1 General introduction

1. Chomsky (1957): a **language** is a set of its grammatical sentences.

1.1. Change "sentences" to "expressions".

1.2. Change "expressions" to "expressions with their structural descriptions".

1.2.1. String similarity is usually misleading:

(1) a. John is easy to please.
    b. John is eager to please.

(2) a. I expected John to leave.
    b. I persuaded John to leave.

(3) a. We expected several students to be at the talk.
    b. We persuaded several students to be at the talk.

(4) a. A unicorn seems to be in the garden.
    b. A unicorn tries to be in the garden.

(5) a. It is easy to play **this sonata** on **this violin**.
    b. **This sonata** is easy to play on **this violin**.
    c. **This violin** is easy to play **this sonata** on.

(6) a. John grows tomatoes.
    b. John destroys tomatoes.

1.3. Item 1.2. brings meaning into picture:

> It is assumed in *LSLT* (as in *SS*) that the theory developed is to be embedded in a broader semiotic theory which will make use of the structure of *L*, as here defined, to determine the meaning and reference of expressions and the conditions on their appropriate use, and will also encompass other investigations (statistical linguistics, etc.). (Chomsky 1975:3)

2. **Grammar** is an explicit system of rules and representations that pairs phonetic forms (sound pathway) with logical forms (meaning pathway).

3. **Acceptability** (data) versus **grammaticality** (theory).

3.1. Factors effective in acceptability are manifold and complex.

3.2. Grammar is just one among many.

3.3. Such idealization is indispensable in science.

3.3.1. Economist Dani Rodrik:

> All models are wrong. They are helpful [when] used in relevant context. Empirics without models yield no understanding.
>
> To clarify, models are wrong in the same sense that a subway map is wrong. Leaves out, misrepresents real world details.
>
> Simplicity in theory is a feature, not a bug. "But the real world is more complicated" is never good riposte. All causal theories simplify. (tweets, March 8, 2017)

4. **Descriptive** versus **explanatory** adequacy.

4.1. Description: "Given a language, what is possible to utter to mean what?"

4.1.1. A descriptively successful grammar for a given language has the widest possible coverage with a minimal set of rules and assumptions.

4.2. Explanation: "What is a possible human language?"

4.2.1. What is the common denominator of all the descriptively successful grammars?

2 Phrase Structure

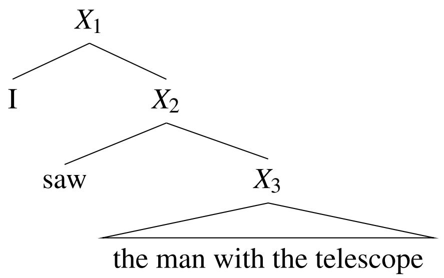
1. "Grammar is about things that go together in meaning" (Mark Steedman)

(7) a. ((Toy car) manufacturer.)
    b. (Toy (car manufacturer.))

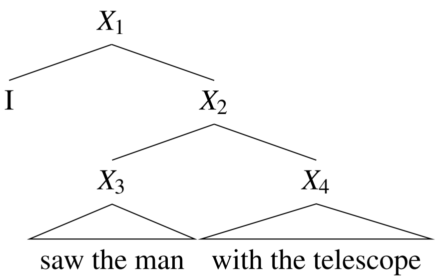
2. A motivation for phrase structure: structural ambiguity.

(8) I saw the man with the telescope.

(9)



(10)



3. A motivation for phrase structure + categories: distributional generalizations.

3.1. The labels X_2 and X_3 are identical when it comes to what can come in the context *I saw* ___.

3.2. Likewise, *I* and *saw* are instances of larger classes that can appear in the places they occupy.

4. Another motivation for phrase structure comes from the need for relational generalizations:

(11) a. John murdered Smith.
b. Smith was murdered by John.

(12) a. The army destroyed the city.
b. The army's destruction of the city.
c. The city's destruction by the army.

(13) a. We expected several students to be at the talk.
b. We expected there to be several students at the talk.

(14) a. John saw the cat that chased the dog.
b. Which dog did John see the cat that chased?

(15) a. John saw the man with the telescope.
b. Which telescope did John see the man with?

(16) a. The man$_i$ who said he$_i$ was tall.
b. *The man$_i$ who he$_i$ said was tall.

5. Such generalizations are best stated over phrase structures.

Complex NP Constraint (Ross 1986):

(17) a. I believed the claim that Otto was wearing this hat.
b. I believed that Otto was wearing this hat.

(18) a. *Which hat did you believe the claim that Otto was wearing?
b. Which hat did you believe that Otto was wearing?

Complex Subject Constraint (Ross 1986):

(19) a. The reporters expected [that the principal would fire some teacher].
b. [That the principal would fire some teacher] was expected.
c. It was expected by the reporters [that the principal would fire some teacher].

Noun phrases in *that*-clauses can be relativized in (19a) and (19c), but not in (19b) (Ross 1986:147).

3 Categories and Features

1. Features start to appear in linguistic generalizations originally in phonetics.

1.1. Some simplified phonetics:

Features: (i) where (**place**) in the vocal tract it is articulated (alveolar, dental, glottal, and so on), (ii) how (**manner**) it is articulated (fricative, plosive, nasal, and so on); and (iii) whether the vocal chords are involved or not (**voice**). A particular consonant can be represented as a set of features.

For example, *t* and *d* are both "alveolar plosives", articulated by restricting (hence plosive) the air flow by tongue and alveolar ridge (hence alveolar). The former is voiceless, while the latter is voiced. Formally, *t* is {[place alveolar],[manner plosive],[voice -]}, while *d* is {[place alveolar],[manner plosive],[voice +]}. We adopt a abbreviating convention, where we omit the words "place" and "manner" and write '[+voice]' instead of '[voice +]', etc.

In its full complexity, the consonant /g/ would be just a short-hand for (Gazdar et al. 1985:17):

(20) $\begin{bmatrix} +\ \text{SEGMENT} \\ +\ \text{CONSONANTAL} \\ -\ \text{SONORANT} \\ -\ \text{SYLLABIC} \\ +\ \text{HIGH} \\ +\ \text{BACK} \\ -\ \text{LOW} \\ -\ \text{ROUND} \\ -\ \text{ANTERIOR} \\ -\ \text{CORONAL} \\ -\ \text{NASAL} \\ -\ \text{CONTINUANT} \end{bmatrix}$

2. Chomsky (1970:207): "[t]he distinction between features and categories is a rather artificial one."

   Chomsky (1970:208): "we might just as well eliminate the distinction of feature and category, and regard all symbols of the grammar as sets of features."

3. Features become useful in various levels of specification: V is enough for tense, but you need subcat for transitive versus intransitive modes of combination.

4. Initially 3 important types of features: projection level (word versus phrase), head features (agreement), combinatory features (valency for now, directionality and order later).

5. Feature structures:

   Definition 3.1 (Feature structures without variables[1])
   Let S_c be the set of symbols standing for complex features and S_a for atomic features. The set of feature structures F is the smallest set satisfying,

   i. If $\alpha \in S_c$ and $\beta \in S_a$, then $\{\langle \alpha, \beta \rangle\} \in F$.
   ii. If $\alpha \in S_c$ and $\beta \in F$, then $\{\langle \alpha, \beta \rangle\} \in F$.
   iii. If F_1 and $F_2 \in F$, and $\mathrm{Dom}(F_1) \cap \mathrm{Dom}(F_2) = \emptyset$, then $F_1 \cup F_2 \in F$.

   □

6. Attribute-value matrices (AVMs) make things more readable.

6.1. Represent $\{\langle \alpha, \beta \rangle\}$ as $\begin{bmatrix} \alpha & \beta \end{bmatrix}$.

6.2. Represent $\{\langle \alpha_1, \beta_1 \rangle, \ldots, \langle \alpha_n, \beta_n \rangle\}$ as $\begin{bmatrix} \alpha_1 & \beta_1 \\ \vdots & \vdots \\ \alpha_n & \beta_n \end{bmatrix}$.

6.3. Here is a definition of feature structures with variables, in AVM notation:

   Definition 3.2
   Let S_c be the set of symbols standing for complex features and S_a for atomic features. The set of feature structures F is the smallest set satisfying,

   i. If $\alpha \in S_c$ and $\beta \in S_a$, then $\begin{bmatrix} \alpha & \beta \end{bmatrix} \in F$.
   ii. If $\alpha \in S_c$ and $\beta \in S_a$ and $n \in \mathbb{Z}^+$, then $\begin{bmatrix} \alpha & \boxed{n}\beta \end{bmatrix} \in F$.
   iii. If $\alpha \in S_c$ and $n \in \mathbb{Z}^+$, then $\begin{bmatrix} \alpha & \boxed{n} \end{bmatrix} \in F$.
   iv. If $\alpha \in S_c$ and $\beta \in F$, then $\begin{bmatrix} \alpha & \beta \end{bmatrix} \in F$.
   v. If $\gamma \in F$ and not formed by this rule, and $n \in \mathbb{Z}^+$, then $\boxed{n}\,\gamma \in F$.

[1] See the appendix for a set-theoretic formulation of feature structures with variables.

vi. If $F_1 = \begin{bmatrix} \alpha_1 & \beta_1 \\ \vdots & \vdots \\ \alpha_m & \beta_m \end{bmatrix}$ and $F_2 = \begin{bmatrix} \gamma_1 & \delta_1 \\ \vdots & \vdots \\ \gamma_n & \delta_n \end{bmatrix}$ for $m, n \geq 1$ are both $\in F$,
and if, disregarding any possible indices heading α_i and γ_j,

$$\{\alpha_1, \dots, \alpha_m\} \cap \{\gamma_1, \dots, \gamma_n\} = \emptyset, \text{ then } \begin{bmatrix} \alpha_1 & \beta_1 \\ \vdots & \vdots \\ \alpha_m & \beta_m \\ \gamma_1 & \delta_1 \\ \vdots & \vdots \\ \gamma_n & \delta_n \end{bmatrix} \in F.$$

□

7. Sag et al.'s (2003) differ from Definition 3.2 by having a type decoration for each feature structure. Type decorations can be fitted into the system of Definition 3.2 by introducing an obligatory feature TYPE for each feature structure. The following equivalence would hold:

$$(21)\quad \begin{bmatrix} \textit{phrase} & \\ \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AGR} & \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{PER} & \text{3rd} \end{bmatrix} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \textit{val-cat} & \\ \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix} \end{bmatrix} \equiv \begin{bmatrix} \text{TYPE} & \text{phrase} \\ \text{HEAD} & \begin{bmatrix} \text{TYPE} & \text{verb} \\ \text{AGR} & \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{PER} & \text{3rd} \end{bmatrix} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{TYPE} & \text{val-cat} \\ \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix} \end{bmatrix}$$

8. Sag et al.'s (2003) type decorations should not be confused with atomic features; their notation is a little confusing in this. For example, the following sentence category,

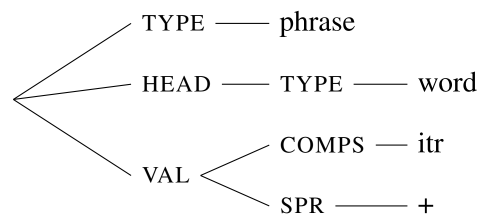
$$(22)\quad \begin{bmatrix} \textit{phrase} & \\ \text{HEAD} & \textit{verb} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix} \end{bmatrix}$$

is actually,

$$(23)\quad \begin{bmatrix} \textit{phrase} & \\ \text{HEAD} & \begin{bmatrix} \textit{verb} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix} \end{bmatrix} \equiv \begin{bmatrix} \text{TYPE} & \text{phrase} \\ \text{HEAD} & \begin{bmatrix} \text{TYPE} & \text{verb} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix} \end{bmatrix}$$

9. Feature structures can also be represented as **directed acyclic graphs** (DAGs) with a unique root constraint. For example the AVM in (23) is,

(24)



10. A path in such a DAG is a successive application of feature names to a feature structure to access its sub-features. For instance, naming the above AVM F, we have,

$$F(<\text{TYPE}>) = \text{phrase}$$
$$F(<\text{VAL COMPS}>) = \text{itr}$$

11. A feature structure F_1 **subsumes** another F_2, shown $F_1 \sqsubseteq F_2$ if F_2 has all the information F_1 has.

Definition 3.3 (Subsumption)
Given F_1 and F_2:

(a) $F_1 \sqsubseteq F_2$, if F_1 and F_2 are atomic features and equal.

(b) $F_1 \sqsubseteq F_2$, if F_1 and F_2 are feature structures such that $\mathrm{Dom}(F_1) \subseteq \mathrm{Dom}(F_2)$ and for each $f \in \mathrm{Dom}(F_1)$, $F_1(\langle f \rangle) \sqsubseteq F_2(\langle f \rangle)$.

□

12. Given two feature structures F_1 and F_2, their **unification** $F_2 = F_1 \sqcup F_2$, if exists, is the smallest feature structure subsumed by both, that is $F_1 \sqsubseteq F_3$ and $F_2 \sqsubseteq F_3$.

References

Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.

Chomsky, N. (1970). Remarks on nominalization. In Jacobs, R. and Rosenbaum, P., editors, *Readings in English Transformational Grammar*, pages 184–221. Ginn, Waltham, Mass.

Chomsky, N. (1975). *The Logical Structure of Linguistic Theory*. Plenum Press, New York.

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Harvard University Press, Cambridge, MA.

Ross, J. R. (1986). *Infinite Syntax!* Ablex Publishing Corporation, Norwood, NJ.

Sag, I. A., Wasow, T., and Bender, E. M. (2003). *Syntactic Theory: A Formal Introduction*. CSLI, Stanford, CA.

A Definitions

Definition A.1 (Feature structures with variables – variable clash possible)
Let S_c be the set of symbols standing for complex features, S_a for atomic features, both disjoint from $\mathbb{Z}^+$. Let K_c be the set of indexed complex feature names defined as $\{\langle n, \alpha \rangle \mid \alpha \in S_c \text{ and } n \in \mathbb{Z}^+\}$, and likewise let K_a be the set of indexed atomic feature names defined as $\{\langle n, \alpha \rangle \mid \alpha \in S_a \text{ and } n \in \mathbb{Z}^+\}$.

The set of feature structures F is the smallest set satisfying,

1. If $\alpha \in K_c$ and $\beta \in K_a \cup \mathbb{Z}^+$, then $\{\langle \alpha, \beta \rangle\} \in F$.
2. If $\alpha \in K_c$ and $\beta \in F$, then $\{\langle \alpha, \beta \rangle\} \in F$.
3. If F_1 and $F_2 \in F$, and $\mathrm{Range}(\mathrm{Dom}(F_1)) \cap \mathrm{Range}(\mathrm{Dom}(F_2)) = \emptyset$, then $F_1 \cup F_2 \in F$.

□