Syntactic Theory 2
Week 1: Introduction, and Minimalism

Dustin A. Chacón

January 16, 2017

- [Syllabus]

1 Goals of Generative Grammar

- Speakers of English can understand a variety of novel expressions:

  (1) My cat Ernie didn't like his diet cat food so he went on a hunger strike

- Grammatical knowledge, or **i-language**, is a cognitive system that allows us to pair up sound and meaning in a systematic way
- We also distinguish **competence** from **performance** – i-language is competence divorced from memory, attention, etc.
- In generative grammar, we suppose that parts of our competence are innate, i.e., biologically pre-determined. Humans are biologically equipped with a **Universal Grammar** (UG)
- An illustration: language is **structure dependent** – rules of grammar are sensitive to grammatical structure, not to linear order:

  (2) a. The eagles can fly
  b. Can the eagles *t* fly?

  (3) Can [$_{DP}$ the eagles that (**t*) swim] *t* fly?

- Chomsky suggests that children ought to infer from (2) that auxiliary-fronting rules are defined over the *first* auxiliary
- However, children seem to correctly identify the rule as one defined over the *main* auxiliary (see Crain & Nakayama 1987 for acquisition evidence)
- The larger issue is that unless children have a constrained hypothesis space (= UG), children will always have infinite hypotheses available for the observed data (cf. Goodman's (1955) "the new riddle of induction").

- On this view, all children are born into an *initial state*, specified by UG. After sufficient exposure to their language input, they grow into adults with a fixed grammar of their language. UG provides the "blueprint" for how the language can look, and the input helps them shape the details.
- If so, we can discuss **principles** of grammars and **parameters**. Principles are those aspects of grammars that are universal and invariable – these are properties of grammars that UG gives to the child "for free". Parameters are those aspects of the grammars that the child must "set" to get the grammar of their linguistic community.
- We have two primary goals in lingusitic theory – **descriptive adequacy** and **explanatory adequacy**. A theory is descriptively adequate if it generates the right structural descriptions for a language; a theory if explanatorily adequate if it explains how a child can leverage their linguistic experience and UG to arrive at that grammar.

2 Why Try Minimalism?

- Chomsky's approach to language has its roots in his 1955 dissertation *The Logical Structure of Linguistic Theory*, which applied a mathematically precise formalism to English grammar. With his associates, it developed into the Standard Theory, then the Extended Standard Theory, and then finally into Government & Binding in the 80's to the mid-90's.
- GB is just one of many "generative grammars" – HPSG, LFG, CCG, TAG, RRG are cousins to GB in some respects.
- GB has a few characteristic properties:
  - GB is **modular** – different components of the grammar interact to determine the structure of a sentence. E.g., X-bar theory, Case Theory, Theta Theory, Binding Theory, EPP, ...
  - GB has four **levels** – D-Structure, S-Structure, LF, and PF. This means a sentence is actually a quadruplet of representations, which are related by transformations (Move α, Deletion, Substitution...). Some modules apply at D-Structure, others at S-Structure, others at LF & PF.
  - GB is **representational** – it characterizes which structures are good, and which are bad. It does not explain *how* those structures are generated. In other words, the different modules are filters on the output of some generative procedures.
- The Minimalist Program was set out in Chomsky's (1995) book *The Minimalist Program*, which set the agenda for how to rethink GB and its properties
- Minimalism forces us to ask: how "small" can we make GB, while maintaining its insights and successes? Additionally, how can we explain the properties of our theoretical formatives?
- Minimalism has the following properties:

- Emphasis on theoretical formatives that are **virtually conceptually necessary** – grammars are structured mappings between sound and meaning. That means there has to be some function that produces structure, and some mapping to sound (= PF) and some mapping to meaning (= LF). That means everything else (= D-Structure, S-Structure) should be dispensed with, unless we find empirical reasons to keep it.
- The **strong minimalist thesis** (SMT) – language is an optimal solution to the interfaces. In other words, syntactic operations occur because the semantics (= LF) or the phonology (= PF) requires it to, not because a grammar-internal module requires it. This means we need to rethink why certain grammatical properties exist (case and agreement are particularly hard nuts to crack here)
- Computational simplicity and **economy** – grammars prefer the most economical derivations. This is a fuzzy notion, but essentially – movements should be as short as possible, derivations should include the fewest steps possible, and operations should only occur if necessary (to satisfy an LF or PF requirement)
- Minimalism is a mix of a representational theory and a derviational theory – some sentences are bad because they are ill-formed at PF and LF, meaning that filters at these levels flag them as ungrammatical. However, some sentences are bad because they violate Economy.

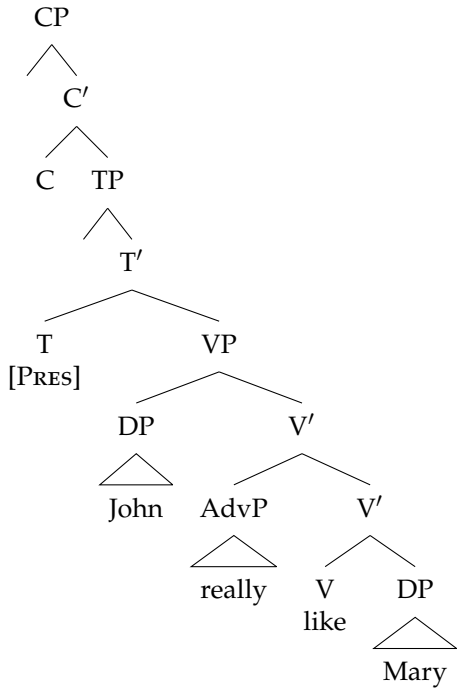
- Minimalism is motivated both by typical scientific meta-principles (e.g., Occam's Razor), and evolutionary concerns. If we take UG to be a biologically pre-determined part of our cognitive capacity that's unique to humans, then it must have *evolved* somehow. By "shrinking" UG and attributing more properties of the language faculty to extra-linguistic cognitive systems (e.g., the conceptual-intentional system (CI) and the articulatory-phonetic system (AP) systems; and general principles of computation), we can begin theorizing about how language evolved in the human species as a matter of one or two small mutations. There is a closely related field called "biolinguistics" which addresses the relation between syntactic theory and evolution (cf. Hauser, Chomsky, & Fitch 2002)

- These goals are very abstract, and are not necessarily "right". Perhaps language looks like GB, and evolved through a complex history. However, by interrogating the "grammar-internal" parts of our theory, we might discover that we made some assumptions that were not warranted, or come up with simpler versions of GB, which simplify the evolution question.

3 An Example: D-Structure and Phrase Structure

- GB proposes that one representation of a sentence is the interface between the syntax and the lexicon, called D-Structure. D-Structure feeds all movement operations, and X-Bar Theory (and Theta Theory) apply here:

(4)



- X-Bar theory states that all phrases adhere to the following pattern:

  (5) a. XP $\rightarrow$ WP X′
  b. X′ $\rightarrow$ X′ YP
  c. X′ $\rightarrow$ X ZP

- D-Structure is not "virtually conceptually necessary". So, let's toss it. Additionally, X-Bar Theory (as stated) is "grammar-internal". Can we re-imagine the properties of X-Bar Theory with the Minimalist agenda in mind?

- Proposal: phrase structure is built by an operation called **Merge**. Merge selets two lexical items α and β, and outputs a complex unit consisting of α and β, with one as the "label":

  (6) a. **Merge**(α,β) = {α, { α, β}}
  b. **Merge**(α,β) = α (with children α and β)

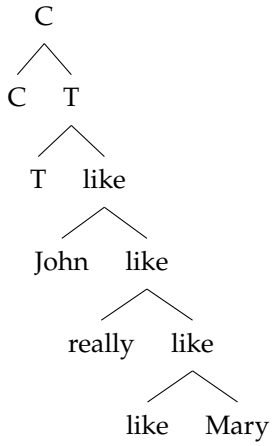
- Abolishing D-Structure means that we can apply each instance of Merge at a time – there is no single representation that satisfies X-Bar theory.

  (7) a. Merge(like, Mary) = [$_{\text{like}}$ like Mary]
  b. Merge(really, like) = [$_{\text{like}}$ really [$_{\text{like}}$ like Mary]]

c. Merge(John, like) = $[_{like}$ John $[_{like}$ really $[_{like}$ like Mary]]]
d. Merge(T, like) = $[_{T}$ T $[_{like}$ John $[_{like}$ really $[_{like}$ like Mary]]]]
e. Merge(C, T) = $[_{C}$ C $[_{T}$ T $[_{like}$ like John $[_{like}$ really $[_{like}$ like Mary]]]]]

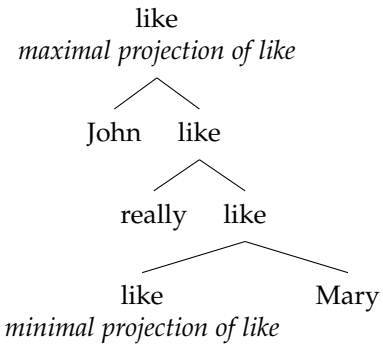
- The final structure above would look something like this in **Bare Phrase Structure**:

(8)



- In X-Bar Theory, branching is always binary. Why? We say that Merge applies to two elements because this is the minimal combinatoric operation. Thus, the fact that phrase structure is largely binary branching follows from considerations of computational simplicity
- In X-Bar Theory, we distinguish three levels of structure – XP, X′, and X^0, which are grammatical-internal notions. With Merge, we dispense with bar-levels entirely, as these are grammatical-internal notions. Rules may apply that target one specific level (head movement only targets X^0, for example). We can say re-imagine rules as targeting the *minimal projection* or *maximal projection* of *like*, without needing the representation to encode these distinctions:

(9)



- In X-Bar Theory, it's simply stipulated that the XP level dominates the X′ level, which dominates the X^0 level. Additionally, it is stipulated that there is only one XP level and

only one X^0 level, but potentially many X' levels that are equivalent. In a Merge-based Bare Phrase Structure theory, these facts follow for free – by definition, the maximal projection must be the largest projection of a word, and cannot be dominated by anything else, and a minimal projection must be the smallest. All intervening projections are non-maximal, non-minimal "intermediate projections".

- We will return to Merge and Bare Phrase Structure later

Bibliography

Chomsky, Noam. 1955. *The Logical Structure of Linguistic Theory*.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: The MIT Press.

Crain, Stephen, Mineharu Nakayama. 1987. Structure Dependence in Grammar Formation. *Language* 63(3), 522–434.

Goodman, Nelson. 1955. *Fact, Fiction, and Forecast*. Cambridge: Harvard University Press.

Hauser, Marc D., Noam Chomsky, W. Tecumseh Fitch. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298(22), 1569–1579.