

October 5, 2004

## CLASS 7: CONSTITUENCY & PHRASE STRUCTURE

### CONSTITUENCY

- (1) The structure of this nice sentence should have concerned everyone.

A string of words is any number of words that are (**string**) **adjacent** to one another. However, not every string of words carries the same meaning: *the structure of this nice sentence* vs. *sentence should*. Those strings of words that do carry some meaning or function, or form a unit of sorts, are called **constituents**. For constituency, there are a number of **tests**.

One constituent can be replaced by another of the same type. Before going into different types, let's play with a sequence of words that can be replaced by a single word; such tests are interchangeably known as **Replacement Test**, **Substitution Test**, or **Pronominalization**:

- (2) a. **It** should have concerned everyone.  
 b. \* The structure of this nice **it** have concerned everyone.

Constituents can be represented by **bracketing**. And since a sentence is also a constituent, it carries brackets as well. In fact, constituents may be made up of further constituents, which we can also test for (more tests to consider are **one-substitution** or **do-substitution**):

- (3) a. [ The structure of this nice sentence should have concerned everyone ]  
 b. [ [ The structure of this nice sentence ] should have concerned everyone ]  
 c. [ [ The structure of this [ nice sentence ] ] should have concerned everyone ]  
 d. [ [ The structure of this [ nice sentence ] ] should have [ concerned everyone ] ]

- (4) The structure of this **one** should have concerned everyone.  
 (5) The structure of this nice sentence should have **done**.  
 (6) **It** should have **done**.

Related to such tests, there is the phenomenon of **ellipsis** (deletion of a constituent):

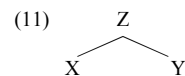
- (7) a. The structure of this nice sentence should have concerned everyone.  
 b. The structure of this nice sentence should have ~~concerned everyone~~.  
 c. The structure of this nice sentence should ~~have concerned everyone~~.

More tests include **movement** (e.g. to the front) and **clfting** (of the form "It is X that..."):

- (8) a. **Everyone**, the structure of this nice sentence should have concerned.  
 b. *It is the structure of this nice sentence that* should have concerned everyone.  
 (9) a. John kissed Mary on the cheek with pleasure.  
 b. **With pleasure**, John kissed Mary on the cheek.  
 c. **On the cheek**, John kissed Mary with pleasure.  
 (10) a. *It is with pleasure that* John kissed Mary on the cheek.  
 b. *It is on the cheek that* John kissed Mary with pleasure.  
 c. \* *It is kissed Mary that* John on the cheek with pleasure.

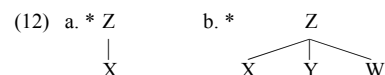
### MERGE

The operation that puts words together to build up a tree is **Merge**. Merge takes two items *X* and *Y* (words, phrases, constituents, etc.), **joins** them, and gives the new **object** a **label** *Z*.

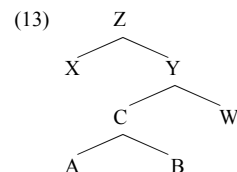


Note that *X* and *Y* can each be more **complex**, of course (i.e. have **internal structure**). The label of the new object is derived from the features of the items joined together, or Merged. The labels of **lexical items** are their **category labels** (D, A, N, P, V, and so on).

And since Merge joins together two items by definition, it is **binary**:



Some **structural relations**: *mother, daughter, sister*  
*contain, immediately contain*  
*branch, node, root node, terminal node*



Merge does not specify **linear order**. We'll address the issue of **linearization** later (Class 9).

Simplest assumption: Merge only applies to root nodes.

### HEADS

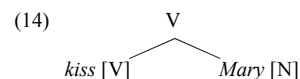
Which syntactic object (string of words, sequence of words, constituent, or other terms) determines the labels of the new object created by Merge then? This will be the **head**. And the name of the object that the head lends its label to is **phrase** — the phrase is a **projection** of the head, or the other way around: **the head projects its label**.

We test for heads through **distribution** (where it can appear within the sentence) and its **sentence frame** (the distribution of the phrase with respect to other elements/constituents — essentially, this is something like its **function** in the sentence: subject, object, and so on).

Two criteria for determining a head:

- (i) *the head of a constituent conditions the distribution of a constituent of which it is a part*  
 (ii) *the head of a constituent is the most important element of the constituent semantically*

This leads us to a **first characterization** of phrases, projections, and the like:



In the next classes, we'll look more into **phrase structure** (but slightly different from *PG*).