Nicholas LaCara · University of Toronto lin331 · 31 July 2018 (Updated 1 August)

The binding principles

1 Binding Theory

- The binding theory accounts for the distribution of reflexives and reciprocals, bound pronouns, and referring expressions (R-expressions).
- Let's begin with some standard definitions.

Under GB, binding domains were defined with regard to Government. We no longer have recourse to such a relation, so we should wonder about the definition in (2).

(1) *Binding Theory*:
   a. *Principle A*:
   An anaphor must be bound in its domain.
   b. *Principle B*:
   A pronoun must be free in its domain.
   c. *Principle C*:
   An R-expression must be free.

(2) *Domain*:
α is the domain for β iff α is the smallest TP containing β and the governor of β.

α can also be the smallest DP containing β and the governor of β. This occurs with possessive DPs, like *Mary's picture of herself*.

(3) *Binding*:
α binds β iff α c-commands and is coindexed with β.

- Each principle accounts for a different kind of element.
- Principle A determines the distribution of reflexives, which must be bound locally.

(4) a. *Mary$_i$ said that [Joe$_k$ liked these pictures of herself$_i$].
   b. Mary$_i$ said that [Joe$_k$ liked these pictures of himself$_k$].

- Principle B precludes pronouns from being locally bound.

Pronouns embedded in PPs do not always cause a strong violation of Principle B.

(5) a. Mary$_i$ said that [Joe$_k$ liked these pictures of her$_i$].
   b. *Mary$_i$ said that [Joe$_k$ liked these pictures of him$_k$].

- Principle C forbids R-expressions from being c-commanded by a conindexed element:

We usually think of these as names, but definite descriptions (such as epithets) are subject to Principle C as well: *Harvey$_i$ says Sally hates the bastard$_{*i}$*.

(6) *She$_i$ said that [Joe$_k$ liked these pictures of Mary$_i$].

- These apply straightforwardly where no A′-movement is involved.
- From a GB point of view, as far as the c-command relations of the elements above are concerned, the binding principles could hold at DS, SS, or LF!

- From the minimalist point of view we should want to say that they hold at LF, since this is the only conceptually motivated level of representation of the three.
- There will be issues with this, of course.
  - Things look good if you restrict your view to Principle A.
  - Adding Principles B and C to the mix creates a number of complications that force us to look specifically at at the syntax of reflexive binding.
  - We will have to introduce new economy constraints to the mix and revise our view of where binding actually holds.

2 Complications from movement

- Movement significantly complicates the view of how the binding theory works.
- We have already seen some of these complications.

See part 3 of Assignment 1.

(7) *John$_i$ wondered which woman liked which pictures of himself$_i$

- If we assume covert *wh*-movement of the full *wh*-phrase *which pictures of himself*, we expect that *John* should be able to bind *himself* in this example.

(8) John$_i$ wondered [[which picture of himself$_i$]$_k$ + [which woman]$_j$ [t_j liked t_k]]

- The proposal we saw was that covert movement need only move the *wh*-word, not the whole *wh*-phrase.
- This prevents *himself* from moving into a position where it could be bound:

(9) *John$_i$ wondered [which$_k$ + [which woman]$_j$ [t_j liked [t_k picture of himself$_i$]]]

- But this solution cannot explain all of the reflexive binding patterns we observe.
- In (10), it appears as though we want to make reference to the base position of *himself* in order to explain how it is bound by *Fred*.
- This is simply done if we assume a level like D-structure:

(10) John$_i$ wondered which picture of himself$_{i/k}$ Fred$_k$ liked.

a. *DS*:
John wondered [$_{CP}$ [$_{TP}$ Fred$_k$ liked [which picture of himself$_k$]].

b. *LF*:
John$_i$ wondered [$_{CP}$ [which picture of himself$_i$]$_m$ [$_{TP}$ Fred liked t_m].

This could also be SS, but since we need LF anyway, lets assume this configuration holds at LF.

- We've seen no independent evidence for DS, however, and have so far assumed that it does not exist.

- To avoid reintroducing DS, we could try assuming that the binding principles apply throughout the course of the derivation rather than at specific levels.

So, as long as the right configuration is met at some point, all will be well.

- But there are empirical problems with such an approach. The reciprocal *each other* in (11) should be able to be bound after *wh*-movement under this view.

(11) The students$_i$ asked [what attitudes about about each other$_{k/*i}$]$_m$ the teachers$_k$ had t_m.

- Thus, it cannot be the case that the binding principles apply at all times.

3 Principle A under the copy theory

- Once we adopt the view that traces are copies of moved elements, we may find a way out of the above conundrum.

We need to do something, anyway, since we don't want superfluous LF copies to introduce Binding Theory violations of their own.

  - Similar to the way PF handles multiple copies, we must reduce the chains that occur at LF so that they behave similar to traces.
  - The solution to the problems above will be to selectively delete parts of the copies at LF.

- Movement leaves behind several copies at LF, and movement of material including an anaphor will create several copies of the anaphor.
- Thus, a more accurate Spell Out representation of (10) is as follows:

(12) John wondered [$_{CP}$ [which picture of himself] [Fred liked [which picture of himself]].

- This means at LF, each copy of *himself* is in an appropriate binding domain, consistent with Principle A.

Note that this requires us to move the whole *wh*-phrase, not just the *wh*-word as proposed above. We will come back to this.

  - The lower copy is bound by *Fred*.
  - The higher copy is bound by *John*.

- So we have the right configuration for each of the possible readings. But how do we reduce the structure to just what we need?

3.1 Operators, variables, and deletion

- Critically, (10) is ambiguous between only two readings: *himself* is bound either by *Fred* or by *John*, but not by both.
  - From this fact, it stands to reason that only one copy is interpreted at LF.

- To interpret (10) correctly, we must convert it into an operator–variable format and get rid of one of the copies of *himself*.
  - A common way of understanding *wh*-movement in GB is that *wh*-elements in SpecCP are quantificational operators that bind variables in the form of *wh*-traces.
  - Since we no longer have traces, we must convert copies to variables at LF.
  - Additionally, we must ensure that there is no repeated material at LF while creating a valid operator–variable structure.

See Fox 1999 for the details on how this operation would work. The discussion below is a sketch. Another approach I really like is Vicente 2009.

- The simplest way to do this is simply to delete the lowest copy at LF, resulting in the interpretation in (14).

~~Double strikeout~~ represents LF deletion.

(13) John wondered [$_{CP}$ [which picture of himself] [Fred liked ~~[which picture of himself]~~]].

Here, *John* binds *himself*.

(14) John wondered which x, x a picture of himself, Fred liked x.

- The binder *which x* is restricted by the material *a picture of himself*.
- However, an alternative exists. We can delete all but the *wh*-element in the higher copy, and then delete the *wh*-word in the lower copy.

Presumably, because LF does not have to worry about issues of linearization, the deletion operation behaves differently than the one at PF.

- The result is the interpretation in (16).

(15) John wondered [$_{CP}$ [which ~~picture of himself~~] [Fred liked [~~which~~ picture of himself]]].

(16) John wondered which x Fred liked x, x a picture of himself.

- Here, the variable is restricted rather than the binder.
- Notice that (15) is very similar to the configuration used to explain (10a) above.
  - The difference here is that we did not move the *wh*-word alone, but selectively deleted material in different copies of the the *wh*-phrase.

But also notice that the explanation for why (10) is ungrammatical requires that it not be possible to interpret the full copy in the higher position, as in (13). We will have to look for an explanation of this fact.

- Indeed, evidence from overt *wh*-movement suggests that this configuration must be interpretable at LF, since it can be generated overtly in some languages.

(17) *French*:

a. [Combien de livres]$_i$ a-t-il consultés t_i?
how.many of book has-he consulted

b. Combien$_i$ a-t-il consultés [t_i de livres]?
how.many has-he consulted of book

(18) *German*:

a. [Was für Bücher]$_i$ hast du t_i gelesen?
what for books have you read

b. Was hast du [t_i für Bücher] gelesen?
what have you for books read

- Thus, if we assume a copy-and-delete approach for LF, we can account for the ambiguity of (10).

3.2 How is this better than what we had before?

- This analysis relies on two innovations: The Copy Theory of Movement, and the deletion of superfluous copies at LF.
  - We already have independent motivation for the Copy Theory.
  - But what about LF deletion?
- We haven't seen any independent motivation for LF deletion. The motivation is completely theory-internal.
- However, even GB/Trace Theory needs some mechanism for dealing with cases such as the following:

(19) Whose goat did you see?
  a. *SS*:
  [$_{CP}$ [whose goat]$_i$ did [you see t_i]]?
  b. *LF*:
  [$_{CP}$ whose$_j$ did [$_{TP}$ you see [t_j goat]]]?

- The question here is not ranging over goats but over people x such that you saw x's goat.

That is, this is a question about people, not goats.

- On the GB approach, one must RECONSTRUCT *goat* into its base position to get this interpretation, literally putting *goat* back into its base position.
- The copy-and-delete approach yields the same results without reconstruction.

(20) a. *LF (CTM)*:
  [$_{CP}$ [whose ~~goat~~] did [you see [~~whose~~ goat]]]?
  b. Who x did you see x's goat?

- Thus, both Trace Theory and the Copy Theory need to do something about this, and deletion seems no worse that reconstruction.

The state of the art view is that copy-and-delete is how reconstruction actually works, just like movement.

3.3 Extension to A-movement

- Though we've so far only looked at *wh*-movement, similar facts can be seen under A-movement.
- As (21a) shows, *them* apparently induces Principle C effects relative to *John and Mary* suggesting it c-commands the subject of the lower clause.

It is thought that *to* is not a real preposition but some sort of Case marker that is part of the DP.

(21) a. *It seems to them$_i$ that [[John and Mary]$_i$ were angry].
  b. [John and Mary]$_i$ seem to each other$_i$ [t_i to be angry].

- To avoid a Principle C violation at LF, it is necessary to assume that the base copy of *John and Mary* is deleted in (22):

We'll come back to the details of Principle C below.

(22) [John and Mary] seem to each other [~~[John and Mary]~~ to be angry].

- This allows *John and Mary* to bind the reciprocal without the reciprocal inducing a Principle C violation by c-commanding the base copy of *John and Mary*.

4 We're not done yet

- The solution to the the ambiguous binding of reflexives above involves partially deleting copies at LF in order to satisfy Principle A in different positions.
- However, we must now consider how this works for Principles B and C.
- Given what we saw above, why can't *Fred* and *him* be coindexed?

(23) John$_i$ wondered which picture of him$_{i/*k}$ Fred$_k$ liked.

- It is fairly easy to get a configuration consistent with Principle B, where *John* and *him* can be coindexed, as in (32a).
- However, nothing so far prevents deletion of the entire low copy, as in (32b), giving rise to the unwanted interpretation.

Notice that *him* is not c-commanded by *Fred* here, so should be able to be coindexed with *Fred*.

(24) John wondered [[which ~~picture of him~~] [Fred liked [~~which~~ picture of him]]].

(25) *John wondered [[which picture of him] [Fred liked [~~which picture of him~~]]].

- The same problem arises for Principle C. Here, neither pronoun can be conindexed with *John*.

(26) He$^1_{i/*j}$ wondered which picture of John$_j$ he$^2_{i/k/*j}$ liked.

John does not c-command *he*^{2} here, so they should be able to be coindexed.

- Again, interpreting the lower copy of *John* leads to the right result, as in (27).
- However, interpreting the higher copy, as in (28) should allow the lower instance of *he* to be coindexed with *John*, contrary to fact.

(27) He wondered [[which ~~picture of John~~] [he liked [~~which~~ picture of John]]].

(28) *He wondered [[which picture of John] [he liked [~~which picture of John~~]]].

5 The Preference Principle

- To account for ambiguities introduced by the interaction of *wh*-movement with Principle A, we have introduced the idea that (subparts of) copies could be deleted at LF to give rise to the correct configurations to satisfy Principle A.

(29) John wondered [$_{CP}$ [which picture of himself] [Fred liked ~~[which picture of himself]~~]].

(30) John wondered [$_{CP}$ [which ~~picture of himself~~] [Fred liked [~~which~~ picture of himself]].

- Critical for this is the assumption that these copies are related by a chain. Chains are reduced by deleting subparts of copies in a chain at LF.

Deletion is usually thought of as a form of Chain reduction really relies on the notion that copying forms chains. Chain reduction is effected by copy deletion of *elements in the same chain*.

- This worked well for Principle A, but it would appear to overgenerate when trying to account for Principles B and C.
- The proposed solution, due to Chomsky (1993), is the Preference Principle:

(31) *Preference Principle*:
Try to minimize the restriction in the operator position.

- What this means is that material other than the *wh*-word should be interpreted in a lower position (if possible).
- This will give us what we want for Principles B and C, but it should also block (29). Is this what we want?
  - First, let's look at how (31) blocks overgeneration with Principles B and C.
  - After this, let's return to (29). We will need to appeal to some specific assumptions about reflexive binding to get the analysis straight.

5.1 Principle B

- In example (32), we want to rule out the interpretation derived from (32b) while keeping the one from (32a).

Notice that *him* is not c-commanded by *Fred* here, so should be able to be coindexed with *Fred*.

(32) John$_i$ wondered which picture of him$_{i/*k}$ Fred$_k$ liked.
  a. John wondered [[which ~~picture of him~~] [Fred liked [~~which~~ picture of him]]].
  b. *John wondered [[which picture of him] [Fred liked [~~which picture of him~~]]].

- Nothing about the deletion operation itself rules out deletion of the entire low copy, as in (32b). This gives rise to the unwanted interpretation.
- However, (31) does give us the right result.
  - The unwanted interpretation is also the one where the *wh*-operator retains a restrictor at LF.
  - If we assume that (31) holds, example (32b) will be ruled out.
  - This leaves (32a), where *John* binds *him* from outside its domain.
- Thus, the Preference Principle achieves the right results for Principle B.

5.2 Principle C

- A similar problem arises in (33).
- Interpreting the lower copy of *John* leads to a Principle C violation, but interpreting the higher copy would allow the lower instance of *he* to be coindexed with *John*, contrary to fact:

(33) $\text{He}_{i/*j}$ wondered which picture of John_j $\text{he}_{i/k/*j}$ liked.

a. He wondered [[which ~~picture of John~~] [he liked [~~which~~ picture of John]]].

b. *He wondered [[which picture of John] [he liked [~~which picture of John~~]]].

- Again, deleting the whole lower copy would contravene (31), so (33a) should be favored over (28).
- So, again, the Preference Principle achieves the right results, ruling out the configuration where *John* c-command *he* at LF.

5.3 Motivation and Principle A (again)

- If the Preference Principle can explain Principles B and C, why does Principle A get an apparent pass?
- The idea is that since Principle A deals directly with anaphors, it must have something specifically to do with the syntax of these elements.
- The solution is to posit that reflexives must move to a position where they can be licensed by their antecedents.
  - Assume that anaphors can only be bound if they agree with their antecedents.
  - The mechanism we have for agreement is checking, which requires movement prior to LF.
  - Thus, if reflexives agree with their antecedents, they must move to a position where they can do so by LF.

This seems reasonable for reflexives, but what about reciprocals? They do not seem to agree.

- Assume for instance that we want the *Fred* to bind *himself*. In this case, *himself* will move to be in a local relation with *Fred*.

(34) *Spell Out*:
John wondered [$_{\text{CP}}$ [which picture of himself] [Fred liked [which picture of himself]].

(35) *LF w/ reflexive movement*:
John wondered [$_{\text{CP}}$ [which picture of himself] [Fred+himself liked [which picture of himself]].

- If this is right, we still need to make sure that we reduce the number of copies of *himself* to just one at LF.

- We must reduce both the *wh*-movement chain and the anaphor movement chain. In other words, there are two steps here:

  (36) Wh-*chain reduction*:
  John wondered [$_{CP}$ [which ~~picture of himself~~] [Fred+himself liked [~~which~~ picture of himself]].

  (37) *Anaphor chain reduction*:
  John wondered [$_{CP}$ [which ~~picture of himself~~] [Fred+himself liked [~~which~~ picture of ~~himself~~]].

A-movement requires reduction, the same as A′-movement.

Because of this step, there is only one copy of *himself*, the one local to its binder *Fred*.

- Now consider what happens if we do reflexive movement for the higher copy of *himself*.
- If we attempt to do chain reduction in keeping with the Preference Principle, things go bad:

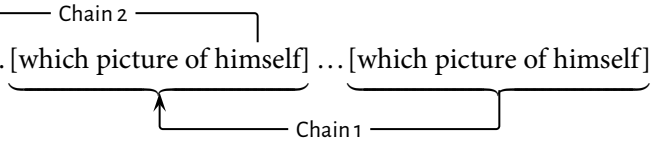
  (38) *LF w/ reflexive movement*:
  John+himself wondered [$_{CP}$ [which picture of himself] [Fred liked [which picture of himself]].

  (39) Wh-*chain reduction*:
  John+himself wondered [$_{CP}$ [which ~~picture of himself~~] [Fred liked [~~which~~ picture of himself]].

- The main issue with (39) is that we wind up with two copies of *himself* at LF.
  - Reducing the higher copy of the *wh*-element in the way we have been so far will delete the base of the reflexive movement chain.
  - The remaining copies of *himself* do not form a chain, though. The copy local to *John* is not part of the *wh*-movement chain.
  - Since chain reduction requires reducing copies *in the same chain*, there is no way to eliminate both copies of *himself* here in keeping with the Preference Principle while maintaining a copy of the reflexive local to *John*.

Additionally, chain reduction of the *wh*-chain as in (39) independently takes care of reduction of the reflexive chain.

(40) John+himself … [which picture of himself] … [which picture of himself]



- The only way to get an acceptable result is to violate the Preference Principle:

  (41) Wh-*chain reduction (Violates (31))*:
  John+himself wondered [$_{CP}$ [which picture of himself] [Fred liked [~~which picture of himself~~]].

  (42) *Reflexive chain reduction*:
  John+himself wondered [$_{CP}$ [which picture of ~~himself~~] [Fred liked [~~which picture of himself~~]].

- This is why the Preference Principle is stated as a preference.

It has the feel of an Optimality Theory constraint (Prince and Smolensky 1993/2004): Only violate this if you need to satisfy more highly ranked constraints.

  - If (31) were an absolute, we would expect (29) to be ruled out.
  - We need to state it as a preference in order to allow (41).

- This relies specifically on the proposed syntax for reflexives, explaining why Principle A gets around the Preference Principle, but not Principles B and C.

  - Without this movement, Principles B and C are subject to the Preference Principle.

Furthermore, the reflexive movement analysis has its antecedents in GB. Here we are adopting the idea into our Minimalist theory of the Binding Theory, so it does not decide between which approach is better.

5.4 Economy and indices

- The Preference Principle is a kind of economy condition.
- Economy conditions are thought to decide only between CONVERGENT derivations (*i.e.*, those that do not crash at the interfaces).
- But if it only chooses between convergent derivations, why does the Preference Principle choose (43a)?

  - There's no grammatical coindexation between *Fred* and *him* in (43a).
  - Example (43b) allows for grammatical coindexation between these elements.
  - So shouldn't the Preference Principle choose (43b) over (43a) in order to allow grammatical coindexation between these two elements?

[I'VE TRIED TO REVISE THIS SECTION A BIT, SINCE I REALIZED I DIDN'T EXPLAIN IT CLEARLY (OR TOTALLY CORRECTLY) DURING THE LECTURE. I HOPE THIS CLEARS UP SOME OF THE CONFUSION —NL]

(43) Mary wondered which picture of $\text{him}_{i/*k}$ Fred_k liked.

  a. Mary wondered [[which ~~picture of him~~] [Fred liked [~~which~~ picture of him]]].
  b. Mary wondered [[which picture of him] [Fred liked [~~which picture of him~~]]].

- However, since coindexation between these two elements is impossible, we know that the Preference Principle *does not* select (43b).

  - So it must be the case that (43a) converges, since otherwise we would expect the Preference Principle to (trivially) select (43b)
  - If (43a) did not converge, the Preference Principle would not consider it.

- So why is (43a) OK? The answer seems to be that coindexation does not factor into determining whether a derivation converges or not.

In other words, indices are not syntactically present.

- Chomsky (1993) suggests that satisfaction of the Inclusiveness Condition requires that indices not be added to syntactic objects during the course of the derivation.
- If indices are not introduced in the course of a derivation, then the problem with (43a) is only an apparent problem.

  - Both of the derivations are, in fact, convergent. Since there are no indices, there is no way to compare derivations with alternative indexations.

  - Since both (43a) and (43b) are convergent, the Preference Principle will choose (43a).

- But this means that we need to reconceptualize the Binding Theory without making reference to coindexation, since indices are inconsistent with the Inclusiveness Condition!

- Fortunately, this isn't too hard to do. We can switch from speaking of *coindexation* to *coreference*:

See Chomsky and Lasnik (1993).

(44) *Binding Theory*:
  a. *Principle A*:
  If α is an anaphor, interpret it as coreferential with a c-commanding phrase in its domain.
  b. *Principle B*:
  If α is a pronoun, interpret it as disjoint from every c-commanding phrase in its domain.
  c. *Principle C*:
  If α is an R-expression, interpret it as disjoint from every c-commanding phrase.

- This is consistent with Inclusiveness and captures the facts we've seen so far.
- It means that binding doesn't hold at LF, but rather at the C–I interface!

This pushes part of the job out of the syntax and into the interpretational component of the grammar. While the syntax creates the configurations for binding, binding itself ceases to be a syntactic phenomenon.

6 Adjunction and Late Merge

- Binding into complement clauses and adjunct clauses shows different behavior:

(45) *Complement clause*:
  *Which claim [that John$_i$ was asleep] did he$_i$ discuss?

(46) *Relative clause*:
  Which claim [that John$_i$ made] did he$_i$ discuss?

- The Preference Principle, however, predicts that these two cases should behave similarly.
- In both cases, *John* should be interpreted below *he*, leading to a Principle C violation:

(47) [[Which ~~claim [that John was asleep]~~] did he discuss [~~which~~ claim [that John was asleep]]]?

(48) [[Which claim [that John made]] did he discuss [~~which~~ claim [that John made]]]?

- So, at first glance, these examples should not behave differently with respect to Principle C.

- Here, we have to employ a trick, based on work in GB by Lebeaux (1988). The idea is that adjuncts can merge later than arguments.
- Recall that the Extension Condition requires that Merge target the root:

(49) *Extension Condition*:
Overt applications of Merge can only target root syntactic objects.

Now that movement is Copy + Merge, we need not distinguish move in this definition.

- The structure in (47) is thus built by repeated applications of Merge, always targeting the root.

(50) a. [$_{CP}$ that John was asleep]
b. [$_{NP}$ claim that John was asleep]
c. [$_{DP}$ which claim that John was asleep]
d. [$_{VP}$ discuss [$_{DP}$ which claim that John was asleep]]
⋮
e. [$_{CP}$ [$_{DP}$ which claim that John was asleep] did he discuss [$_{DP}$ which claim that John was asleep]]

- If adjuncts can undergo so-called Late Merge, then they are not subject to the Extension Condition.

(51) a. [$_{CP_1}$ Op_i that John made t_i]
[$_{CP_2}$ did he discuss [$_{DP}$ which claim]]
b. [$_{CP_1}$ Op_i that John made t_i]
[$_{CP_2}$ [which claim] did he discuss [$_{DP}$ which claim]]
c. [$_{CP_2}$ [[which claim][$_{CP_1}$ Op_i that John made t_i]] did he discuss [$_{DP}$ which claim]]

- If we do this, then it will not be possible to reconstruct *John* into a position below the pronoun.
- We are borrowing technology from GB again, but we still need to refer to any levels of representation beyond LF.

Note that within GB, Lebeaux's proposal seems to undermine the entire concept of DS, since not all of the material in the derivation need be present at that level.

6.1 Empirical problems

- Given that adjunction has been problematic before, we might just say that the exceptionality of adjunction with respect to the Extension Condition is just how adjunction works.
- But this is just a stipulation; we should want this fact to fall out from something else, if it is even true.
- Consider, for example, the impossibility of extraction out of an adjunct:

(52) *Which book$_i$ did you talk to Sally [before buying t_i]?

- Let's assume for the moment that this is the result of the copying operation being unable to target material in adjuncts.
- The problem is that if adjuncts like *before buying* don't merge until later in the derivation, it should be possible to copy *which book* out of the adjunct *before* it becomes an adjunct.

Remember: Under our current systems, being an adjunct is determined relationally when the element merges.

- Presumably, we would build the adjunct separately from the rest of the clause before we merge it:

(53) a. K = [$_{PP}$ before PRO buying [which book]]
b. L = [did you [$_{VP}$ talk to Sally]]

- At this point, the operation Copy should still be able to copy *which book* because the PP has not yet become an adjunct.

And it won't become an adjunct until it merges.

(54) a. K = [$_{PP}$ before PRO buying [which book]]
b. L = [did you [$_{VP}$ talk to Sally]]
c. M = [which book]

- Now that we have copied *which book*, we should be able to merge M and L:

(55) a. K = [$_{PP}$ before PRO buying [which book]]
b. N = [[which book] did you [$_{VP}$ talk to Sally]]

- And now we late merge the adjunct PP and reduce the chain:

(56) [[which book] did you [$_{VP}$ [$_{VP}$ talk to Sally] [$_{PP}$ before PRO buying ~~[which book]~~]]]

- Nothing so far prevents this, but it allows us to derive the ungrammatical (52).
  - And if we take seriously the idea that adjuncts can merge late, it's not clear what could prevent it.

6.2 A sideward solution

- These cases of so-called SIDEWARD MOVEMENT are a real problem for the Late Merge hypothesis.

See Nunes 2004.

- However, if we take sideward movement seriously and restore the Extension Condition, it turns out that things still work out alright.
- Let us reconsider (46), repeated here:

(57) Which claim [that John$_i$ made] did he$_i$ discuss?

- Again, assume that the adjunct and the main clause are built in parallel.

There is nothing weird about this if we assume the Extension Condition.

(58) a. K = [did he discuss [which claim]]
b. L = [Op_k that John made t_k]

- As above, we can copy *which claim*:

(59) a. K = [did he discuss [which claim]]
b. L = [Op_k that John made t_k]
c. M = [which claim]

- But now we can merge L and M, merging the relative clause with the *wh*-phrase:

(60) a. K = [did he discuss [which claim]]
b. N = [[which claim] [Op_k that John made t_k]]

- The resulting element can now merge with K. Overall, this has the same effect as late merger of the relative clause, but the whole derivation obeys the Extension Condition:

(61) [[which claim] [Op_k that John made t_k]] [did he discuss ~~[which claim]~~]

- Importantly, this cannot derive (45), repeated here:

(62) *Which claim [that John$_i$ was asleep] did he$_i$ discuss?

- The problem here is that the complement clause is the complement of *claim*. If we wait to copy *which claim* and then try to merge the complement clause with the *wh*-phrase, we will not wind up with the correct structure.

(63) *Build structures*:
a. K = [did he discuss [which claim]]
b. L = [$_{CP}$ that John was asleep]

(64) *Copy* wh-*element*:
a. K = [did he discuss [which claim]]
b. L = [$_{CP}$ that John was asleep]
c. M = [which claim]

(65) *Merging L and M generates wrong structure*:
a. K = [did he discuss [which claim]]
b. N = [[which claim] [$_{CP}$ that John was asleep]]

We want: [which claim [that John was asleep]]

- Merging the CP with *which claim* as a complement would violate the Extension Condition.
  - The only alternative is to assume that the complement clause merges with *claim* as part of the main derivation.

References

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 1–52. MIT Press.

Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An International Handbook of Contemporary Research*, ed. J. Jacobs, A. von Stechow, Wolfgang Sternefeld, and T. Vennemann, 506–69. Berlin: Walter de Gruyter.

Fox, Danny. 1999. *Economy and Semantic Interpretation*. Number 35 in Linguistic Inquiry Monograph. Cambridge, Mass.: The MIT Press.

Lebeaux, D. 1988. Language acquisition and the form of the grammar. Doctoral Dissertation, University of Massachusetts Amherst.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Number 43 in Linguistic Inquiry Monographs. Cambridge, Mass: MIT Press.

Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, Mass, and Oxford, UK: Backwell.

Vicente, Luis. 2009. A Note on the Copy vs. Multidominance Theories of Movement. *Catalan Journal of Linguistics* 8:75–97.