to the same category is that they can be substituted for one another without affecting grammaticality.

Substitution principle: The result of substituting a word of a category C for another word of the same category does not change the grammaticality of the phrase or sentence in which it appears (Note: This is distinct from impacting meaning, e.g., substitution can create nonsensical sentences)

Discussion: How, or where, does the traditional approach to PoS breakdown?

2. A new approach to PoS: Beyond semantics

- Semantically-based criteria are unreliable.
- Words are like chameleons: they change meaning depending on their surroundings

JABBERWOCKY

Lewis Carroll

(from Through the Looking-Glass and What Alice Found There, 1872)

`Twas **brillig**, and the **slithy toves**
Did **gyre** and **gimble** in the **wabe**:
All **mimsy** were the **borogoves**,
And the **mome raths outgrabe**.

"**Beware** the **Jabberwock**, my **son**!
The **jaws** that **bite**, the **claws** that **catch**!
Beware the **Jubjub bird**, and **shun**
The **frumious Bandersnatch**!"

He **took** his **vorpal sword** in **hand**:
Long time the **manxome foe** he **sought** --
So **rested** he by the **Tumtum tree**,
And **stood** awhile in **thought**.

And, as in **uffish** thought he stood,
The Jabberwock, with **eyes** of **flame**,
Came whiffling through the **tulgey wood**,
And **burbled** as it **came**!

One, two! One, two! And **through** and through
The vorpal **blade** went **snicker-snack**!
He **left** it **dead**, and with its **head**
He went **galumphing** back.

"And, has thou **slain** the Jabberwock?
Come to my **arms**, my **beamish boy**!
O **frabjous day**! Callooh! Callay!'
He **chortled** in his **joy**.

`Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

Key point: Far more reliable for determining PoS are ***morphological*** and ***syntactic*** criteria.

Morphological criteria:

- Certain types of *inflectional* morphemes attach only to specific categories.
- We can identify individual categories according to the range of inflections which they permit

Syntactic / distributional criteria:

- Certain categories appear only in certain sentence positions
- Syntactic distribution provides some of the strongest empirical evidence for different categories.

Take-away message: A *word-level category* is a set of words which share a common set of morphological and/or syntactic properties. Categories are not to be defined semantically.

2. Lexical Sub-Categories

- All words bear a category (e.g., noun, verb, adjective, etc.).
- Some words bear *sub-categories*

Sub-categorization

Lexical items differ according to how many and what types of things they can/must combine with in order to make complete phrases.

For instance, a verb like *kiss* requires a direct object. A common locution is to say that a given verb 'sub-categorizes for' a certain phrase, such as a direct object, meaning that it combines with such as phrase, e.g., *kiss* subcategorizes for a direct object

Argument: general term for the phrase that is sub-categorized for by a word

(Note: arguments are contrasts with adjuncts, i.e., non-obligatory modificational material)

Example: Verb sub-categorization
Verbs can be divided into two subcategories, based on ***valence*** and ***transitivity***

Valence: Total number of arguments that a predicate can take (by analogy with the chemical term).

Transitivity: Total number of arguments that follow the verb

Importantly, subcategorization restricts not only the number of arguments, but also the categories of those arguments.

Take-away message: Some of the rules that we ascribe to how words combine together come from words themselves.