Of Neglect[‡]

Shrayana Haldar shaldar@mit.edu
Massachusetts Institute of Technology

April 18, 2024
Since November 2, 2021

Abstract

In this paper, I argue that PF and LF Neglect, in the sense of Sportiche (2016), are not independent of each other. Furthermore, independent LF Neglect has properties that are also underivable from Sportiche (2016). I argue this by recapitulating an observation from Fox (2017), and also based on data from Sauerland (1998). Moreover, while partial Neglect is stipulative, total Neglect suffers from its own undergeneration problems that are not alleviated even when supplemented with Fox's (2017) multidominant grammar. I conclude, therefore, that Neglect only stipulates and doesn't have explanatory power.

Keywords: Neglect, Late Merge, reconstruction, extraposition, modularity, multidominance, linearization, presupposition projection.

1 Introduction

This paper is a reflection on the operation dubbed *Neglect* in Sportiche (2016). Neglect is an interface operation — applicable at both LF and PF — that "neglects" copies, or, in other words, makes them fail to be interpreted by the interfaces. This is not obviously different from the deletion of a copy at an interface, as envisioned in Chomsky (1993). The concept can also be found under the name *distributed deletion* in Fanselow and Ćavar (2002) and under the name *scattered deletion* in Bošković (2015). The purpose of this paper is to argue that the kind of approach that is known as *Neglect* doesn't have explanatory power. To that end, in this introduction, I will summarize why Neglect is needed and how Neglect works, pivoting on the notion of reconstruction feeding Binding Theoretic considerations. Once this is done, the remainder of the paper will be devoted to arguing why I believe this approach doesn't work. Therefore, let's begin with the following statement of Condition A of the Binding Theory in (1).

(1) **Condition A, version 1**
The structurally highest copy of an anaphor must be bound in its binding domain.

‡. I'm grateful to Danny Fox, Kyle Johnson, Amir Anvari, and Zachary Satoshi Feldcamp for their comments and the discussions I had with them on this material and several other previous versions of this. All errors are mine alone.

This statement clearly seems to be wrong, given examples like (2) because the higher copy *which side of itself* c-commands *the sofa* and therefore, there's no way for *itself* to get bound by *the sofa* if there's an occurrence of *itself* in a position that c-commands it.[1]

(2) Which side of itself$_1$ did the sofa$_1$ fall on?

Given this, we could remedy the situation by jettisoning (1) and stating Condition A differently. Such a version of Condition A can be stated as in (3).

(3) **Condition A, version 2**
At least one occurrence of an NP containing an anaphor must be bound in its binding domain.

Let's point out at the very outset that even with this version of the Condition A, semantic composition itself overgenerates. With Condition A, version 2, when (4), below, is evaluated at LF for the purposes of Binding Theory, Condition A should be satisfied. What goes wrong is the interpretation of the higher copy. When this higher copy of the *wh*-chain, containing a free variable *itself*, is interpreted under a standard Heim and Kratzer (1998) semantics, the index of *itself* would be mapped to what the assignment function maps this index to in the absence of λ-abstraction, because λ-abstraction makes the interpretation of otherwise assignment-dependent bound variables assignment-independent. Therefore, the index of *itself* in the higher copy could be mapped, perhaps, to a book that's salient. So, assuming that the assignment function is g, Condition A, version 2 predicts that (4) should be able to mean something like "Which side of the book that is $g(1)$ is such that every sofa, x, fell on the side of x?" This is not an attested reading.

(4) Which side of itself$_1$ did every sofa$_1$ fall on?

Even if we are charitable and hope this can be ruled out by an independent assumption, we know, from Engdahl (1986), Heim (2019), that *wh*-questions like (4) have "functional" readings, *i.e.*, "Which f_{ee} is such that f is a way of mapping sofas to their sides and every sofa, x, fell on $f(x)$?" A possible answer to a question with a reading like this can be exemplified by, "Its front side". This is all good, and we know, from the literature cited above, of some ways to derive these functional readings. The problem is that, when we combine Condition A, version 2 with the possibility of functional readings, we run into problems accounting for minimal pairs like (5). In (5a), both surface scope and inverse scope readings are possible. But in (5b), the inverse scope reading is unavailable. On the face of it, such a reading may not even be conceivable, but the following is a reasonable functional meaning one could want to express: "For every way f_{ee} of mapping students to aspects of those students, there exists a potentially different student x to whom I introduced $f(x)$. For instance, I introduced John to John's lack of self-esteem; I introduced Mary to Mary's childhood trauma, . . . " But this is not a reading available from (5b). Condition A,

1. Inanimate anaphors cannot be exempt anaphors. Exempt anaphors must be anteceded by logophoric centers, while plain anaphors don't. Plain anaphors are subject to the Binding Theory, while exempt anaphors are not (Charnavel 2020). Because of this, examples with an animate anaphor coreferent with an R-expression often lead speakers to parse the sentences as having logophoric centers. However, that's a parse I want to avoid, because I'm dealing with plain anaphora, not exempt anaphora. This is why I will systematically suppress examples where the possibility of exempt anaphora hasn't been eliminated, unless for the purposes of citing examples found in previous literature and of recapitulating/rehashing accounts of such examples in this paper. I would therefore caution the reader to take any judgements they might have about examples with coreference between an animate anaphor and an R-expression with a grain of salt. Also see Charnavel and Sportiche (2016).

version 2, in (3), would allow *picture of himself* to be bound only in its lower position, giving rise to such a reading.[2]

(5) a. ***Context:*** Spoken by an art history teacher.
I introduced a student to every aspect of the painting.
✓∃ > ∀
✓∀ > ∃

b. ***Context:*** Spoken by a therapist.
I introduced a (*different) student$_1$ to every aspect of himself$_1$.
✓∃ > ∀
*∀ > ∃

Given this problem, I conclude that Condition A, version 2 can't be the path to go down, and I'll pursue another option: to devise a way to make *side of itself* not be in the higher copy. Once we take this path, we're in a position where we know what structure we need for (2), which is given in (6), but we don't know how to derive it.[3]

(6) [which] did the sofa$_1$ fall on [side of itself$_1$]

Given this, we can take the following theoretical stock. The common wisdom, currently, is that if two objects are sisters at PF, then they are syntactic sisters, and, also, if two objects are sisters at LF, then they are syntactic sisters there as well. However, (2) shows that this cannot be the case, given (1). That is, it seems that, for (2), *which* and *side of itself* must be syntactic sisters at PF, and also, mustn't at LF.

One approach to this problem has been to say that, when movement happens in syntax, each copy produced by that movement operation is maximally articulate, in that each of them contains every part originally present in the base-generated copy. Therefore, according to this approach, syntax produces something like (7a). An approach like this, called *Neglect* and developed in Sportiche (2016), selectively "neglects" (parts of) copies at PF and LF, creating the effect of the neglected parts never having existed in the neglected positions at their respective interfaces. This is shown in (7b-c). In (7b), we see that PF neglects the lower occurrence of *which side of itself*, therefore, the higher occurrence of it is pronounced and the pronunciation of (2) is derived. In (7c), we see that LF neglects the higher occurrence of *side of itself* and preserves its lower occurrence, so Condition A can be satisfied.[4]

(7) a. ***Syntax:***
[which side of itself$_1$] did the sofa$_1$ fall on [which side of itself$_1$]

b. ***PF:***
[which side of itself$_1$] did the sofa$_1$ fall on ~~[which side of itself$_1$]~~

c. ***LF:***
[which ~~side of itself$_1$~~] did the sofa$_1$ fall on [~~which~~ side of itself$_1$]

2. It could be claimed that *every* and other non-*wh*-quantifiers can't give rise to functional readings. The first thing to say to that claim is: why?
3. This is very much along the lines of suggestions made in Heim (2019). I will refer the reader to that for more.
4. The neglect of *which* is entangled with the question of Trace Conversion (Fox 2002), among other things. I will come back to it later in the paper. For now, the lower occurrence of *which* has been shown to be struck through simply to encode the fact that *which* isn't interpreted downstairs, but upstairs.

This is how and why Neglect works: it neglects copies at interfaces so that they can interpret chains properly, given the independent needs of the two interfaces.

However, the more crucial question is: how is Neglect constrained? At the very least: how do we block Neglect operations like the ones in (8) (regardless of the interface)?

(8) ~~[which side of itself$_1$]~~ did the sofa$_1$ fall on ~~[which side of itself$_1$]~~

To answer this question, Sportiche proposes to block the Neglect of all copies of a chain from Chomsky's (1995) **Principle of Full Interpretation (FI)**. Chomsky informally states this as a ban against there being any superfluous symbols in a syntactic object. That is, all of the copies in a chain can't be deleted/neglected at any interface because if they could be, then that means that they never needed to be merged/generated via the Copy operation in the first place, thus violating the ban against superfluous symbols. Based on this, Sportiche formalizes Neglect as (9), while stating FI as in (10) (slightly modified from Sportiche's version). Here, *syntactic object* is to be understood not as individual occurrences of a chain, but as an entire chain itself. That is, it amounts to saying that at least one occurrence of each chain must be interpreted at each interface. The Neglect operations in (8) violate FI because none of the copies of the *wh*-chain is interpreted.

(9) **Neglect**
Any material at any interface can be ignored up to crash.

(10) **Principle of Full Interpretation (FI)**
Interpret every syntactic object at least once.

What FI does is put a constraint on *how much* of a chain can be neglected. That is, it blocks too much Neglect, so that at least one occurrence of all parts of the moving object is interpreted at both interfaces. **What FI doesn't do, however, is predict PF-LF correlations in Neglect and distinct patterns in Neglect at an interface,** both of which are attested in natural language. I observe such patterns below in the context of both partial PF and partial LF Neglect. In the end, I conclude that Neglect doesn't have explanatory power. I don't provide any way out in this paper. Work on a possible solution is underway.

The following is how the paper is structured. I begin by discussing the issues that exist for partial Neglect in section 2. In section 2.1, I recapitulate an argument from Fox (2017) that, when we consider what Neglect must systematically capture for facts about relative clause (henceforth, "RC") extraposition, there's a clear PF-LF modularity issue. That is, there are systematic correlations between PF Neglect and LF Neglect, making it seem that the two need to communicate in some way; and this is a problem for modularity. In section 2.2, I argue that, even when we ignore such problems of PF-LF correlations and just look at cases where no such correlation is empirically attested, there are monotonic patterns of reconstruction in *wh*-questions that arise independently for LF Neglect, while there's a constant, unchanging pattern of PF Neglect, blind to what's happening at LF. This, I argue, shows that partial LF Neglect is only stipulative and not explanatory. Given partial LF Neglect is thus inadequate for these cases, I consider whether total LF Neglect can derive these cases in section 3 and finally conclude that total LF Neglect, although able to rule out unattested patterns of reconstruction, also rules out the attested, monotonic patterns thereof. That is, it undergenerates. Therefore, I attempt to supplement total LF Neglect with a multidominant model of grammar developed in Fox (2017) that seems to be of help, at least initially, in that no partial Neglect operation is required. Once I lay down the basic system of

Fox (2017) in section 4, and address a semantic detail in section 5, I come back to our crucial cases in section 6 and consider various possible ways of exactly deriving the attested pattern. In section 6.1, I argue that most, if not all, attempts to derive these patterns, making only innocent assumptions, fail, and, in section 6.2, I argue that the only conceivable way we can derive the necessary structures requires doing something outlandish, resulting in unattested, non-trivial presuppositions for the sentences. In section 7, I conclude that Neglect doesn't have explanatory power, which can be appreciated only when we look at patterns of reconstruction across examples, and not just one single datum. I close the paper by alluding to ongoing work where I address such patterns of reconstruction within a system that has predictive power.

2 Partial Neglect

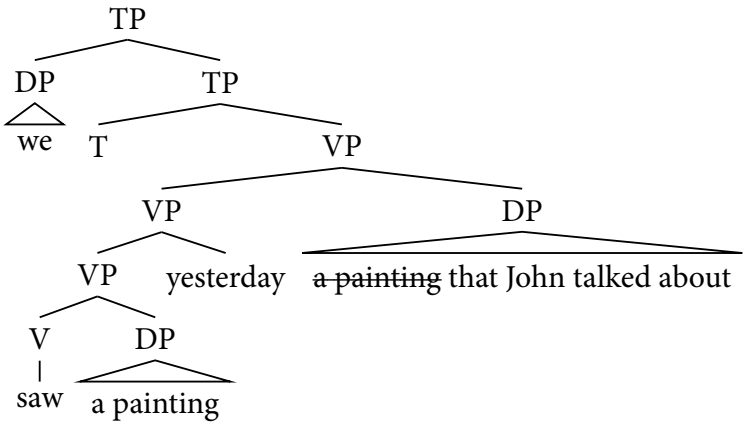
There are two kinds of phenomena I will be concerned with in this paper: RC extraposition and *wh*-movement. I will discuss RC extraposition in the context of a PF-LF modularity issue that arises for this phenomenon and *wh*-movement in the context of partial LF Neglect. For RC extraposition, I'll recapitulate an observation from Fox (2017) to the effect that there are systematic correlations between PF and LF partial Neglect that can't be captured without stipulations. For some *wh*-question data from Sauerland (1998), I'll argue that, even when there's no such correlation, there are monotonicity properties to LF partial Neglect which also can't be captured without stipulations.

2.1 Problem from a PF-LF Modularity Issue

I will begin by discussing an issue brought up in Fox (2017), which is based on observations made in Fox and Nissenbaum (1999) (henceforth, "F&N"), but I will elaborate fully on what he wants to say because he mentions the issue very briefly in that paper. According to F&N, we have the derivation in (11b) for a case of RC extraposition like (11a): the DP is QRed, the higher copy is deleted because the movement is covert, and then the modifier *that John talked about* is "late merged" (henceforth, "LMed"), that is, countercyclically merged to the restrictor of the higher copy.

(11) a. We saw a painting yesterday that John talked about.

b.



[Fox (2017), (6): 28]

F&N presents extensive evidence in favor of such an analysis, in turn providing evidence for what has now come to be known as the *single-component grammar* or the *single-output/single-cycle model* of syntax (Bobaljik 1995, 2002; Pesetsky 2000; Fox 2002; Bhatt and Pancheva 2004, *inter multa alia*). I reproduce (12) from them to instantiate the spirit of their diagnostics. What (12) specifically shows is that the RC is interpreted in a position higher than all the elements inside the VP because there's no disjoint reference effect between the pronoun and *John*.

(12) a. I gave him$_1$ a picture yesterday from John's$_1$ collection.
(*Cf.* ??/* I gave him$_1$ a picture from John's$_1$ collection yesterday.)

b. I gave him$_1$ an argument yesterday that supports John's$_1$ theory.
(*Cf.* ??/* I gave him$_1$ an argument that supports John's theory yesterday.)

c. I told you that he$_1$ will accept the argument when you and I last spoke that I presented to John$_1$ yesterday.
(*Cf.* * I told you when you and I last spoke that he$_1$ will accept the argument that I presented to John$_1$ yesterday.)

[F&N, (11): 8-9; indices mine]

The other part of the phenomenon we need to observe is that the determiner heading the DP containing the NP restrictor modified by the extraposed RC is also interpreted higher than all the elements in the VP. This is shown in (13). (13c-d) are control cases.

(13) ***"Free choice"* any *is licensed in the scope of the verb* look for.**

a. I looked very intensely for anything that would help me with my thesis.

b. * I looked for anything very intensely that will/would help me with my thesis.

c. I looked for something very intensely that will (likely) help me with my thesis.

d. I would buy anything without making a fuss that will/would help me with my thesis.

[F&N, (5): 5; indices mine]

Overfelt (2015b) has shown the same thing by exploiting the Strawson-downward-entailing (SDE) nature of the restrictor of *every*, which licenses NPIs. Because this has been known since Ladusaw (1979), Overfelt presents this as *Ladusaw's Generalization*, given in (14). The data corroborating this are given in (15).

(14) ***Ladusaw's Generalization***
An NPI is licensed by *every* only if that NPI is generated in the restrictor argument of *every*.

[Overfelt (2015b), (24): 157]

(15) a. Every [$_{NP}$ guest who ate *any* of the potato salad] [$_{VP}$ became ill].

b. * Every [$_{NP}$ guest who became ill] [$_{VP}$ ate *any* of the potato salad].

[Overfelt (2015b), (13): 154]

Given this, we can make sense of the RC extraposition data in (16) only if *every* is interpreted as high as the RC, where the NPI can be interpreted inside its restrictor position.

(16) a. We met $[_{DP}$ every biker$]_1$ yesterday
$[_{CP}$ who had *ever* ridden on these trails$]_1$.

b. The company considered $[_{NP}$ every applicant$]_1$ last month
$[_{CP}$ who was from *any* of the local temp agencies$]_1$.

[Overfelt (2015b), (26a), (27a): 158]

These effects are just special cases of what is dubbed *Williams's Generalization*, given in (17), in Fox (2002), based on observations in Williams (1974).

(17) ***Williams's Generalization (WG)***
When an adjunct β is extraposed from a "source DP" α, the scope of α is at least as high as the attachment site of β (the extraposition site).

[Fox (2002), (19): 71]

In (12), α corresponds to *a picture*, *an argument*, and *the argument*, and β corresponds to the RC being extraposed. In (13), α corresponds to *anything*, and β corresponds to the RC being extraposed. In none of the cases can α be interpreted inside the VP. If it could, then all the examples in (12) would be ungrammatical because of Condition C, and (13b) would be grammatical because the licensing requirement of free choice *any* would be met.

For the sake of completeness, also consider (18). This serves to show that the NP restrictor *book about John* in (18b) must be interpreted inside the VP. If it could somehow escape interpretation inside the VP, then the Condition C effect in (18b) wouldn't arise and the starred coindexation would be grammatical. (18a) is the baseline.

(18) a. I gave John$_1$ a book about himself$_1$ yesterday that Mary found.

b. I gave him$_{2/*1}$ a book about John$_1$ yesterday that Mary found.

Therefore, any account of Williams's Generalization must systematically capture (19).

(19) a. ***The NP Requirement***
The NP restrictor being modified by an extraposed RC must be interpreted semantically and phonologically inside the VP. (Recall (18): **I gave him$_1$ a book about John$_1$ yesterday that Mary found*.)

b. ***The Determiner Requirement***
When D^0 head of the host DP (*i.e.,* the DP from which the RC is extraposed) must be interpreted semantically at least as high as the RC but phonologically inside the VP. (Recall (13a)-(13b): *I looked very intensely for anything that would help me with my thesis* v. **I looked for anything very intensely that will/would help me with my thesis*.)

c. ***The RC Requirement***
When an RC is extraposed from inside a VP, it must be interpreted semantically and phonologically outside the c-command domain of everything inside the VP.

(Recall (12a): *I gave him$_1$ a picture yesterday from John's$_1$ collection* v. ??/* *I gave him$_1$ a picture from John's$_1$ collection yesterday*.)

Within a theory that espouses an operation such as Neglect, (20) is the pair of Neglect operations we would need at the two interfaces for a sentence like (11a).

(20) a. ***PF:***
. . . [~~a [painting~~ [$_{RC}$ that John talked about]]]
saw [a [painting ~~[$_{RC}$ that John talked about]~~]]

b. ***LF:***
. . . [a [~~painting~~ [$_{RC}$ that John talked about]]]
saw [~~a~~[5] [painting ~~[$_{RC}$ that John talked about]~~]]

The problem with this becomes apparent when we look at Heavy NP Shift (henceforth, "HNPS"). Under an account that assumes Neglect, HNPS will be derived by following PF and LF Neglect operations in (21).

(21) a. ***PF:***
. . . [a [painting [$_{RC}$ that John talked about]]]
saw ~~[a [painting [$_{RC}$ that John talked about]~~]]

b. ***LF:***
. . . ~~[a [painting [$_{RC}$ that John talked about]]]~~
saw [a [painting [$_{RC}$ that John talked about]]]

Such LF Neglect is motivated by HNPS facts like (22), which shows that that the material making the shifted nominal "heavy" enough for the purposes for HNPS can trigger disjoint reference effects with pronouns inside the VP. Crucially, compare this minimal pair with the minimal pair involving RC extraposition in (23), where R-expressions inside the extraposed RC don't trigger similar disjoint reference effects. Such cases are noted in F&N and Fox (2002) as well.

(22) a. I gave John$_{1/2}$ yesterday the book that he$_1$ said he$_1$ had been looking for.

b. I gave him$_{2/*1}$ yesterday the book that John$_1$ said he$_1$ had been looking for.

(23) a. I gave John$_{1/2}$ the book yesterday that he$_1$ said he$_1$ had been looking for.

b. I gave him$_{1/2}$ the book yesterday that John$_1$ said he$_1$ had been looking for.

Therefore, we can now combine our observation from RC extraposition and HNPS, and arrive at the observation in (24). These are implicational relationships between PF and LF Neglect operations that can't be captured without stipulation under Sportiche's (2016) account.

5. This *a*, which has been struck through, is not really partial LF Neglect, but a shorthand for Trace Conversion (Fox 2002), which replaces a quantificational determiner with a silent definite determiner. The same would apply for struck-through *which*es and other quantificational determiners in a lower copy in diagrams and LFs given below. Also see footnote 4.

(24) a. ***For RC extraposition:***
If the determiner and the NP restrictor of the moving nominal is neglected upstairs, and RC is neglected downstairs at PF, then the RC and the determiner of the moving nominal is neglected downstairs, and the NP restrictor is neglected upstairs at LF.

b. ***For HNPS:***
If the whole moving nominal is neglected downstairs at PF, then all of it must be neglected upstairs at LF.

This is the central point that Fox (2017) makes. That is, there's no in-built tool in this Neglect framework that explains the systematic correlation of what is neglected at PF and what is neglected at LF. In Fox's (2017) own words, Sportiche's proposal "relies on independent neglect at the phonological component to derive the word order of extraposition, and therefore does not explain the correlation with semantic neglect . . . " (*ibid.*, p. 28).

This section has recapitulated from Fox (2017) the issue of why partial Neglect doesn't explain WG, but only stipulates it. I will now argue, along the same conceptual lines, that partial LF Neglect is vulnerable to the same objection of simply stipulating the results.

2.2 Problem from Independent Partial LF Neglect

We've seen, in the previous section, that PF and LF seem to need to somehow communicate, which Neglect stipulates but doesn't explain. Here, I will show that, in some cases, while there's nothing extraordinary to be stipulated at PF, LF needs stipulative Neglect operations which don't explain the facts. We find this in the case of stacked RC modifiers in *wh*-movement.

But let's begin with some baselines. It'll help to contextualize the issue of partial LF Neglect in *wh*-movement around partial reconstruction, resulting in Condition C connectivity effects. For instance, it has long been established in the literature that the NP restrictor of a DP undergoing $\overline{\text{A}}$-Movement is reconstructed. This is shown by the fact that, in (25), *wh*-movement can't bleed Condition C (van Riemsdijk and Williams 1981, Freidin 1986, Lebeaux 1988, Fox 1999, *inter multa alia*).

(25) *Which aspect of Alma$_1$ does she$_1$ despise?

However, it's also been shown that the entirety of this NP restrictor isn't subject to reconstruction. For instance, in (26a), the R-expression *Alma* in the CP *that Alma got covid* causes a Condition C violation, as would be expected from (25); however, in (26b), the R-expression *Alma* in the CP *that Alma wrote* doesn't.[6]

(26) a. *Which **proof** that Alma$_1$ got covid did she$_1$ object to?

b. Which **paper** that Alma$_1$ wrote did she$_1$ later publish?

There is a variety of explanations that have been proposed for this. The traditional explanation, championed in Lebeaux (1988), and represented in Lebeaux (1990, 2000, 2009), Chomsky (1993), Fox (1999, 2002, 2017), Fox and Nissenbaum (1999), Overfelt (2015a), *inter alia*, is that there is LM of the RC to the NP of the moved DP *which paper*, as shown in (27).

6. For recent skepticism about these reconstruction facts, see Adger, Drummond, and van Urk (2017) and Bruening and Al Khalaf (2019). For counterarguments against this body of work, see Stockwell, Meltzer-Asscher, and Sportiche (2021, 2022).

(27) Which [paper [$_{RC}$ that Alma$_1$ wrote]]
did she$_1$ later publish ⟨which paper⟩?

A similar derivation is unavailable for (26a). This is because complements (*e.g.,* the CP in (26a)) must be merged before movement, as soon as possible, in contrast to adjuncts (*e.g.,* the RC CP in (26b)) and this is forced by the projection principle (also dubbed the *Local Predicate Saturation* by Sportiche 2016, (39): 16). And this is what derives the asymmetry in (26). The problem with literal countercyclic Merge is, in short, overgeneration. Such overgeneration cases are noted in detail in Sportiche (2018). One example is what the LM in (28) teaches us about (29).[7] (28) shows us that LM of the boldfaced RC must be LMed at least under a complex NP island (and if the ellipsis dots contain island boundaries, then under those island boundaries as well). Of course, there's a conceptual problem of something being merged inside an island. But, even if we set that aside, this example does teach us the empirical necessity of unboundedly countercyclic LM. The problem is that, if we allow this, this kind of unbounded countercyclic LM overgenerates cases like Late *Re*merge, as in (29). In (29), the trace can be LMed in the lower position and the ellipsis dots can stand for any depth of embedding, including potential island boundaries, since LM needs to be capable of being unboundedly countercyclic. One could say that LM is okay as long as it's External LM, and not Late Remerge (one could also call it *Internal LM*) but that would be just a stipulation. This is one of the reasons Sportiche concludes LM should not be permissible.

(28) [Whose [$_{NP}$ claim [$_{CP}$ that . . . John read [$_{DP}$ the [$_{NP}$ paper [$_{CP}$ **that Mary$_2$ wrote**]]]]]]$_1$
did she$_2$ object to t_1?

(29) Near Paris John thinks . . . that you live *t*

[Sportiche (2018), (5): 419]

Sportiche's (2016) explanation of the LM phenomenology is Neglect. According to Sportiche (2016), (26b) is good because the entire lower copy is neglected at PF, as in (30a), and *paper* in the higher copy and *which* and the RC in the lower copy are neglected at LF, as in (30b). Note, crucially, that the necessary PF Neglect operation shown in (30a) is not partial: it's total Neglect of the lower copy of *which paper that Alma wrote*. Therefore, the kind of modularity issue that arose in the case of extraposition doesn't seem to arise here because PF needs to independently neglect the lower copy anyway, regardless of what happens at LF.

(30) a. ***PF:***
Which [paper [$_{RC}$ that Alma$_1$ wrote]]
did she$_1$ later publish ~~which [paper [$_{RC}$ that Alma$_1$ wrote]]~~?

b. ***LF:***
Which [paper [$_{RC}$ that Alma$_1$ wrote]]
did she$_1$ later publish ~~which~~ [paper ~~[$_{RC}$ that Alma$_1$ wrote]~~]?

7. Sportiche (2018) uses the example in (i). Since it's not easily parseable, I use (28).

(i) [Whose criticism of [Mary's rendition of (. . .) the claim [that you [formulated (. . .) the hypothesis **[that Henri [visited the villages near Picasso's$_2$ estate]**]]]]]$_1$ did he$_2$ endorse t_1?

[Sportiche (2018), (5), (4)]

There's still one conceptual issue that merits mention. Consider (28) again. If the necessary LF for it is to be generated by Neglect, then we'll need the following LF Neglect operation. This Neglect operation needs to target something embedded under a complex NP island boundary. Note that such island-insensitive Neglect is never what is necessary for any standard Neglect operation required to neglect some part of some copy in a regular movement chain, because movement is never supposed to cross island boundaries. That is, although we can get around the issue of (29), the introduction of Neglect as a solution to (28) creates a locality issue that I'm not aware any current understanding of locality at either of the interfaces would know how to tackle. However, because I want to see how far we can take Neglect as an account, I would be charitable and set this issue aside.

(31) ***Required LF Neglect operation for (28)***
[whose [$_{NP}$ claim [$_{CP}$ that . . . John read [$_{DP}$ the [$_{NP}$ paper [$_{CP}$ **that Mary$_2$ wrote**]]]]]]$_1$
did she$_2$ object to
[whose [$_{NP}$ claim [$_{CP}$ that . . . John read
[$_{DP}$ the [$_{NP}$ paper ~~[$_{CP}$ that Mary$_2$ wrote]~~]]]]]$_1$

With this background, let's consider the following examples with multiple stacked RC modifiers. To my knowledge, these were first discussed in Sauerland (1998). (34) is a similar contrast. Under an LM-based account, the effect shown is that when there are two adjuncts modifying an NP, only the outer RC can be LMed and, if it can't — *e.g.,* because it contains a pronoun that must be bound by a quantifier downstairs — then the inner one can't either. I will call the (a) examples ***the only-inner sentences,*** and the (b) examples ***the only-outer sentences.***[8]

(32) a. [Which [[computer [compatible with his$_2$]] that Mary$_1$ knew how to use]]$_3$ did she$_1$ tell every boy$_2$ to buy t_3?

b. * [Which [[computer [compatible with Mary's$_1$]] that he$_2$ knew how to use]]$_3$ did she$_1$ tell every boy$_2$ to buy t_3?

(33) a. ? Tell me which books describing Kant's$_1$ views that were published every woman said he$_1$ agreed with.

b. * Tell me which books describing Kant's$_1$ views that she$_2$ published every woman$_2$ said he$_1$ agreed with.

[Sauerland (1998), (2.37-2.38): 52; (2.40b-c): 53]

(34) a. [Which [[book [that he$_1$ hates]] [which Mary$_2$ likes a lot]]]$_3$ is she$_2$ asking every boy$_1$ to reread t_3?

b. ?? [Which [[book [that Mary$_2$ hates]] [which he$_1$ likes a lot]]]$_3$ is she$_2$ asking every boy$_1$ to reread t_3?

[Stanislao Zompì's examples used in an assignment, Spring 2022]

8. In the only-inner sentences, which are acceptable, how to interpret the higher copy with a functional reading, if there's LM of the modifying containing the R-expression, is a non-trivial question. This is tied to the question of how to derive and formalize functional readings of *wh*-questions in general, which I address briefly in section 5. However, I don't go into the specific question of interpreting the higher copy in these Sauerland cases. This is a very complex issue that deserves its own paper. I'm exploring this question in ongoing work.

Sauerland (1998) explains the ungrammaticality of the only-outer sentences by modifying the original Lebeuxian LM explanation a little bit. Borrowing what Tada (1993: 63-70) says about similar examples in Japanese, Sauerland says that, although LM is countercyclic, there's a limit to its countercyclicity. The reason, Sauerland says, (35) is grammatical is because when LM happens to the SpecCP position, that instance of LM is cyclic with respect to that SpecCP. But, in the only-outer sentences, that can't be the case because the first RC must be merged in its base position so that the pronouns in it can get bound by the universal quantifiers, and then, because of this constraint, when LM is to happen in SpecCP, it must apply to a node rather deeply embedded inside this SpecCP, unlike in (35), and this countercyclicity with respect to the SpecCP where LM is happening is responsible for the ungrammaticality of the only-outer sentences.

(35) Which argument that John_1 published did he_1 later disprove?

However, there's some substantial vagueness in this explanation. *Viz.,* how is it that there's no countercyclicity in the LM involved in (35)? That is, this LM must happen to the NP embedded inside the DP sitting in the SpecCP, not to the DP itself. Then, how is the countercyclicity involved in the only-outer sentences also not involved in (35)?

Moreover, there have recently been reports of exceptionally deeply embedded LM effects in Fox (2017). (Also see (28) and footnote 7.)

(36) a. I'll [[explain [a paper that was recommended by a linguist] when we meet] who teaches at UCLA].

b. I'll [[explain [every paper that was recommended by any linguist] when we meet] who teaches at UCLA].

c. I'll [[talk to you about [every paper that was recommended by any linguist] when we meet] who teaches at UCLA].

d. I'll [[do [nothing that was prohibited by anyone] when I am a student] who would have any authority over me].

[Fox (2017), (20a), (21a), (23)]

If Tada's explanation for (32-34) is the correct one, then it should make the sentences in (36) ungrammatical because of the depth of LM happening here, which they are not. Therefore, Tada's explanation cannot be the right one.[9]

Let's now see what Neglect operations would be needed to yield the necessary LFs. I'll schematize the sentences with stacked modifiers as below.

(37) $[\text{which } [[\text{NP } \text{MOD}_1] \text{ MOD}_2]]_3 \ldots t_3$?

Recalling the pattern observed above: only the outer modifier can be LMed, and if, for some reason, it can't be, then the inner one can't either. Therefore, the following should be predicted to be the permissible and impermissible Neglect operations, if we're to derive this pattern.

9. I won't be going into how the sentences in (36) work here. See Fox (2017) for it. He ends up using a way of implementing multidominance very similar to Johnson (2009, 2012, 2018), Fox and Johnson (2016). Moreover, Fox uses, although he doesn't say it explicitly, what is dubbed *horizontal non-bulk sharing* in Gračanin-Yuksek (2007) and exploited in Citko and Gračanin-Yuksek (2013a,b, 2016), Gračanin-Yuksek (2013).

(38) ***PF (same for all LFs):***
[which [[NP MOD$_1$] MOD$_2$]]$_3$
. . . ~~[which [[NP MOD$_1$] MOD$_2$]]$_3$~~

a. ***LF (equivalent to LMing only the outer modifier; permissible):***
[which [[NP ~~MOD$_1$~~] MOD$_2$]]$_3$
. . . [~~which~~ [[NP MOD$_1$] ~~MOD$_2$~~]]$_3$

b. ***LF (equivalent to LMing both modifiers; permissible):***
[which [[NP MOD$_1$] MOD$_2$]]$_3$
. . . [~~which~~ [[NP ~~MOD$_1$~~] ~~MOD$_2$~~]]$_3$

c. * ***LF (equivalent to LMing only the inner modifier; impermissible):***
[which [[NP MOD$_1$] ~~MOD$_2$~~]]$_3$
. . . [~~which~~ [[NP ~~MOD$_1$~~] MOD$_2$]]$_3$

However, Sportiche's concept of Neglect doesn't derive this; given FI, we predict all of the examples in (38) are allowed because in none of them are all of the copies of *wh*-movement totally neglected, but we don't predict why (38a) and (38b) are allowed, while (38c) isn't. Therefore, even though partial LF Neglect is a necessary operation to ignore certain parts of certain copies at LF, it overgenerates.

3 Total LF Neglect

Total Neglect of a copy of a chain — either at LF or PF — would look like either of the two examples in (39).

(39) a. [~~X_1~~ . . . X_1]

b. [X_1 . . . ~~X_1~~]

We've seen that partial LF Neglect isn't sufficient to derive the pattern in (32)-(34). I will now show that total LF Neglect is able to rule out the only-outer sentences, but it also rules out the only-inner sentences. That is, it undergenerates.[10] Recall the schema from above, repeated below.

(37) [which [[NP MOD$_1$] MOD$_2$]]$_3$ . . . t_3?

The following syntax and Neglect operations can generate this string.

(40) a. ***Syntax:***
[which$_4$ [[which$_4$ [[NP MOD$_1$] MOD$_2$]]$_3$ . . . [which [[NP MOD$_1$] MOD$_2$]]$_3$]]

b. ***PF:***
[~~which$_4$~~ [which$_4$ [[NP MOD$_1$] MOD$_2$]]$_3$ . . . ~~[which [[NP MOD$_1$] MOD$_2$]]$_3$~~

c. ***LF:***
[which$_4$ ~~[which$_4$ [[NP MOD$_1$] MOD$_2$]]$_3$~~ . . . [~~which~~ [[NP MOD$_1$] MOD$_2$]]]$_3$

In syntax, the entire *wh*-phrase moves, and then, *which* moves out of it. At PF, only the higher copy of the *wh*-phrase is interpreted and everything else is neglected. At LF, the higher copy of the *wh*-phrase and the *which* head of the lower copy of the *wh*-phrase are neglected and the rest is

10. This observation is owed to Danny Fox (p.c.).

interpreted. This rules out the only-outer sentences, but it also rules out the only-inner sentences because, in (40c), MOD$_2$ fails to be interpreted in the higher position and escape reconstruction. For the same reason, total LF Neglect like this would also rule out the most basic cases of LM phenomenology, as in (26b) (*Which paper that Alma$_1$ wrote did she$_1$ later publish?*).

Therefore, we've seen that while partial Neglect overgenerates, total Neglect undergenerates. However, it should be noted that this doesn't mean that total Neglect must be an impermissible operation. All we can say is that total Neglect is permissible, while additional machinery is needed. In fact, total Neglect is a viable operation, especially because of the well-known *de re*/*de dicto* ambiguity in raising as in (41). **And, crucially, nothing needs to be stipulated because there's no PF-LF correlation in the Neglect operations to be carried out. There's constant, independent lower copy Neglect at PF and alternate possibilities at LF, depending on whether we have *de re* or *de dicto* interpretation.**

(41) [An Austrian]$_1$ is likely to t_1 win the gold medal.
PF:
[an Austrian]$_1$ is likely to ~~[an Austrian]$_1$~~ win the gold medal

a. ✓∃ > likely
LF:
[an Austrian]$_1$ is likely to [~~an~~ Austrian]$_1$ win the gold medal

b. ✓likely > ∃
LF:
~~[an Austrian]$_1$~~ is likely to [an Austrian]$_1$ win the gold medal

[Sauerland and Elbourne (2002), (1): 284]

If total Neglect is a viable option, then one path we could take is to have total Neglect, and look for other machinery to yield the only-inner sentences. **Crucially, this other machinery can't be LM in the strict sense of countercyclic Merge, because, as Sportiche (2018) has noted, LM is unconstrained and vastly overgenerates.** Then, is there some other way of doing DP movement that can give us the net output of LM without there being countercyclic Merge *per se* and also **predict, and not stipulate,** the patterns we've seen in adjunct extraposition and stacked modifiers in *wh*-questions? Such a notion of Merge and Move can apparently be found in Fox (2017). I will now present the relevant parts of it below and then see if it can help us.

4 Fox (2017)

4.1 The System

The spirit of Fox's (2017) system can be summarized by the following long quote:

> Assume that there is no movement rule and that a single operation of merge yields multi-dominance structures — structures in which a single constituent might have more than one mother. In other words, assume (along the lines of much recent literature) that multi-dominance can be derived not only by a rule of Internal Merge and is therefore not limited to cases where one mother of a multi-dominated phrase dominates all other mothers (cases where one position c-commands all others). Such an assumption might license the derivation of the necessary syntactic representations (LM structures) without a counter-cyclic operation of late merge. However, this comes with an obvious cost: the constraints on displacement (e.g. locality) cannot be linked

directly to movement (Internal Merge), since there will be no specific rule responsible for displacement. We, thus, cannot say that the source of locality follows from requirements on Internal Merge. (For example, it would not suffice to say that an object cannot be Internally Merged, unless it is made visible by agreement.)

I, therefore, suggest (in the spirit of Johnson 2012, 2014) that the constraints on displacement follow more indirectly. When multi-dominance arises (when a single constituent has more than one position in a linguistic representation), I will adopt what I think is a rather simple assumption, namely that linearization (which is required for spell-out) fails, unless a specific device is in place which tells the phonology which of the two positions occupied by a single constituent is to be ignored or deleted (Nunes 2001, 2004, Johnson 2012, 2014). I will add to this the assumption that phonological deletion (of the relevant sort) is parasitic on agreement, which is, in turn, conditioned by locality. So movement does not exist as a special rule, and, hence, LM does not exist. Multi-dominance, however, does exist, and subsequently the structures that I will be arguing for can be derived. Merge creates many multi-dominance structures that are not attested. As usual, the hope is that this problem of over-generation can be dealt with by independently motivated constraints (in particular, TC [Trace Conversion] in semantics, linearization in phonology and probably a syntactic residue, e.g. the case filter).

[Fox (2017: 1-2)]

Let's think about how such a conception of Merge will deal with (11a), repeated below in (42). Understanding what the superscripts in the LF in (42b) is crucial. **Every single superscripted character in this LF represents multidominance. That is, for any node *X* and any character α, X^{α} stands for a single syntactic object that occupies every syntactic position in the LF where X^{α} is written.** Therefore, *painting* comes to be multidominated when the RC *that John talked about* is **externally merged** to it to create the new node NP^2, which is then externally merged to a^3, resulting in the multidominance of a^3. That is, NP^1 is present both inside DP^4 and DP^5, so is a^3, but, crucially, **the RC is present only inside DP^5.** This captures the ability of the RC to escape the c-command domain everything inside the VP.

But the presence of the quantificational *a* inside the VP creates well-known problems of type-clash, because a verb of type $\langle e, \langle e, t\rangle\rangle$ can't combine with *a painting*, which is of type $\langle\langle e, t\rangle, t\rangle$. This is fixed during semantic composition by stipulation through Trace Conversion (Fox 2002). Fox (2017) implements Trace Conversion as in (42d), while an indexed determiner can be understood as in (42e). With the help of these tools, semantic composition can proceed as in (42c).

(42) a. We saw a painting yesterday that John talked about.

b. ***LF***
$[_{DP^5}\ a^3\ [_{NP^2}\ [_{NP}\ \text{painting}]^1\ [_{CP}\ \text{that John talked about}]]]$
$[\text{we}\ [\text{saw}\ [_{DP^4}\ a^3\ [_{NP}\ \text{painting}]^1]\ \text{yesterday}]]$

c. ***Semantic composition***
$[\![(42b)]\!]^g$
$= [\![\text{a painting that John talked about}]\!]^g(\lambda x_e\ .\ [\![\text{we saw the}_1\ \text{painting}]\!]^{g[1 \to x]})$
$= [\![\text{a painting that John talked about}]\!]^g(\lambda x_e : x \text{ is a painting}\ .\ [\![\text{we saw } x]\!]^g)$

[Fox (2017), (9): 7]

d. **Trace Conversion (simplified)**

$D_{ett} \longrightarrow \text{the}_i$

e. ***Indexed definite determiners***
$[\![\text{the}_i]\!]^{g[i \to x]} = \lambda P_{et} : x \in P \,.\, x$
(*i.e.,* "the function that takes a predicate, P, is undefined if $x \notin P$ and returns x if $x \in P$" (Fox 2017, footnote 12))[11]

The next question, then, is: how do we make sure that locality is preserved in movement, if Move is unconstrained by Agree? As Fox says, this can be done by making PF Neglect parasitic on Agree, since Agree itself is sensitive to locality. Therefore, in the case of an LF like (42b), Agree would take place between the probe on the head of the VP and DP^4. This would essentially establish a chain relationship between DP^5 and DP^4. Note, therefore, that the "head" and the "tail" of the chain are not identical in this structure: the "head" is larger and contains the RC, as well as all the elements in the "tail". Here, the Agree relation to be established between the head of the VP and DP_4 is local; therefore, the chain is linearizable.

Once the linearizability of the chain is confirmed, we need to ask: how do we actually linearize these structures into the strings that are pronounced? Fox writes an algorithm for that, given in (43), with an auxiliary definition in (44). Here, "delete" can be understood as "neglect".

(43) Let $[YP_2\ [_{\overline{X}}\ X^0\ [\ldots\ YP_1\ \ldots]]]$ be a phrase in which X^0 agrees with Y^0, the head that is shared by YP_1 and YP_2:

a. **Overt Neglect:**
For every terminal, t, dominated by YP_1 and for every position of t, P, if $\overline{X} \in P$, delete P.

b. **Covert Neglect:**
For every terminal, t, dominated by YP_1 and for every position of t, P, if $\overline{X} \notin P$, delete P.

[Fox (2017), (47): 31]

(44) **Position**
A position of a phrase X, in a structure S, is a path that goes from X to S, that is, a sequence of phrases $\langle a_1, \ldots, a_n \rangle$, where $a_1 = X$, $a_n = S$, and a_{j+1} is a mother of a_j, for any j.

[Fox (2017: 30)]

Next, we'll see that this notion of Merge and this algorithm can account for the extraposition facts and partial reconstruction in *wh*-movement, assuming only total Neglect.

4.2 Extraposition and *Wh*-Movement

(43a) will become relevant only when we get to *wh*-movement because it's for overt movement, so I won't elaborate on it right now. (43b) is immediately relevant for us, because this is what we need in order to linearize the extraposition chain in (42b), since this involves covert QR. Let's first identify all the elements. YP_1 is DP_4, YP_2 is DP_5, X^0 is V^0, $\overline{X}$ is $\overline{V}$, and Y^0 is ${D_4}^0$. The terminals in DP^4 are a^3 and all the nodes dominated by NP^1, *i.e., painting*. a^3 and *painting* each have two positions. For the sake of simplicity, I won't write the full paths. It's easy to see informally that

11. See also Fox and Johnson (2016) and Anvari (2024).

the paths that include the sister of $\overline{X}$ do not include $\overline{X}$ itself; the lowest node included in that path is $\overline{X}$'s sister's — and $\overline{X}$'s — mother. That is, for each of these paths, P, $\overline{X} \notin P$. That means that, according to (43b), a^3 and *painting* must be deleted inside the sister of $\overline{X}$, but pronounced inside VP. This derives the attested string *We saw a painting yesterday that John talked about*.

So, it seems we didn't have to stipulate any PF-LF correlation and still were able to capture the NP Requirement, the Determiner Requirement, and the RC Requirement, given in (19). This is because the RC is merged only in the head of the chain, so there's no way for it to be interpreted inside the VP; the determiner, in our specific case, *a*, will be replaced with an indexed definite determiner in its lower position during semantic composition, so the quantifier *a* will be interpreted as a quantifier only in its higher position, that is, higher than all the elements inside the VP; the NP restrictor, *painting* is merged inside the VP and also interpreted there. And a linearization algorithm, the one in (43), **independent of any LF fact,** can yield the attested string. Therefore, the modularity issue we worried about seems to be gone.

Let's now ask whether we can account for partial reconstruction in *wh*-movement. Let's take our initial example in (2), repeated in (45). Recall that we decided to do without partial LF Neglect. So, I've only shown how to derive the LFs with total Neglect. It turns out it works. The syntax in (45a) is exactly as in (40a). At LF, we have total Neglect of the higher occurrence of the *wh*-phrase after the covert movement of *which*. Therefore, only the lower occurrence of the *wh*-phrase is interpreted at LF, while *which* would be turned into an indexed definite determiner during semantic composition. **Crucially, (43) makes sure that the PF side works out too.** Because the movement of the whole *wh*-phrase is overt, (43a) applies and the reader can verify for themselves, along the same lines as the extraposition example we had before, that it's the lower occurrence that would be neglected. On the other hand, the movement of *which* is covert, so, (43b) applies and the higher occurrence of it is neglected. This exactly derives the attested string.

(45) Which side of itself$_1$ did the sofa$_1$ fall on?

a. ***Syntax:***
[which^3 [[which^3 [picture of itself$_1$]]]^2 did the sofa$_1$ fall on [which [picture of itself$_1$]]^2]]
covert
overt

b. ***LF; total Neglect; no modularity issue:***
[which^3 [[~~which^3 [picture of itself$_1$]]~~]^2
did the sofa$_1$ fall on
[which [picture of itself$_1$]]^2]]

c. ***PF; predicted by Fox's algorithm:***
[~~which^3~~ [[which^3 [picture of itself$_1$]]^2
did the sofa$_1$ fall on
~~[which [picture of itself$_1$]]^2~~]]

But this is not the end of our consideration of partial reconstruction in *wh*-movement. We still need to see what happens to the only-inner and only-outer sentences. But before that, let's sort out a semantic issue in the next section.

5 A Tangent: Unary *Which*

I've been showing LFs where *which* seems to be a unary function that moves to attach to a whole question CP, for instance, in (40a) and (45a). We want to understand whether this is something

reasonable or independently required. Such a unary entry for *which* has been proposed in Heim (2019). So, I'll briefly discuss the relevant points of her paper.

In terms of Hamblin (1973), Karttunen (1977), the existential quantifier in the question denotation of (46a) can't range over *e*-type entities, but, as we know from Engdahl (1986), the quantification must be over $\langle e, e\rangle$-type functions; informally speaking, over functions that map girls to pictures of themselves. Engdahl (1986) would do this by having the pronoun *herself* bound by a covert variable binder cum type-shifter. I follow Heim's (2019) way of presenting Engdahl's ideas for expository reasons.

(46) a. Which picture of herself$_1$ did no girl$_1$ submit?

b. ***LF for (46a)***
$[\lambda_p$ [which $[E_y$ [picture$_{w@}$ of herself$_y$] λ_f [[$Q(p)$]
$[\lambda_w$ [[no girl$_{w@}$] $[\lambda_x$ [t_x submit$_w$ $t_{f(x)}$]]]]]]]]

c. ***Denotation for the LF in (46b)***
$\{p : \exists f_{\langle e,\, e\rangle}\ [\forall x\ .\ \textsc{picture-of}_{w@}(f(x), x) = 1 \wedge$
$p = \lambda w_s\ .\ \neg\exists x[\textsc{girl}_{w@}(x) = 1 \wedge \textsc{submit}_w(x, f(x)) = 1]]\}$

d. **Q *encodes Karttunen's "proto-question" formation***
$[\![Q]\!] = \lambda p_{st}\ .\ \lambda q_{st}\ .\ p = q$

e. **Which *is an existential quantifier***
$[\![\text{which}]\!] = \lambda f_{et}\ .\ \lambda g_{et}\ .\ \exists x[f(x) = 1 \wedge g(x) = 1]$

f. **E *is a covert variable binder and type-shifter***
$[\![E_y\ \zeta]\!]^g = \lambda f_{\langle e,\, e\rangle}\ .\ \forall x\ .\ [\![\zeta]\!]^{g^{x/y}}(f(x)) = 1$

[Heim (2019: 283-285); slightly modified]

For reasons Heim (2019) elaborates in her section 5, having the pronoun bound upstairs and making it not depend on the ϕ-features and structural position of the quantificational DP *no girl* give rise to problems. Because of this, Heim proposes to remodel the functional readings of sentences like (46a) by interpreting the *wh*-restrictor *picture of herself in situ*. I won't present the entirety of Heim's system here, but it can generate the LF in (47b). See Heim's paper for details on why and how this system is developed. Notice, crucially, that these considerations end up motivating a unary *which*, given in (47d). I'll take this to be the theoretical justification for having movement operations as in (40a) and (45a).

(47) a. Which picture of herself$_1$ did no girl$_1$ submit?

b. ***LF for (47a)***
$[\lambda_p$ [which $[\lambda_f$ [[$Q(p)$] $[\lambda_w$ [[no girl$_{w@}$] $[\lambda_y$ [t_y submit$_w$
[the picture$_{w@}$ of herself$_y$ ident [f *pro*$_y$]]]]]]]]]]

c. ***Denotation for the LF in (47b)***
$\{p : \exists f[\forall y[\textsc{girl}_{w@}(y) = 1 \rightarrow y \in \text{dom}(f) \wedge \textsc{picture-of}_{w@}(f(y), y) = 1] \wedge$
$p = \lambda w_s\ .\ \neg\exists y[\textsc{girl}_{w@} = 1 \wedge \textsc{submit}_w(y, f(y)) = 1]]\}$[12]

d. ***Polymorphic* which**
$[\![\text{which}]\!] = \lambda P_{\sigma t}\ .\ \exists x_\sigma[P(x)]$

12. Heim (2019: 292) writes the union of this set and $\{\varnothing\}$ as the denotation, for reasons related to presupposition projection and the pathological element of propositional type ($\#_{st}$). I'll omit that since that complication isn't necessary for my purposes here.

where σ is any type.

e. ***Polymorphic THE***
$[\![\text{THE}]\!] = \lambda P_{\sigma t} : \exists! x_\sigma[P(x)] \,.\, \iota x_\sigma[P(x)]$
where σ is any type.

f. ***Polymorphic IDENT***
$[\![\text{IDENT}]\!] = \lambda x_\sigma \,.\, \lambda y_\sigma \,.\, x = y$
where σ is any type.

[Heim (2019: 291-292); slightly modified]

Having sorted this out, we can now consider what happens to *wh*-movement with stacked modifiers, when we consider them in the context of Fox (2017) and the possibility of total Neglect.

6 Stacked Modifiers, Fox (2017), and Total Neglect

6.1 Attempts at Solution with Innocent Assumptions Fail

Let's recall the only-inner and only-outer examples and their schema, repeated below.

(32) a. ***Only-inner sentence***
[Which [[computer [compatible with his$_2$]] that Mary$_1$ knew how to use]]$_3$ did she$_1$ tell every boy$_2$ to buy t_3?

b. ***Only-outer sentence***
* [Which [[computer [compatible with Mary's$_1$]] that he$_2$ knew how to use]]$_3$ did she$_1$ tell every boy$_2$ to buy t_3?

(37) [which [[NP MOD$_1$] MOD$_2$]]$_3$ . . . t_3?

In this subsection, I will try some possible derivations of the only-inner sentences. These derivations will involve mostly innocent assumptions, and all of them will fail. So, by the end of this subsection, we will show that most, if not all, of the derivations with no outlandish assumptions fail at deriving the only-inner sentences. In the next subsection, I will consider a derivation that will succeed, but, as I will argue, requires an outlandish derivational step.

Given we have total Neglect and Fox's multidominance now, we're at least going to have the movement shown in (48) in the syntax, which is the *wh*-movement. What has happened here is that MOD$_2$ has been externally merged to the already merged node [$_{\text{NP}^7}$ NP MOD$_1$]^5, resulting in the multidominance of the latter and creation of the new node NP^6. (NP^7 is just the label of the node that's superscripted "5". These are not different syntactic objects.) This node NP^6 is then externally merged to *which*^{4}, resulting in the multidominance of the latter.

(48) [[which^4 [$_{\text{NP}^6}$ [$_{\text{NP}^7}$ NP MOD$_1$]^5 MOD$_2$]]$_3$ . . . [which^4 [$_{\text{NP}^7}$NP MOD$_1$]^5]]

But this can't be the end of the derivation. There needs to be total LF Neglect of the head of the *wh*-chain because we need to achieve the effect of MOD$_1$ being "reconstructed" because it contains the bound pronoun. But this head of the *wh*-chain can't just be neglected totally as it is in (48) because it contains MOD$_2$ as well, which has to be interpreted but will go uninterpreted if the whole head of the chain is neglected at LF. That means that, besides moving *which*^{4} out of

the head of the chain, we need to first move out MOD_2 as well, before we can totally neglect the head of the chain. These two movements can be done as follows.

(49) $[\text{MOD}_2{}^8\ [\text{which}^4\ [[\text{which}^4\ [_{\text{NP}^6}\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5\ \text{MOD}_2{}^8]]_3 \ldots$

covert

covert

$\ldots\ [\text{which}^4\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5]]]]$

After this, we can have total Neglect at the interfaces as follows. As the reader can verify, Fox's algorithm again predicts the PF Neglect operations, because of the covertness of the movements of both *which*^{4} and $\text{MOD}_2{}^8$. This derives the attested string of the only-inner sentences.

(50) a. ***LF; total Neglect; no modularity issue:***
$[\text{MOD}_2{}^8\ [\text{which}^4\ [[\sout{\text{which}^4\ [_{\text{NP}^6}\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5\ \text{MOD}_2{}^8]]_3} \ldots$
$[\text{which}^4\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5]]]]$

b. ***PF; predicted by Fox's algorithm:***
$[\sout{\text{MOD}_2{}^8}\ [\sout{\text{which}^4}\ [[\text{which}^4\ [_{\text{NP}^6}\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5\ \text{MOD}_2{}^8]]_3 \ldots$
$\sout{[\text{which}^4\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5]}]]]$

But the LF in (50a) isn't interpretable after all. After *which*^{4} combines with its sister at its highest position, the resulting node is of type *t*. This node is what $\text{MOD}_2{}^8$ attaches to. $\text{MOD}_2{}^8$, crucially, is a function that characterizes $\langle e, e\rangle$-type functions. That is, its characteristic set contains $\langle e, e\rangle$-type functions that map entities to entities that Mary knows how to use. So, $\text{MOD}_2{}^8$ has to be of type $\langle\langle e, e\rangle, t\rangle$. This can't combine with a sister of type *t*.

So perhaps, $\text{MOD}_2{}^8$ tucks in under *which*^{4}, as in (51). The same Neglect operations as in (50) would apply in this case as well.

(51) $[\text{which}^4\ [\text{MOD}_2{}^8\ [[\text{which}^4\ [_{\text{NP}^6}\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5\ \text{MOD}_2{}^8]]_3 \ldots$

covert

covert

$\ldots\ [\text{which}^4\ [_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5]]]]$

In this LF, the sister of the head of the *wh*-chain is of type $\langle\langle e, e\rangle, t\rangle$. This can combine with $\text{MOD}_2{}^8$ via generalized Predicate Modification.[13]

But things are still not entirely okay. Notice that we have Heim's unary *which* in our LF. So, its sister has to be of type $\langle\langle e, e\rangle, t\rangle$. But if we're being really careful, is the sister of $\text{MOD}_2{}^8$ really of type $\langle\langle e, e\rangle, t\rangle$ as well or should the combination of these two sister nodes even involve Predicate Modification? Notice that $\text{MOD}_2{}^8$ has moved to attach to its sister, so there should be a λ-binder that attaches right below it, turning it sister into something of type $\langle\langle\langle e, e\rangle, t\rangle, \langle\langle e, e\rangle, t\rangle\rangle$. But is there a trace anymore for this $\text{MOD}_2{}^8$? No, because its trace is inside the totally

13. This raises the question of how to get MOD_2 to have a meaning of type $\langle\langle e, e\rangle, t\rangle$, instead of type $\langle e, t\rangle$. This is an independent question that any theory that tries to account for functional readings of *wh*-questions with stacked modifiers with reconstruction patterns like the only-inner sentences needs to answer and not a side-effect solely of my considerations in this paper. I will therefore put this issue aside. Perhaps, there's a type-shifter like the one defined in (i). This will type-shift MOD_2 into a characteristic function of a set of functions each of which is of type $\langle e, e\rangle$ and always maps entities to entities that Mary knows how to use.

(i) $[\![\mathcal{E}]\!] = \lambda P_{et}\,.\,\lambda f_{ee}\,.\,\text{codom}(f) \subseteq P$

neglected *wh*-phrase, which is supposed to not be interpreted. But a trace must be interpreted as a trace, otherwise, λ-binding doesn't make sense. So, we're stuck. Do we find ourselves equally stuck when the unary *which* moves? This is not something Heim (2019) answers. For her, the attachment of *which* to its sister resulting in unary composition without there being any λ-binder involved in the process actually must be stipulated, especially the λ-binder-free nature of the process. And the *wh*-phrase being totally neglected doesn't cause any problems in the case of *which* because no trace needs to be interpreted, since, by Heim's stipulation, there's no λ-binder binding it.

Essentially, the question we're asking is: does it make sense to have covert movement for the purposes of interpretation (*e.g.,* type reasons, Neglect, perhaps among others) that doesn't introduce any λ-binder? Even if we respond affirmatively to this question, we do know, because of the existence of QR, that there is also covert movement that must introduce λ-binders. That is, **we began by seeking to eliminate the stipulations surrounding partial Neglect and PF-LF correlations in Neglect, and we've ended up realizing that we need to stipulate which covert movements can introduce λ-binders and which ones can't**. It seems we haven't made progress.

There's still another route we could take, which is to have an LF like (52). This differs critically from (48), where the entire $[_{\text{NP}^7}\ \text{NP}\ \text{MOD}_1]^5$ node was multidominated. What happens in (52) is that just the smallest NP node, *i.e.,* NP^3 is multidominated and pulled into the "higher copy" of the *wh*-chain. This results in essentially nothing having to be neglected for the purposes of "reconstruction" because MOD_1, which contains the bound pronoun, is successfully c-commanded and bound by the quantifier (omitted by the ellipsis dots) and it doesn't need to be prevented by Neglect from being interpreted upstairs. This removes the need for moving MOD_2 out of the "higher copy" of the chain because the "higher copy" doesn't need to be neglected.

(52) $[\text{which}^4\ [\text{which}^4\ [\text{NP}^3\ \text{MOD}_2]]\ \ldots\ [\text{which}^4\ [_{\text{NP}^7}\ \text{NP}^3\ \text{MOD}_1]^5]]$

covert

The problem arises at PF. Fox's algorithm predicts the following PF Neglect operations.

(53) ***PF leading to crash for the LF in (52); predicted by Fox's algorithm:***
$[\sout{\text{which}^4}\ [\text{which}^4\ [\text{NP}^3\ \text{MOD}_2]]\ \ldots\ \sout{[\text{which}^4\ [_{\text{NP}^7}\ \text{NP}^3\ \text{MOD}_1]^5]}]$

In this PF, the entire lower copy of the *wh*-chain, **crucially, including MOD_1,** is neglected. This happens despite the fact that MOD_1 is not multidominated because **(43a) involves universal quantification over all terminals inside a lower copy, regardless of whether all of them are multidominated.** However, MOD_1 isn't multidominated and therefore, **neglecting the only position where it's merged violates FI.** Because of these contradictory requirements, this PF should crash. So, we've seen that (52) doesn't help us either. That is, we've gone through attempts to derive the reconstruction pattern while making only innocent assumptions, and we've confirmed that these attempts fail.

6.2 The Only Conceivable Solution is Arguably Outlandish

In this section, I will lay out another derivation of the only-inner sentences that will succeed, **because of a crucial step in the derivation, which I will argue is outlandish.** This will help us make precise what we need to succeed at deriving, because we'll be able to clearly characterize the outlandish nature of that step. The derivation in question is summarized in (54).

(54) a. ***LF; total Neglect; no modularity issue:***
[[$_{NP^{11}}$ which^4 [$_{NP^{12}}$ NP^3 MOD$_2$^{8}]] ~~[which^4 [$_{NP^{10}}$ [$_{NP^7}$ NP^3 MOD$_1$^{6}] MOD$_2$^{8}]]~~
. . . [which^4 [$_{NP^7}$ NP^3 MOD$_1$^{6}]^5]]

b. ***PF; predicted by Fox's algorithm:***
[~~[$_{NP^{11}}$ which^4 [$_{NP^{12}}$ NP^3 MOD$_2$^{8}]]~~ [which^4 [$_{NP^{10}}$ [$_{NP^7}$ NP^3 MOD$_1$^{6}] MOD$_2$^{8}]]
. . . ~~[which^4 [$_{NP^7}$ NP^3 MOD$_1$^{6}]^5]~~]

Here's how the LF is built. First, [which^4 [$_{NP^7}$ NP^3 MOD$_1$^{6}]^5] is base-generated downstairs. To form the higher copy, we merge NP^7 with MOD$_2$^{8} and create the node NP^{10} upstairs. This results in NP^7 being multidominated. NP^{10} is then merged with *which*^{4} and thus, *which*^{4} becomes multidominated. In the next step, we play a crucial trick. We take NP^3 and MOD$_2$^{8} — which don't form a constituent yet — and merge them together to form the node NP^{12}. In NP^{12}, NP^3 and MOD$_2$^{8} now form a constituent. Next, NP^{12} is merged with *which*^{4}, resulting in something that can be characterized as the covert movement of [$_{NP^{11}}$ which^4 [$_{NP^{12}}$ NP^3 MOD$_2$^{8}]], despite the fact that [$_{NP^{11}}$ which^4 [$_{NP^{12}}$ NP^3 MOD$_2$^{8}]] didn't use to be a constituent before the combinatorial trick was played.

At LF, we will have total Neglect of [which^4 [$_{NP^{10}}$ [$_{NP^7}$ NP^3 MOD$_1$^{6}] MOD$_2$^{8}]]. This lets MOD$_1$ be interpreted only downstairs, which is the result we needed to achieve, to ensure "reconstruction". Since MOD$_2$ was never downstairs, we replicate the effect of LM and obviate Condition C. This single Neglect operation is completely independent of any Neglect operation that needs to take place at PF, so there's no modularity issue.

At PF, Fox's algorithm gives us the right result. When [which^4 [$_{NP^{10}}$ [$_{NP^7}$ NP^3 MOD$_1$^{6}] MOD$_2$^{8}]] is built, it's built as the higher copy of an overt movement because we're concerned with overt *wh*-movement. Therefore, the Overt Neglect algorithm forces the lower copy, that is, [which^4 [$_{NP^7}$ NP^3 MOD$_1$^{6}]^5], to be deleted. As indicated above, when NP^{11} is built, it's covert movement. Therefore, the Covert Neglect algorithm forces NP^{11} to be neglected upstairs, and the result is that it's as if NP^{11} were never formed, as far as PF is concerned, that is, as if only the first step of movement has happened. Therefore, it seems we have derived the only-inner sentence with total Neglect without creating any modularity issue.

However, what I would like to caution against in this derivation is the combinatorial trick we played, when we combined NP^3 and MOD$_2$^{8}. Here's how we can get an informal idea of why such a trick might be the outlandish part I alluded to in the beginning of this subsection. Suppose we're going through the derivation of an RC extraposition structure and we've built VP^8 in (55a) in an ordinary fashion. After building this, we build DP^{10} in (55b) by playing the same trick we played in the derivation above: we leave out the adjectives *expressionist*^{4} and *sepia*^{6} downstairs and merge the bolded RC^{12} in (55b) to just the noun *self-portrait*^{7}, thereby resulting in the multidominance of just *self-portrait*^{7}, without *expressionist*^{4} and *sepia*^{6} being multidominated as well. Next, DP^{10} and VP^8 are merged, resulting in the final VP structure with extraposition, as shown in (55c). That is, the restriction of *every*^{2} upstairs is the intersection of *self-portrait* and the RC, while the restriction of *every*^{2} downstairs (to be converted to the definite determiner via Trace Conversion) is the intersection of *expressionist*^{4}, *sepia*^{6}, and *self-portrait*^{7}.

(55) a. [$_{VP^8}$ [$_{VP^9}$ I found [$_{DP^1}$ every^2 [$_{NP^3}$ expressionist^4 [$_{NP^5}$ sepia^6 self-portrait^7]]]] yesterday]

b. [$_{DP^{10}}$ every^2 [$_{NP^{11}}$ self-portrait^7 [$_{RC^{12}}$ **that John was looking for**]]]

c. [[$_{DP^{10}}$ every^2 [$_{NP^{11}}$ self-portrait^7 [$_{RC^{12}}$ **that John was looking for**]]]
. . . [$_{VP^8}$ [$_{VP^9}$ I found [$_{DP^1}$ every^2 [$_{NP^3}$ expressionist^4 [$_{NP^5}$ sepia^6 self-portrait^7]]]] yesterday]]

The possibility of a derivation like this shows that, if we admit the possibility of the trick we played, we predict that a sentence like *I found every expressionist sepia self-portrait yesterday that John was looking for* can mean something like "For every self-portrait that John was looking for, *x*, I found the self-portrait *x*, ***x* being presupposed to be expressionist and in sepia**". This presupposition arises from the definite determiner with which Trace Conversion replaces the downstairs *every* during semantic composition. Assuming universal projection from under *every*, we derive that this sentence would have the presupposition that **every self-portrait that John was looking for is expressionist and in sepia.** This presupposition, crucially, is non-trivial, that is, it isn't going to be trivially met. And, this presupposition is clearly not an attested presupposition of this sentence, which we can verify from the felicity of a discourse like the following: *I don't know whether every self-portrait that John was looking for was expressionist and in sepia, but I can tell you this much: I found every expressionist sepia self-portrait yesterday that John was looking for*. Therefore, the trick we played above predicts unattested presuppositions and this is why I argue this is outlandish. That is, **the only kind of derivation that can generate the only-inner sentences requires doing something which we can empirically show must be prevented.** A formalization of how to prevent this is a separate issue. The fact that this problem exists is enough to rule out the possibility in (54) on empirical grounds.

7 Conclusion

What we've seen in this paper is that, when we're careful about the patterns and correlations across multiple examples (and not just interpretations of only one datum), partial Neglect creates a multitude of problems. It creates modularity issues, stipulates, and doesn't explain. Furthermore, total Neglect, equipped with multidominance so we're able to produce LM-like structures without countercyclic Merge and its attendant overgenerative power, faces issues when we consider reconstruction patterns in stacked modifiers in *wh*-movement. **It must be noted, however, that we still haven't seen evidence against total Neglect *per se*; it's still necessary for accounting for reconstruction for quantifier scope in raising constructions. All we've really seen is that total Neglect, combined with Fox's system, can't fix the problems partial Neglect was meant to.** Partial Neglect, at best, shows us which problems need to be solved. In fact, they were always problems to be solved. In future work, I plan to build a predictive account to explain the facts discussed here, with any necessary assumptions made explicit and independently justified, whenever possible. I will close this paper with this.

References

Adger, David, Alex Drummond, and Coppe van Urk. 2017. Is there condition C reconstruction? In *Proceedings of NELS 47*, ed. Andrew Lamont and Katerina Tetzloff, 21–31. Amherst, Massachusetts: GLSA.

Anvari, Amir. 2024. Revisiting *e-type* theory. Talk given at MIT LingLunch on February 15, 2024.

Bhatt, Rajesh, and Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35:1–46.

Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge.

Bobaljik, Jonathan. 2002. A-chains at the PF-interface: copies and 'covert' movement. *Natural Language and Linguistic Theory* 20:197–267.

Bošković, Željko. 2015. On multiple left-branch dislocation: Multiple extraction and/or scattered deletion. In *Proceedings of FASL 23*, ed. Małgorzata Szajbel-Keck, Roslyn Burns, and Darya Kavitskaya, 20–35.

Bruening, Benjamin, and Eman Al Khalaf. 2019. No argument-adjunct asymmetry in reconstruction for binding condition C. *Journal of Linguistics* 55:247–276.

Charnavel, Isabelle. 2020. Logophoricity and locality: a view from French anaphors. *Linguistic Inquiry* 51.4:671–723.

Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor binding: What French inanimate anaphors show. *Linguistic Inquiry* 47.1:35–87.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Ken Hale and Jay Keyser, 1–52. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

Citko, Barbara, and Martina Gračanin-Yuksek. 2013a. Multiple coordinated relative clauses. Talk presented at the annual meeting of the Linguistic Society of America.

Citko, Barbara, and Martina Gračanin-Yuksek. 2013b. Towards a new typology of coordinated *wh*-questions. *Journal of Linguistics* 49:1–32.

Citko, Barbara, and Martina Gračanin-Yuksek. 2016. Multiple (coordinated) (free) relatives. *Natural Language and Linguistic Theory* 34.2:393–427.

Engdahl, Elisabet. 1986. *Constituent questions*. Dordrecht, The Netherlands: D. Reidel Publishing Company.

Fanselow, Gisbert, and Damir Ćavar. 2002. Distributed deletion. In *Theoretical approaches to universals*, ed. Artemis Alexiadou, 65–107. John Benjamins Publishing Company.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

Fox, Danny. 2017. Embedded Late Merge and the theory of displacement. Unpublished manuscript, Massachusetts University of Technology.

Fox, Danny, and Kyle Johnson. 2016. QR is restrictor sharing. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, ed. Kyeong-min Kim, Pocholo Umbal, Trevor Block, Chan Queenie, Tanie Cheng, Finney Kelli, Mara Katz, Sophie Nickel-Thompson, and Shorten Lisa, 1–16.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: a case for overt QR. In *Proceedings of the West Coast Conference on Formal Linguistics 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 132–144. Somerville, Massachusetts: Cascadilla Press.

Freidin, Robert. 1986. Fundamental issues in the theory of binding. In *Studies in the acquisition of anaphora*, ed. Barbara Lust, volume 1, 151–188. Dordrecht: Reidel.

Gračanin-Yuksek, Martina. 2007. About sharing. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Gračanin-Yuksek, Martina. 2013. Linearizing multidominance structures. In *Challenges to linearization*, ed. Theresa Biberauer and Ian Roberts, 269–294. de Gruyter.

Hamblin, Charles. 1973. Questions in Montague grammar. *Foundations of Language* 10:41–53.

Heim, Irene. 2019. Functional readings without type-shifted noun phrases. In *Reconstruction effects in relative clauses*, ed. Manfred Krifka and Mathias Schenner, 283–301. Berlin, Germany: de Gruyter.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, Massachusetts: Blackwell.
Johnson, Kyle. 2009. Why movement? Unpublished manuscript, University of Massachusetts at Amherst.
Johnson, Kyle. 2012. Toward deriving differences in how *Wh* movement and QR are pronounced. *Lingua* 122:529–553.
Johnson, Kyle. 2018. Movement as Multidominance. Lecture notes from a course taught at the University of Crete, Rethymnon.
Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.
Ladusaw, William A. 1979. Polarity sensitivity as inherent scope relations. Doctoral Dissertation, University of Massachusetts, Amherst, Amherst, MA.
Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral Dissertation, University of Massachusetts, Amherst.
Lebeaux, David. 1990. Relative clauses, licensing, and the nature of the derivation. In *Proceedings of the North East Linguistic Society 20*, ed. Juli Carter, Dechaine Rose-Marie, Bill Philip, and Tim Sherer, 318–332. University of Massachusetts at Amherst: Graduate Linguistic Student Association.
Lebeaux, David. 2000. *Language acquisition and the form of the grammar*. John Benjamins Publishing Company.
Lebeaux, David. 2009. *Where does binding theory apply?*. Cambridge, Massachusetts: MIT Press.
Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32:303–344.
Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Linguistic Inquiry Monographs. Cambridge, Massachusetts: MIT Press.
Overfelt, Jason. 2015a. Extraposition of NPIs from NP. *Lingua* 164:25–44.
Overfelt, Jason. 2015b. Rightward movement: A study in locality. Doctoral Dissertation, University of Massachusetts at Amherst.
Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge: MIT Press.
van Riemsdijk, Henk, and Edwin Williams. 1981. NP-structure. *The Linguistic Review* 1:171–217.
Sauerland, Uli. 1998. The meaning of chains. Doctoral Dissertation, Massachusetts Institute of Technology.
Sauerland, Uli, and Paul Elbourne. 2002. Total reconstruction, PF-movement, and the derivational order. *Linguistic Inquiry* 33:283–319.
Sportiche, Dominique. 2016. Neglect. Ms. UCLA. https://ling.auf.net/lingbuzz/002775.
Sportiche, Dominique. 2018. Somber prospects for late merger. *Linguistic Inquiry* 50.2:416–424.
Stockwell, Richard, Aya Meltzer-Asscher, and Dominique Sportiche. 2021. There is reconstruction for condition c in english questions. In *Proceedings of NELS 51*, ed. Alessa Farinella and Angelica Hill, 205–214. Amherst, Massachusetts: GLSA.
Stockwell, Richard, Aya Meltzer-Asscher, and Dominique Sportiche. 2022. Experimental evidence for the condition C argument-adjunct asymmetry in English questions. In *Proceedings of NELS 52*, ed. Özge Bakay, Breanna Pratley, Eva Neu, and Peyton Deal, volume 2, 145–158. Amherst, Massachusetts: GLSA.
Tada, Hiroaki. 1993. A/A-bar partition in derivation. Doctoral Dissertation, Massachusetts Institute of Technology.
Williams, Edwin. 1974. Rule ordering in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.