Syntactic Theory 2
Week 4: Minimalism

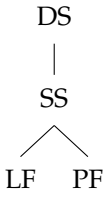
Dustin A. Chacón

January 29, 2017

1 Introduction

- After having reviewed GB to death, let's remind ourselves of some of its properties.[1]

(1)



- Only LF and PF are "virtually conceptually necessary", thus DS and SS ought to be eliminated, and their empirical work distributed to the LF and PF levels.
- **D-Structure**: Theta Theory, X′-Theory
- **S-Structure**: Case Theory, Binding Theory, Subjacency, distinctions between overt/covert movement

2 Eliminating S-Structure

(2) a. DS: [$_{\text{TP}}$ was [$_{\text{VP}}$ seen 3.SG.M]]
b. SS: [$_{\text{TP}}$ he$_{[\text{NOM}]i}$ was [$_{\text{VP}}$ seen t_i]]

(3) a. DS: [$_{\text{TP}}$ T [$_{\text{VP}}$ seems [$_{\text{TP}}$ to [$_{\text{VP}}$ be likely [$_{\text{TP}}$ 3.SG.M to [$_{\text{VP}}$ win]]]]]]
b. SS: [$_{\text{TP}}$ he$_{[\text{NOM}]}$ T [$_{\text{VP}}$ seems [$_{\text{TP}}$ t'_i to [$_{\text{VP}}$ be likely [$_{\text{TP}}$ t_i to [$_{\text{VP}}$ win]]]]]]

- When does Case Theory apply?
- DS? Clearly not – otherwise, no DP could start in a Case-less position.

[1] This lecture is largely based off of Hornstein, Grohmann, & Nunes (2006).

- PF? Nope – A-movement for Case feeds binding, and therefore affects the semantics of a sentence. Thus, it must happen somewhere in the narrow syntax (DS→SS→LF):

  (4) He$_i$ is likely t'_i to seem to himself$_i$ t_i to be perfect, (that narcissist).

- Additionally, A-movement affects the scope relations of sentences:

  (5) a. Many arrows didn't [$_{\text{VP}}$ t_i hit the target].
  There were many arrows that didn't hit the target
  (but perhaps there were also many arrows that DID.)
  b. The target$_i$ wasn't [$_{\text{VP}}$ hit t_i by many arrows]
  It is not the case that many arrows hit the target

- LF? No. Case-assignment feeds the phonology, because the morphological form of a pronoun is determined by its syntactic position. Thus, it must be in the "overt syntax" (DS→SS→PF)

  (6) He/*him likes *he/him

- Thus, we need to assign Case at some point in the derivation after DS that feeds both LF and PF. Looks like we're stuck with SS?

- Chomsky (1995) points out that this reasoning assumes that Case is *assigned* – DPs are unmarked for Case until they appear in their Case position.

- Let's instead suppose that DPs enter the derivation with their Case already specified. On this view, DPs must move to *check* their Case:

  (7) a. DS: [$_{\text{TP}}$ was [$_{\text{VP}}$ seen he$_{[\text{Nom}]}$]]
  b. SS: [$_{\text{TP}}$ he$_{[\cancel{\text{Nom}}]i}$ was [$_{\text{VP}}$ seen t_i]]

- On this view, it is no longer necessary to suppose that Case must be checked at PF! Now, we can rethink the Case Filter as follows:

  (8) **Case Filter**: All DPs must have their Case checked by LF

- The form of explanation for Case phenomena is the same as before – DPs now violate the Case Filter not because they lack Case, but because their Case filters go unchecked.

- This is a theoretical shell game in some respects. However, any time we can explain the same set of data by restating something as an LF or PF constraint, we "win" according to Minimalist desiderata. It's an extra cherry on top if we find increased empirical coverage with Minimalist technology.

- And, cherries, there are. Consider existential constructions:

  (9) There is a cat on the mat.

- Apparently, *a cat* receives Case of some kind. The only Case assigner in the sentence is *is*, but *a cat* is not in Spec,TP, the canonical position for Nominative Case assignment.
- To account for this, GB theorists had postulated that Case is not assigned to DPs, but to chains. Chains were typically understood as sections of a movement relation. But, we enrich the construct of chain to also include "expletive-associate" pairs. We mark these with superscripting[2]. Thus, Nominative Case is assigned to the chain ⟨there, a cat⟩.

  (10) $[_{\text{TP}}$ there^i is $[_{\text{PP}}$ a cat^i on the mat]]

- On a Case-checking approach, *a cat* has its Nominative Case feature pre-specified, and it must be checked at LF. Luckily, we have a mechanism for this – covert movement:

  (11) LF: $[_{\text{TP}}$ a cat$_{\text{[}\cancel{\text{Nom}}\text{]}i}$ there is $[_{\text{PP}}$ t_i on the mat]]

- Is there evidence for this? Yes – associates can bind from their LF position. (Doesn't Binding Theory apply at SS too?: We'll eventually re-imagine Binding Theory as LF conditions as well, don't worry.)

  (12) a. There seems to himself$_i$ to be a man$_i$ in the mirror.
  b. LF: a man$_i$ there seems to himself$_i$ t'_i to be t_i in the mirror.

- The idea that Case features must be "checked" is a general idea in Minimalism. We will suppose that phrases enter the derivation with many features specified, and must be "checked" in order to be legitimate PF or LF objects.[3]
- We have not yet explained why DPs have Case features in the first place, nor why they must be checked, nor why (in general) movement to check Case occurs in the overt syntax.

2.1 Binding Theory

(13) a. He$_{*i}$ greeted Mary after John$_i$ walked in.
b. After John$_i$ walked in, he$_i$ greeted Mary.

(14) DS: [he$_i$ greeted Mary $[_{\text{CP}}$ after John$_i$ walked in]]
SS/LF: [$[_{\text{CP}}$ after John$_i$ walked in]$_j$ he$_i$ greeted Mary t_j]

- If Principle C applied at DS, then both sentences in (13) should be ruled out, as they both share a DS. So, it must be SS or LF. Which is it?
- Let's take an excursus into covert *wh*-movement. Recall that pair-list readings are analyzed as covert *wh*-movement:

  (15) a. SS: Who ate what?

[2]Importantly, NOT subscripting/indexing – otherwise, this would be a Principle C violation!

[3]In recent parlance, features that lead to ungrammaticality if unchecked are called "derivational time bombs", thanks to Omer Preminger.

b. LF: Who$_i$+what$_j$ t_i ate t_j?
For which pairs $\langle x, y \rangle$, x ate y?
'Ernie ate some kibble, Frisky ate some milkbones, Lacey ate some tuna . . . '

- With this in place, let's examine the coreference relations of the three sentences below:

(16) Which picture that Harry$_i$ bought did he$_i$ like?
a. DS: He$_i$ did like which picture that Harry$_i$ bought
b. SS: [Which picture that Harry$_i$]$_j$ bought did he$_i$ like t_j?
c. LF: [Which picture that Harry$_i$]$_j$ bought did he$_i$ like t_j?

(17) *He$_i$ liked this picture that Harry$_i$ bought.
a. DS: He$_i$ liked this picture that Harry$_i$ bought
b. SS: He$_i$ liked this picture that Harry$_i$ bought
c. LF: He$_i$ liked this picture that Harry$_i$ bought

(18) *Which man said he$_i$ liked which picture that Harry$_i$ bought?
a. DS: Which man said he$_i$ liked which picture that Harry$_i$ bought
b. SS: [Which man]$_j$ t_j said he$_i$ liked which picture that Harry$_i$ bought
c. LF: [which picture that Harry$_i$ bought]$_k$ [which man]$_j$ t_j said he$_i$ liked t_k

- We've already established that Principle C cannot apply at D-Structure. However, examining (18), we see that it cannot apply at LF either. If we accept this covert movement story for pair-list interpretations, then the moved DP *which picture that Harry bought* is no longer c-commanded by *he*. If Principle C were checked at LF, then there should be no violation. But, there is a violation – therefore, Principle C must be checked before LF – i.e., SS.

- In Minimalism, we will assume that grammars do the minimal work possible. On this view, why should the *entire wh*P covertly move?:

(19) which$_k$ [which man]$_j$ t_j said he$_i$ liked t_k picture that Harry$_i$ bought

- If this is the LF representation, then *Harry* is still c-commanded by *he* at LF. Thus, it is still a Principle C violation – this is a correct prediction.

- But then, why is this sentence bad?:

(20) *Which which man said he liked picture that Harry bought?

- Perhaps the unacceptability of (20) is due to the presence of an illegitimate PF object. By hypothesis, covert operations don't care about PF legibility (and vice versa). As we saw in HW1, ellipsis can heal errant PF representations:

(21) John said he liked a picture that Harry bought, but I don't remember . . .

a. *which$_i$ John said he liked a picture that Harry t_i bought.
b. which$_i$ ~~John said he liked a picture that Harry t_i bought~~.

- Furthermore, there *are* languages are these kinds of movements are permitted:

(22) a. Combien de livres$_i$ a-t-il consultés t_i?
How many of books hase-he consulted
b. Combien$_i$ a-t-il consultés t_i de livres?
How many has-he consulted of books
'How many books has he consulted?' (French)

2.2 Overt vs. Covert Movement

(23) a. What$_i$ did Bill buy t_i?
b. Bill mai-le shenme?
Bill bought-Asp what
'What did Bill buy?' (Mandarin Chinese)

(24) a. John often drinks wine.
b. Jean bois$_i$ souvent t_i du vin.
Jean drinks often of wine
'Jean often drinks wine'

- Traditionally, the assumption is that languages "look alike" at DS and LF – Mandarin covertly moves *shenme* to Spec,CP at LF, and English covertly moves *drinks* to T^0.
- Traditionally, features that are "morphologically strong" need to be checked before SS, presumably because they have some morphological realization, and therefore PF needs to "see" it[4].
- However, this actually does not motivate a singular syntactic level at which we check whether all strong features have been checked in the overt syntax. All that it requires is some split-off point in which one copy of the representation is fed to PF and another is fed to LF – call this **Spell-Out**.
- Eventually, we will develop a theory that forces features to be checked "as soon as possible"[5].

(25) English:
a. DS: $C_{[WH,STRONG]}$ Bill bought what$_{[WH]}$
b. "SS": what$_{[WH]i}$ $C_{[\sout{WH}]}$ Bill bought t_i

(26) Mandarin:
a. DS: $C_{[WH,weak]}$ Bill bought what$_{[WH]}$
b. LF: what$_{[WH]i}$ $C_{[\sout{WH}]}$ Bill bought t_i

- The strong/weak distinction can be attributed to an Economy condition called Procrastinate:

[4]Of course, this raises the question of whether some movement might actually occur at PF, and have no semantic effect, as explored in HW1!
[5]And in recent work, "simultaneously."

(27) **Procrastinate:** Do not move a phrase overtly if you can wait to move it covertly.

- Whose features need to be checked in *wh*-movement – $C_{[WH]}$, or *wh*? Is there a strong "C-feature" that *wh*-phrases have in English?

(28) a. Who t_i gave what to whom?
b. *Who what to whom gave?

- When we move one *wh* phrase, the C^0's [WH]-feature is checked.
- Bulgarian might have strong C-features on *wh*-operators:

(29) a. *Koj dade kakvo na kogo?
who gave what to whom
b. Koj kakvo na kogo dade?
who what to whom gave
'Who gave what to whom?' (Bulgarian)

- On this view, we can completely dispense of the EPP as a module of the grammar, and reinterpret it as a strong D-feature on T^0. It's *as stipulative* as the original EPP, but is couched within a more general technical vocabulary.
- This also raises new interesting technical issues regarding features – who's responsible for movement? We'll turn to this issue later in the semester.
- We also have another issue to resolve – why Procrastinate? Why is it "cheaper" to move covertly?
- Finally – is Spell-Out just another name for SS? Possibly. We'll see that Spell-Out can be understood as an *operation* (sometimes called **Transfer**) that applies conditionally at multiple points in the derivation, and likely is linked to other considerations of computational simplicity and domain-restricted operations. Stay tuned.

3 Eliminating D-Structure

- D-Structure: X′-Theory, Theta Theory, phrase structure, input to Move α
- Minimalist challenge – does there need to be a single representation where all four of these conditions are satisfied simultaneously?
- **Proposal**: Phrase structure is iteratively built through an operation Merge:

(30) **Merge**$(\alpha, \beta) = [_\alpha\ \alpha\ \beta]$

- This is an operation that accepts two arguments, and outputs a structured object composed of both arguments and "labels" them. For now, we will take for granted that the labels comply with X′-theory – but this will quickly be abandoned.

(31) John said that Bill saw Mary.

(32)

Merge(saw,Mary)	$[_{V'}$ saw Mary]
Merge(John,V′)	$[_{VP}$ John $[_{V'}$ saw Mary]]
Merge(T,VP)	$[_{T'}$ $[_{VP}$ John $[_{V'}$ saw Mary]]]
...	

- In effect, we're abandoning the idea that there is a unified representation with all theta-roles discharged, and replacing it with an iterative procedure that builds structures that comply to those constraints on DS that we know from GB. This is a small move, but it opens the way to think about old ideas in new ways.
- Furthermore, to maintain the generalization that DPs start off in their theta position, let's rethink the Theta Criterion as follows:

(33) **Theta-Role Assignment Principle** (TRAP): theta roles can only be assigned under a Merge operation

- Let's examine raising & control:

(34) a. John$_i$ wants e$_i$ to be sleeping
b. John$_i$ seems e$_i$ to be sleeping

(35) a. It seems that John is sleeping
b. John$_i$ seems e$_i$ to be sleeping
c. *It wants that John is sleeping

(36) a. The doctor wanted e$_i$ to examine the patient
b. The patient wanted e$_i$ to be examined by the doctor
c. The doctor seemed e$_i$ to examine the patient
d. The patient seemed e$_i$ to be examined by the doctor

- The primary observation is that in control structures, the matrix subject seems to receive its thematic interpretation from both the main clause verb and the embedded verb. In GB, the empty category in (34-a) cannot be a trace, because D-Structure is both the level of theta-role assignment *and* the input to movement operations. That is, there's no way to "move into" a theta position. For this reason, D-Structure assumptions in GB force us to postulate PRO, which then forces us to postulate a Control Module and derive the PRO Theorem to explain its distribution. We will interrogate the need for PRO in the third part of the semester.
- The Theta Criterion is quite easy to reimagine as an LF constraint in Minimaslim, because we assume that movement leaves a trace – that is, the thematic structure is actually represented throughout the course of the derivation, so nothing is lost by moving Theta Theory to the end
- However, where does this leave movement? Traditionally, D-Structure is the input to movement operations. But now, there is no D-Structure. Suppose that we can alternate Merge and Move – what does this buy us?

(37) Mary laughed at whom she was looking

- This is a problem for GB, because we would need to postulate **??** as the D-Structure, but this does not capture the fact that the *at* in *at whom* satisfies the selectional criterion of *laugh*, and it implies that *laugh* can take CP arguments, contrary to fact:

  (38) a. D-Structure: *[$_{TP}$ Mary laughed [$_{CP}$ she was looking at whom]]
  b. *Mary laughed [$_{CP}$ that she was looking at Jay]

- Interleaving Merge (the operation under which theta roles can be assigned/selectional criteria can be checked) and Move fixes this:

  (39) [$_{CP}$ C$_{[+WH]}$ she was looking at whom]
  [$_{CP}$ at whom$_i$ C$_{[+WH]}$ she was looking t_i]
  [$_{VP}$ laughed [$_{CP}$ at whom$_i$ C$_{[+WH]}$ she was looking t_i]]
  [$_{TP}$ Mary [$_{VP}$ laughed [$_{CP}$ at whom$_i$ C$_{[+WH]}$ she was looking t_i]]]

- We can make a similar move for *tough*-movement:

  (40) *Moby Dick* is hard to read

- *Moby Dick* is both the thematic object of *read* and the subject of *hard to read*. This looks like regular A-movement –

  (41) *Moby Dick is hard for Bill to decide when to read

- So, *prima facie*, we should have the following S-Structure representation:

  (42) [$_{TP}$ *Moby Dick*$_i$ [$_{VP}$ is [$_{AP}$ hard [$_{CP}$ for [$_{TP}$ Bill to [$_{VP}$ read t_i]]]]]

- However, that should be a Principle A violation (if we take A-traces to be anaphors), or a hyperraising construction (because we move over *Bill*). Furthermore, it's unclear why the DP should move in the first place, given that it can stay *in-situ* and receive Case:

  (43) It is hard for Bill to read *Moby Dick*

- Chomsky (1981) tells us that this is actually A′-movement of an operator. This explains why it can be long-distance, and but is island-sensitive:

  (44) *Moby Dick* is hard [$_{CP}$ OP$_i$ for Bill to {say that/*wonder whether} Mary likes t_i]

- This doesn't quite work, though. First of all, we know the matrix subject position isn't a theta position, since it can be filled by an expletive, as in (43). We have to say something like the matrix subject position is only a theta position *after* operator movement:

  (45) Merge(read,OP): [$_{VP}$ read OP]
  Merge…: [$_{CP}$ C$^0_{[+WH]}$ [$_{TP}$ Bill to [$_{VP}$ read OP]]

Move OP:	$[_{CP}\ OP_i\ C^0_{[+WH]}\ [_{TP}$ Bill to $[_{VP}$ read $t_i]]]$
Merge(hard,CP):	$[_{A'}$ hard $[_{CP}\ OP_i\ C^0_{[+WH]}\ [_{TP}$ Bill to $[_{VP}$ read $t_i]]]]$
Merge(Moby Dick, A′)	$[_{AP}$ MD $[_{A'}$ hard $[_{CP}\ OP_i\ C^0_{[+WH]}$ Bill to read $t_i]]]]$

- So far, we've been consistently merging at the root of the tree. Let's suppose that this is the case:

  (46) **Extension Condition**: Overt applications of Merge and Move can only target the root of the tree[6]

- With this in place, we actually derive what's called the cyclicity condition:

  (47) *I wonder what you asked how John fixed?
  D-Structure: $[_{TP}$ I wonder $[_{CP}\ C^0\ [_{TP}$ you asked $[_{CP}\ C^0\ [_{TP}$ John $[_{VP}$ fixed what] how]]]]]
  S-Structure: $[_{TP}$ I wonder $[_{CP}$ what$_k$ $C^0\ [_{TP}$ you asked how$_i$ $C^0\ [_{TP}$ John fixed $t_k]\ t_i]]]$
  * by Subjacency

- However...

  (48) D-Structure: $[_{TP}$ I wonder $[_{CP}\ C^0\ [_{TP}$ you asked $[_{CP}\ C^0\ [_{TP}$ John $[_{VP}$ fixed what] how]]]]]
  $[_{TP}$ I wonder $[_{CP}\ C^0\ [_{TP}$ you asked $[_{CP}$ what$_k$ $C^0\ [_{TP}$ John $[_{VP}$ fixed $t_k]$ how]]]]]
  $[_{TP}$ I wonder $[_{CP}$ what$_k$ $C^0\ [_{TP}$ you asked $[_{CP}\ t'_k\ C^0\ [_{TP}$ John $[_{VP}$ fixed $t_k]$ how]]]]]
  $[_{TP}$ I wonder $[_{CP}$ what$_k$ $C^0\ [_{TP}$ you asked $[_{CP}$ how$_i$ $t'_k\ C^0\ [_{TP}$ John $[_{VP}$ fixed $t_k]\ t_i]]]]]$[7]

- This operation is "counter-cyclic" – we can't go back to an embedded clause and move things up after having performed operations at a higher clause. For this reason, there have been stipulations that operations must all first target the most embedded clause, then the next clause up, and so on. With the Extension Condition, this is no longer necessary.

- In Minimalism, we need some way to know when the derivation is "done". D-Structure provides us with all the lexical items that will enter the derivation, but without D-Structure, we need something else to do that work.

- Chomsky (1995) proposes the **numeration** – an unordered set of lexical items plus an index, where the index is the number of times it will be Merged. Each application of Merge reduces the index by 1. When the entire numeration is exhausted, the derivation is "done". Numerations will also be how we define classes of derivations that we compute Economy for – i.e., for a derivation N, we want the derivation that exhausts N using the simplest/smallest operations possible.

[6]The Extension Condition is derived by Chomsky's (2000) "No Tampering Condition", which says that Merge(X,Y) cannot change the internal structure of X or Y. This assumes that Move is actually an instance of the operation Merge, and that traces are not left under movement. We'll work up to this.

[7]This may be out because there are two *wh*-phrases in the embedded Spec,CP. However, there are many proposals that intermediate traces of arguments may be deleted, including Lasnik & Saito (1992).

(49)

Merge(see,John)	$[_{V'}$ see John$]$
N = $\{C^0_1$,[Pst]$_1$, Mary$_1$, see$_0$, John$_0\}$	
Merge(Mary,V′)	$[_{VP}$ Mary $[_{V'}$ see John$]]$
N = $\{C^0_1$,[Pst]$_1$, Mary$_0$, see$_0$, John$_0\}$	
Merge([Pst], VP)	$[_{T'}$ [Pst] $[_{VP}$ Mary $[_{V'}$ see John$]]]$
N = $\{C^0_1$,[Pst]$_0$, Mary$_0$, see$_0$, John$_0\}$	
Move	$[_{TP}$ John$_i$ $[_{T'}$ [Pst] $[_{VP}$ Mary $[_{V'}$ see $t_i]]]]$
N = $\{C^0_0$,[Pst]$_0$, Mary$_0$, see$_0$, John$_0\}$	
Merge(C^0,TP)	$[_{CP}$ C^0 $[_{TP}$ John$_i$ $[_{T'}$ [Pst] $[_{VP}$ Mary $[_{V'}$ see $t_i]]]]]$

- We will posit that there may some internal structure to Merge called *phases* later in the semester. Numerations are not important objects in modern theorizing, however, the phase theory is quite important in modern Minimalism, and for that reason, we will discuss it in more detail.

Bibliography

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

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In R. Martin, D. Michaels, & J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honour of Howard Lasnik*, 89–155. Cambridge: MIT Press.

Lasnik, Howard, Mamoru Saito. 1992. *Move α*. Cambridge: MIT Press.

Hornstein, Norbert, Kleanthes Grohmann, and Jairo Nunes. 2006. *Understanding Minimalism*. Cambridge, MA: Cambridge University Press.