For *Minimalist Syntax*,
Edited by Randall Hendrick,
Blackwell Publishers
First draft, May 2001, Second draft May 2002

On Logical Form*

Danny Fox, Massachusetts Institute of Technology

A Logical Form (LF) is a syntactic structure that is interpreted by the semantic component. For a particular structure to be a possible LF it has to be possible for syntax to generate it and for semantics to interpret it. The study of LF must therefore take into account both assumptions about syntax and about semantics, and since there is much disagreement in both areas, disagreements on LF have been plentiful. This makes the task of writing a survey article in the field fairly difficult, a difficulty that is amplified by the amount of material that needs to be covered if the result is going to be in any way representative. My response to this difficulty is to limit my objectives. As a start, I will confine myself to issues relating to the syntactic positions of Quantificational Noun Phrases (QNPs) at LF and to various interpretive consequences. But even within these relatively narrow confines, I will not attempt anything close to a comprehensive survey. Instead my goal will be restricted to the presentation of one leading idea and to the discussion of some evidence that might bear on it.[1]

Much research on the nature of LF has consisted in attempts to account for the meaning of sentences containing QNPs:

(1) a. A girl is tall.
    b. Many girls are tall.
    c. Every girl is tall.
    d. No girl is tall.

---

* Thanks to Jonathan Bobaljik, Noam Chomsky, Irene Heim and Jon Nissenbaum.

[1] The editor of this volume (Randall Hendrick) asked me to dedicate half of this paper to a discussion of some of the traditional arguments in favor of QR and to use the other half to present some of my own work. There is no clear point where the first half ends and the second begins, but the transition is somewhere in the middle of section 3.

Since it is generally assumed that the meaning of sentences is related to the meaning of their parts, it might be useful to come up with a hypothesis about the meaning of QNPs.

So let's think about the sentences in (1) and how their meanings might result from the meaning of the predicate *tall* combined with the meaning of the QNP (a *girl* in (1)a, *many girls* in (1)b, *every girl* in (1)c, and *no girl* in (1)d). Basic predicates express properties of individuals. The predicate *tall* expresses the property that an individual has if the individual is tall, and not otherwise. This can be modeled if we think that semantically the predicate is a function that maps individuals to truth values. Under standard terminology, the predicate *denotes* a function that maps an individual to TRUE if the individual is tall and to FALSE otherwise (to TRUE if and only if the individual is tall). If the subject of the predicate denotes an individual, the predicate can take the subject as its argument. But there is no straightforward way to think of QNPs as *denoting* individuals. So what can QNPs denote in order combine with the predicate *tall*?

One way to deal with this problem is to assume that QNPs denote second order predicates. They convey information about basic (first order) predicates like *tall*; they tell us something about the set of individuals that a given (first order) predicate is true of. So in the sentences in (1) the relevant predicate is *tall*. And, given the meaning of the specific QNPs, the sentences convey the information that the predicate is true of at least one girl, (1)a, many girls, (1)b, every girl, (1)c, or no girl, (1)d.[2]

Assuming that this is correct, we will say that QNPs denote functions that take predicates as arguments and map these predicates to truth-value:

(2) a. *a girl* denotes a function, f, (from predicates to truth values) that maps a predicate, P, to TRUE if and only if (iff) there is at least one girl, g, such that P(g) = TRUE.

b. *many girl* denotes a function, f, (from predicates to truth values) that maps a predicate, P, to TRUE iff there are many girls, g, such that P(g) = TRUE.

[2] The treatment of quantification as second order predication is due to Frege (see Dummett 1977). The application to natural language semantics together with the observation that QNPs are restricted quantifiers comes from Montague. The presentation in this introduction follows the logic of Heim and Kratzer (1998).

c. *every girl* denotes a function, f, (from predicates to truth values) that maps a predicate, P, to TRUE iff every girl, g, is such that P(g) = TRUE.

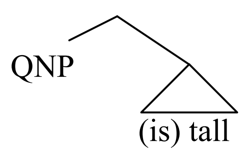
d. *No girl* denotes a function, f, (from predicates to truth values) that maps a predicate, P, to TRUE iff no girl, g, is such that P(g) = TRUE.

These statements raise an immediate question regarding the relationship between structure and interpretation:

(3) How is the argument, P, of a QNP determined?

In the cases in (1) the argument of the QNP is its sister, the predicate *tall*:[3]

(4)



Therefore, if we limit our attention to cases of this sort, we have a fairly simple candidate for an answer to the question in (3):

(5) The argument of a QNP is always the sister of the QNP.

This answer is not only a simple answer given the cases in (1). It is a simple answer given the fact that functions in general take their sisters as arguments. However, the answer (simple as it may be) is inadequate for the full range of constructions in which QNPs seem to occur. To see this, consider the following sentence:

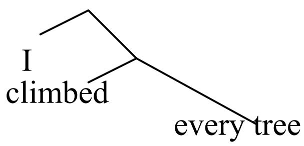
(6) I climbed every tree.

For a slightly different way of thinking about the semantics of QNPs, see Larson and Segal (1995; chapter 7), as well as Heim (1997).

[3] I am ignoring the copula *is* assuming that it is semantically vacuous. To simplify the discussion, I am also ignoring "functional categories" that are not vacuous, such as tense.

The meaning of the QNP is a function that maps its argument, P, to a truth value: TRUE iff P is true of every tree (see (2)c above). This truth value would correspond to speakers judgments about the sentence if P were the predicate that is true of exactly those things that I climbed (λx. I climbed x).[4] However, the sister of *every tree* doesn't seem to express a predicate of this sort:[5]

(6)'



The answer given in (5) is sufficient only if the sister of a QNP is a one-place predicate (a function from individuals to truth values; the type of function that P in (2) has to be if the definitions are to make sense). Whenever this is not the case, it is not clear how the argument of the QNP is to be determined.

This problem becomes more acute when we attempt to account for the interpretations available for the sentence in (7). This sentence has two meanings; it can assert the existence of a single boy who climbed all of the trees or alternatively it can assert that for every tree there is a (perhaps different) boy such that the boy climbed the tree.

(7) A boy climbed every tree.

This ambiguity, commonly referred to as a scope ambiguity, receives a very specific characterization given our assumption about the meaning of QNPs. The first interpretation results from the QNP *a boy* having as its argument or scope (terms that I will henceforth use interchangeably) the predicate which is true of exactly those individuals who climbed every tree (λx. x climbed every tree). The second interpretation results for the QNP *every tree* having as its argument (or scope) the predicate which is

[4] λx.ϕ, where ϕ is a sentence, should be read as a function from individuals to truth values that assigns TRUE to x, iff ϕ is true.

[5] The problem generalizes to all cases in which the QNP is in the non-external position of a predicate that has two or more arguments.

true of exactly those things that a boy climbed (λx. a boy climbed x). The problem is to find a procedure that will determines arguments/scopes for QNPs and will account for this ambiguity. More specifically we need a procedure that (in certain cases) allows for two possible interpretations: the subject QNP can be interpreted as part of the argument/scope of the object QNP (in which case we will say that the object QNP has wide scope) or alternatively the object QNP can be interpreted as part of the argument/scope of the subject QNP (in which case we will say that the subject QNP has wide scope).

So we started with a fairly simple mapping between syntax and semantics, (5), which allows QNPs to be interpreted whenever their sister is a one place predicate. However this procedure leaves us with two open questions:

(8) a. How does a QNP find its argument when its sister is not a one place predicate (e.g., when the QNP is generated in object position)?
b. How are arguments determined in constructions that involve multiple quantification so as to account for scopal ambiguities such as the one exemplified in (7)?

The questions in (8) take their particular form as a consequence of the assumptions that we made. It is definitely worthwhile to pursue alternative assumptions and to see whether the questions will subsequently be easier to address. In particular, one could attempt to relax our constrained assumptions about the syntax semantics interface from which (5) follows or to revise our assumptions about the meaning of various constituents.

However, it is also interesting to search for answers that keep to our fairly minimal assumptions about the system. My goal in this paper is to present one such answer and some of the evidence that has been gathered in its support. It is my impression that the evidence is fairly significant, and to the extent that I am correct, not only will the answer be corroborated but also the minimal assumptions that determined the form of the questions.

1. The Relevance of Movement

Consider the sentence in (9). This sentence has two interpretations. Under one interpretation, the speaker of the sentence asserts that he made a certain demand of his addressee, namely the demand that the addressee avoid reading any books. Under the other interpretation, the speaker of the sentence does not assert that he has made any demands of his addressee. All that he asserts is that there is no particular book such that he demanded that the addressee read it.[6]

(9) I demanded that you read not a single book.

Given our assumptions, this ambiguity can be attributed to two possible arguments/scopes that the QNP *not a single book* can take. Let's start with the second interpretation. This interpretation would result if the QNP could take as its argument the predicate which is true of exactly those things that I demanded that you read (λx. I demanded that you read x).[7] The result of the composition is the proposition that is true iff there is not a single book that satisfies the predicate, that is, not a single book that I demanded that you read. The first interpretation would result from *not a single book* taking as its argument the predicate which is true of exactly those things that you read (λx. you read x). The result of the composition is a proposition that is true iff you read not a single book. This proposition, in turn, serves as an argument of the matrix predicate *demanded* resulting in a predicate which is true of exactly those individuals that made a demand that you read not a single book. Finally, this predicate takes *I* as argument resulting in a proposition that is true iff I made the demand that you read not a single book.

[6] See Klima (197?), Kayne (1981). The latter interpretations is forced when *not a single book* is replaced with *no particular book*.

[7] Notice that I am using the terms *I* and *you* to describe the meaning of the sentences (in the "meta language"). This practice would not be a good idea if it were adopted more generally. Sophisticated things need to be said in order to explain the way that first and second person pronouns depend on context to determine their meaning. Also, much more needs to be said about the meaning of verbs such as *demand*. To account for their scopal behavior, it is standard to tread them as quantifiers that range over possible worlds.

This characterization is quite similar to the one that we provided for the ambiguity of the sentence in (7). In (7) there were two interpretations one resulting from the subject QNP taking wide scope over the object QNP and the other resulting from the reversal of the scopal relation. Similarly, here one interpretation results from the QNP, *not a single book*, taking scope over the verb *demand*, and the other resulting from the reversal of this scopal relation. Assuming that this characterization is more or less correct, the same questions arise. What is the procedure that allows the QNP to find its argument given that its sister is not a one place predicate, (8)a? And, why is it that the relevant procedure can apply in two different ways, yielding two different arguments for the QNP, (8)b?

The sentence in (10), which is derived from the sentence in (9) by movement of the QNP, suggests a possible answer.

(10) Not a single book did I demand that you read t.

This sentence is restricted to the interpretation resulting from wide scope for the QNP. It expresses the proposition that there is no particular book that I demanded that you read, and cannot be used to assert that I made the demand that you avoid reading any book.[8]

What could this restriction follow from? To answer this question, let's look at the constituent that serves as sister of the QNP at its landing site (the derived sister of the QNP). This constituent is the IP *I demanded that you read t* (ignoring the effects of *do*-support and I to C movement). How is this IP interpreted? The restriction on the interpretation of the whole sentence would follow, if the IP were interpreted as the predicate *λx. I demanded that you read x*. Let us therefore assume that this is how things work, and state it (somewhat) explicitly in the following way:

[8]Notice, though, that what makes the semantic effects of movement visible is not very well understood. Once a broad array of movement constructions is taken into account, it is clear that a moved constituent can sometimes be interpreted as if it hasn't moved (scope reconstruction). Furthermore, it is not obvious (to me, at least) what blocks scope reconstruction in (10). But there are also arguments (for the rule in (11)) that do not depend on unavailability of scope reconstruction. These arguments rely on environments in which (again for ill-understood reasons) QR – the operation that we will introduce shortly – is more restricted than overt movement (see Fox 2000:144 and references therein). Also, strong evidence comes from the effects of scrambling in scope rigid languages (see Kuroda 1971 and Hoji 1985) or from other languages in which QR might be overt (see note 14).

(11) In a structure formed by DP movement, $DP_n[_{\varphi}\ldots t_n\ldots]$, the derived sister of DP, φ, is interpreted as a function that maps an individual, x, to the meaning of φ[x/n].

φ[x/n] is the result of substituting every constituent with the index n in φ with some NP referring to the individual x.[9]

What we have learned is that a QNP can start out in a position other than the position in which it is interpreted. Movement takes the QNP to a position where its sister is a one place predicate (given the rule in (11)) and it is in this new position that the QNP is interpreted. Can this help us address the questions in (8)? Well, it might, as long as we are willing to entertain the possibility that movement can sometimes occur without visible effects on word order. Such movement, which is commonly referred to as "covert movement" or Quantifier Raising (QR), will allow a QNP to find its arguments when its base generated sister is not a one place predicate, thus providing an answer to (8)a. Similarly, this movement could account for scopal ambiguities, (8)b, under the assumption that more than one possible landing site is available.[10]

Stated somewhat differently, the rule in (11) allows a moved QNP to take its derived sister as its argument/scope.[11] This allows a QNP that is base generated in a position where its sister cannot serve as its argument/scope to move to a position where its argument/scope can be determined. In other words, it is predicted that the scope of a QNP will be its derived sister (i.e., its c-command domain in the structure derived by

[9] According to Sauerland and van Stechow (2000), this rule is due to Cooper (19??), though the idea that the semantic objects formed by this rule should be thought of as arguments of quantifiers goes back to Frege. The statement that I provide in (11) yields undesirable results when certain constructions that involve more than one operation of movement are taken into account. These problems could be circumvented by a constraint against "meaningless coindexation" of the type argued for in Heim (1997). See Heim and Kratzer (1998) for extensive discussion of a different implementation. On other concerns associated with reference to substitution, see Larson and Segal (1995: 7.2) and references therein.

[10] The hypothesis that QNPs move to "their scope position" (Quantifier Raising) is due to Chomsky (1976) and May (1977). It is, of course, related to the hypothesis of Mckawley (1970) that QNPs start out in their scope position and move from there to their surface position (Quantifier Lowering). In fact, the argument in favor of QR based on the Coordinate Structure Constraint (see section 2.1. below) was originally framed in terms of Quantifier Lowering. See also Postal (1974). Quantifier Raising also bears some resemblance to the rule of Quantifying in proposed by Montague. For a recent comparison of the various approaches, see Jacobson (in press).

[11] As long as the result is interpretable. This will be the case as long as the derived sister denotes a one place predicate, i.e. as long as φ[x/n] denotes a truth value.

movement). If movement can be covert, i.e., if QR exists, and if there is some flexibility in landing sites, the questions in (8) would be answered.

To see this in more detail, consider the sentence in (6) repeated below. The problem with this sentence is that in the base generated structure, the sister of the QNP *every tree* (the base generated sister) is not a one place predicate.

(6) I climbed every tree.

However, if we postulate covert movement, this does not have to be true of the structure that is interpreted by the semantic system. The QNP can move covertly, yielding the LF structure in (6)'.

(6)' [every tree]$_1$ [I climbed t_1]

In this new structure the derived sister of the QNP is interpreted by (11) as the predicate that for every individual, x, yields the semantic interpretation of the derived sister, when t_1 is substituted by some name of the individual, x. This derived predicate is *λx. I climbed x,* and, as mentioned in the previous section, this is exactly the predicate we need as an argument of *every tree* in order to derive the correct truth conditions for the sentence.

Let's now go back to the sentence in (9).

(9) I demanded that you read not a single book.

Also in this sentence the QNP cannot be interpreted in its base position, since, again, the base generated sister is not a one place predicate. However, there are two positions to which the QNP can move where its derived sister would be interpreted as a one place predicate (by (11)):

(9)' a. I demanded [that [Not a single book]$_1$ **[you read t_1]**]

b. [Not a single book]$_1$ **[I demanded that you read t_1]**

It is easy to convince oneself that given our assumption in (11), the structures in (9) provide us with the basis to derive the two interpretations. Specifically, the derived sisters in (9)a and (9)b (which are marked in boldface) are interpreted as the one place predicates *λx. you read x* and *λx. I demanded that you read x*, respectively. And as we reasoned above, these are precisely the predicates that are needed to derive the two interpretations.

Finally, consider the scopal ambiguity in (7).

(7) A boy climbed every tree.

Here, just as in (6), the object QNP cannot be interpreted in its base position. However, once again, this problem can be addressed with the aid of covert movement. The object can raise covertly over the subject yielding the structure in (7)':

(7)' [every tree]$_1$ [a boy climbed t_1]

Here the derived sister is interpreted as the predicate *λx. a boy climbed x* and after this predicate is interpreted as the argument of the universal quantifier one of the two interpretations is derived, namely the interpretation that we have called the object-wide-scope interpretation (sometimes also called Inverse Scope). To derive the alternative subject-wide-scope (or Surface Scope) interpretation, we would need to move the subject over the derived position of the object, yielding the representation in (7)''.

(7)'' [A boy]$_2$ [[every tree]$_1$ [t_2 climbed t_1]]

Here the subject QNP takes as its argument the predicate which is true of an individual, y, if the QNP, *every tree*, yields TRUE when it combines with the predicate *λx. N climbed x*, where *N*, is some name for y. In other words, the derived sister of the subject QNP is the predicate which is true of an individual if the individual climbed every tree. Consequently, the sentence ends up being true iff there is a boy who climbed every tree.

So let's see where we are. We ended the previous section with two problems. Given the assumption that a QNP is a second order predicate (i.e., a function from first order predicates to truth values) and given a restrictive theory of the syntax semantics interface (which stated that that the argument of a function must be its syntactic sister), there was (a) no way to interpret a QNP in object position, and (b) no way to account for scope ambiguities. We started this section with the observation that overt movement of a QNP has consequences for the argument/scope of the QNP. We provided an account of this fact, with the aid of the semantic rule in (11). This semantic rule predicts that the derived sister of a moved QNP will serve as its argument/scope. In other words, it is predicted that the scope of a QNP is its c-command domain at LF. We then noticed that the puzzles in (8) can be resolved under the hypothesis that syntax allows movement that has consequences for LF but no consequences for phonology (covert movement). Our goal in the next section is to introduce additional considerations that might bear on the truth of the hypothesis.

But before we move on, I would like to make a short digression and say a few words about the hypothesis that subjects always move to a position higher than their base position. We will see that this hypothesis, known as the VP internal subject hypothesis (VPI), is relevant for determining some of the consequences of covert movement. Consider, again, the sentence in (7), *a boy climbed every tree*. Given our assumptions, the object QNP can be interpreted only if it moves to a position c-commanding the subject position:

(7)' [every tree]$_1$ [a boy climbed t_1]

This, as we have said, yields the Inverse Scope interpretation. In order to get the Surface Scope interpretation, the subject must move over the derived position of the object, yielding the representation in (7)''.

(7)'' [A boy]$_2$ [[every tree]$_1$ [t_2 climbed t_1]]

But notice that if we assume VPI, the subject moves over the derived position of the object independently of scope considerations. This means that (7)', which does not include subject raising, is not a possible LF.[12] Given VPI, the two scopal relationships would result from the following LFs:

(7)''' a. Surface Scope

$[_{IP}$ [A boy]$_2$ $[_{VP}$ [every tree]$_1$ $[_{VP}$ t_2 climbed t_1]]]

b. Inverse Scope

$[_{IP}$ [every tree]$_1$ $[_{IP}$ [A boy]$_2$ $[_{VP}$ t_2 climbed t_1]]]

The important piece of this discussion is that our semantic framework in conjunction with VPI allows objects to be interpreted in a VP internal position. This position, which is higher than the base position of the subject but lower than its surface position, will play an important role in some of the discussion that follows.

2. Arguments in favor of Covert Movement

A principle such as (11) needs to be incorporated into our semantic system in order to account for the interpretation of constructions involving overt movement. Once the principle is in place, it turns out that certain correlations between sound and meaning are explained by the system as is only if QR is part of grammar. However, one could imagine other ways to enrich the system in order to account for the relevant correlations. Indeed, the literature is full of competing proposals.

I will not attempt to evaluate these competitors. Instead I will try to go through various arguments that have been presented in favor of the idea that QR is the correct enrichment. This decision is due to space limitations, and is a significant shortcoming since arguments in favor of a proposal cannot be fully appreciated without an understanding of the nature of potential competitors. Nevertheless, I hope to be able to

[12] Notice that this is not the case once scope reconstruction is acknowledged (see note 8). See Hornstein (1995) and Johnson and Tomioka (1997) for an interesting exploitation of this observation. See also notes

present a good picture of what has been achieved by the postulation of QR and of the challenges that ought to be considered when evaluating alternatives.[13]

How can we tell whether a certain operation takes place in the syntax? Well, we have to see whether there are any predicted consequences. Predictions, however, won't follow from the operation alone but only in conjunction with various other theoretical assumptions. We, therefore, have to search for various components of the theory that will yield predictions when combined with the operation. There are at least three components that could be relevant: phonology, semantics and syntax. The claim that QR is a covert operation is essentially the claim that there are no predicated consequences that follow from phonology. So we need to ask whether there are predicted consequences given properties of semantics or syntax. We have seen in the previous section that certain consequences are predicted from the rules of semantics. Specifically, given the rule in (11), QR predicts that object QNPs will be interpretable and that scopal ambiguities will sometimes be attested. But the question we are asking now is whether there are further predictions.

So what we need to do is to identify additional principles of syntax or semantics that could tell us whether or not movement has taken place when there are no consequences for phonology. Two types of principles come to mind. First there are various constraints on movement that can serve as "movement detectors": if it turns out that scope is restricted by the movement constraints, this will suggest that scope is determined by movement. Second, there are various constraints on structure that can be taken as "structure diagnostics". We can use these diagnostics to see whether the output of QR provides the correct structural description of various sentences.[14]

2.1. Properties of Movement

---

17 and 29.

[13] See Hendriks 1993, Jacobson 1994, Brody 1995, Williams 1986, and Barker 2001.

[14] I will not discuss the argument in favor of QR presented in Chomsky (1976), which was based on Weak Crossover. I will also not present various argument that based on languages in which Quantifier Raising might be overt. Among the relevant references are Kiss (1990), Szabolcsi (1998), Pesetsky (1987, 2000) Hoji (1985), Huang (1982), and Johnson (2000).

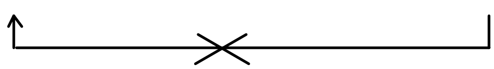
The claim that the scope of a QNP is determined by a (covert) movement operation leads to the expectation that constraints on movement will reveal themselves as constrains on scope. If this expectation turns out to be correct, it can be taken as evidence in favor of a theory that incorporates QR.

One case where the predictions of the theory are corroborated is the case of movement out of one of two coordinated conjuncts. Such movement is impossible as demonstrated in (12).

(12) a. Which professor does John like t?
b. *Which professor does John [[like t] and [hate the dean]]?

This constraint (the Coordinate Structure Constraint, CSC, Ross 1967) can be used to test the predictions of a theory that postulates covert movement. The predictions are corroborated by the following contrast noted in Lakoff (1970) and Rodman (1976):

(13) a. A (different) student likes every professor. $(\exists > \forall)$ $(\forall > \exists)$
b. A (#different) student [[likes every professor] and [hates the dean]]



$(\exists > \forall)$ $*(\forall > \exists)$

(13)a is a case of scope ambiguity that receives the analysis provided for (7) above. (13)b, by contrast, is restricted to the interpretation in which the object QNP is interpreted within the scope of the subject QNP: the sentence can express the proposition that there is a student who likes all of the professors and hates the dean, but not the proposition that for every professor there is a (perhaps different) student who likes the professor and hates the dean.[15] This restriction is a direct consequence of a theory that incorporates QR. Under such a theory, Inverse Scope requires the object to move by QR over the surface position of the subject, and this movement violates the CSC.

---

[15] The use of the adjective, *different*, makes the difference between surface and inverse scope fairly transparent.

Of course, a theory that incorporates QR must also explain how Surface Scope is possible in (13)b. As noted in May (1985), this requires the assumption that QNPs can be interpreted within the VP. As mentioned, this assumption is a direct consequence of the rule in (11) once VPI is adopted. To see this in more detail, let's go over the derivation of the sentence, starting at a point at which two VPs are conjoined, (14)a. At this point, the object moves (covertly) by QR to a position internal to the first conjunct where it can be interpreted, (14)b. Since this movement does not cross a coordinated structure, the CSC is not violated. Finally, the subject moves (overtly) "Across the Board" to its surface position, [Spec, IP] yielding the LF representation in (14)c.[16]

(14) a. [$_{VP}$[$_{VP}$ a student likes every professor] and [$_{VP}$ a student hates the dean]]

b. [$_{VP}$ [$_{VP}$ **[every professor]$_1$** [$_{VP}$ a student likes t_1]] and
[$_{VP}$ a student hates the dean]]

c. [$_{IP}$ **[A student]$_2$** [$_{VP}$ [$_{VP}$ [every professor]$_1$ [$_{VP}$ t_2 likes t_1]] and
[$_{VP}$ t_2 hates the dean]]]

This is the representation of Surface Scope since the sister of the subject QNP (its argument/scope) contains the object QNP. To get the Inverse Scope interpretation the subject (in [spec, IP]) must be dominated by the derived sister of the object. The only way that this could be achieved is if the object were to move out of the coordinate structure violating the CSC.[17]

The observation that scope obeys the CSC serves as an important piece of evidence in favor of a theory that incorporates QR.[18] One would like to ask whether scope obeys other constraints on movement. Various people have argued that it does, but unfortunately I cannot go over the arguments in this context. Among the important references are Huang (1982),May (1977, 1985), Kayne (1981), Pesetsky (1987, 2000), Reinhart (1991, 1997), Aoun and Li (1993), and Richards (2001). If these arguments are

[16] This derivation is cyclic and as such is incompatible with models of grammar in which covert operations must "follow" all overt operations. This incompatibility is, of course, not significant; the relevant structures could be derived in two different cycles; one overt and one covert. But see section 4.
[17] One might wonder whether scope reconstruction (cf. notes 8 and 12) could circumvent the effects of the CSC. See Fox (2000:58-62).

successful, we might be able to conclude that all movement constraints indicate that QR should be dealt with by covert movement.

However, the situation is complicated by various factors, and there are many open questions for further research. For example, it has been argued that QR is more restricted than various forms of overt movement. But these arguments have received different interpretations, in part due to the fact that there are various restrictions that apply to some forms of overt movement but not to others. (Relevant references include Wilder (1996), Hornstien (1995), Kennedy (1997), Johnson and Tomioka (1997), Johnson (2000), Bruning (2001), Sauerland (2001x), and Fox (2000:2.4).) Furthermore, some researchers have argued that the variability in overt movement is mirrored by a variability in covert movement (see, among others, Beghelli and Stowell (1998), Szabolcsi (1998) and Aguero-Bautista (2001)). Finally, some have suggested that in certain environments QR is more flexible than overt movement and have attempted to provide an explanation for this fact (see Reinhart (1991) and Fox (2000:2.4)).

These complications are inevitable given that the constraints on movement are not fully understood. Nevertheless, the fact that QR is demonstrably sensitive to certain specific constraints such as the CSC is encouraging. Still one would like to identify tests for QR that are based on better grounded auxiliary assumptions. As mentioned, one thing we can do is search for constraints on structure that can tell us whether or not the output of QR provides the correct structural description for various sentences. One such constraint, which we will use as a "structure diagnostic" in the next section, is a constraint on VP deletion, sometimes called Parallelism.

2.2. Parallelism as a "structure diagnostic"

The second sentence in (15) is missing a phonologically visible VP. Nevertheless, the sentence receives an interpretation. Specifically, it is interpreted as if it contained a VP identical to the VP of the preceding sentence. (I.e., the interpretation is identical to that of *Then, Bill talked to Mary*.)

---

[18] Ruys (1992) observes that the CSC applies somewhat differently in the case of overt and covert movement. See Lin (2001) for arguments that overt A-movement patterns with covert movement.

.

(15) First, John talked to Mary. Then, Bill did.

This interpretation is the only interpretation available: the second sentence cannot be interpreted as if it contained a VP which was different in interpretation from the preceding VP (at least as long the discourse does not contain additional sentences).

A fairly simple way to characterize the situation would be to say that it's possible for a VP not to be pronounced but that this possibility is constrained by an identity condition:

(16) **Parallelism**: VP_1 can be deleted only if the discourse contains a pronounced VP_2 (the antecedent VP) such that VP_2 is syntactically identical to VP_1.[19]

In light of Parallelism, consider the sentences in (17) (in which the deleted VPs are written inside angel brackets, <deleted VP>).

(17) a. I [$_{\text{VPantecedent}}$ read [a book [that you did <read t>]]].
b. I [$_{\text{VPantecedent}}$ gave [every book [that you did <give t to Mary>]] to Mary].
c. I [$_{\text{VPantecedent}}$ wanted to be in [the city [that you did <wanted to be in t>]]]

It appears that in these sentences the deleted VP is contained within its antecedent. This is reflected in the term commonly used to refer to the type of deletion involved: Antecedent Contained Deletion (ACD). But if Parallelism is to be satisfied, there can be no antecedent containment; the antecedent VP must be identical to the elided VP, but a syntactic constituent cannot be identical to one of its sub-constituents.[20] Therefore, if Parallelism holds, the deleted VP can't contain the elided VP.

---

[19] I haven't said what it means for two VPs to be identical. There are very difficult questions relating to the way the condition should apply to traces and pronouns that I cannot go over in this context. What I have in mind is something along the lines of the theory of Parallelism in Rooth (1992) as it is outlined in Heim (1997).

[20] Since syntactic constituents are finite.

So the sentences in (17) cannot have the structures hinted at by the square brackets. What is their structure then? QR provides a possible answer. More specifically QR yields a structure in which the antecedent and the elided VP are identical:

(17)' a. [a book [that you did <read t>]]

I [$_{\text{VPantecedent}}$ read t]

d. [every book [that you did <gave t to Mary>]]

I [$_{\text{VPantecedent}}$ gave t to Mary].

e. [the city [that you did <wanted to be in t>]]

I [$_{\text{VPantecedent}}$ wanted to be in t]

QR thus solves the problem that is posed by apparent cases of antecedent containment. Parallelism predicts that antecedent containment should be impossible.[21] If QR is not part of our theory, we have to posit structures that violate Parallelism, obviously not a good result. However, if we incorporate QR into the theory, the structures do not violate Parallelism, and the problem is eliminated.

Stated somewhat differently, assuming that Parallelism is established on independent grounds, we can use it as a “structure diagnostic”. This diagnostic indicates that the output of QR yields a possible structural description for the sentences in (17). On the other hand, it is not obvious how the structural description might satisfy Parallelism if QR is not postulated.[22]

But the argument can be made even stronger. Right now QR has several independent effects in our system. If it turns out that these independent effects correlate with each other in non-trivial ways, it will suggest that we are on to something real. One significant correlation was pointed out in Sag (1976) (and discussed in Larson and May (1990) as well as much subsequent work). Given that QR yields a larger argument (or wider scope) for a QNP, and given that QR is necessary if ACD is to satisfy Parallelism, it is predicted

[21] Notice that the predictions of Parallelism go beyond the exclusion of antecedent containment. So it is not enough to postulate a structure for the sentences where antecedent containment does not exist; one has to make sure that the postulated structure satisfies Parallelism.

[22] There are, however, alternative proposals. One type of proposal involves a relaxation of Parallelism (Vanden Wyngaerd and Zwart 1991 and Brody 1995). Another involves the assumption that the

that the two would correlate. More specifically it is predicted that the argument/scope of a QNP in an ACD construction would have to contain the antecedent VP. To see that this prediction is made consider the representation that an ACD construction has "before" QR:

(18) …[$_{\text{VPAntecedent}}$ …[$_{\text{QNP}}$…<deleted VP>]…]…

In order for the construction to satisfy Parallelism the QNP must move by QR and vacate the antecedent VP. This movement will place the QNP in some position where its derived sister contains the antecedent VP:

(18)' [$_{\text{QNP}}$…<deleted VP>…][$_{\text{DerivedSister}}$…[$_{\text{VPAntecedent}}$ …t…]…]

This derived sister will be interpreted as the argument/scope of the QNP by the rule in (11), and since the antecedent VP is part of the Derived Sister, the resulting interpretation is the one we describe as involving wide scope for the QNP.

The prediction has been tested and corroborated in various ways. But they all share the following property: they involve a sentence with a QNP that can be interpreted either insider or outside of a certain VP; when this VP serves as the antecedent in an ACD construction (and when the deleted VP is dominated by the QNP), it is predicted that only the second option would be available.

Consider, from this prospective, the sentence in (19).

(19) John refused PRO to read every book that I recommended.

Given our assumptions, there are at least two positions where the QNP can be interpreted. One position would result from QR to the matrix IP and the other from QR to the embedded IP:[23]

---

constituant that is elided is not a traditional VP but something else (Jacobson 1994). Space limitations do not allow me to discuss these alternatives (but see note 24).

[23] Given the VPI and the consequence that a QNP can be interpreted in VP, there are other positions that would yield the same interpretation, namely the matrix and embedded VP. The argument would not be

(19)' a. [[every book that I recommended]$_1$ [John refused PRO to read t_1]].

b. John refused [[every book that I recommended]$_1$ [PRO to read t_1]].

Two different interpretations result from the two syntactic options (given the rule in (11)). If the QNP targets the matrix IP (wide scope for the QNP), as in (19)'a, the sentence asserts that every book that I recommended is a book that John refused to read (i.e., a book that the derived predicate *λx.John refused to read x* is true of). Under the alternative structure (19)'b, in which the derived sister of the QNP (the scope) is smaller (narrow scope for the QNP), the sentence asserts that John refused to the suggestion that he read all of the books that I recommended.

So we have an ambiguity here that can be used to test the prediction we characterized abstractly above. We have a QNP that can be interpreted either inside or outside of a certain VP ((19)'b or (19)'a), and, as mentioned, it is predicted that only the second option is available when the VP serves as the antecedent in an ACD construction.

Consider the following.

(20) I [refused to read [every book [that you did <read t>/<refuse to read t>]]].

In this sentence there are two interpretations for the elided VP. It could be interpreted with the embedded or matrix VP as antecedent. (I.e., the elided VP could be *read* or *refuse to read*.) Our prediction is that the matrix VP can be the antecedent, only if the QNP is interpreted outside of this VP. (This prediction follows since the QNP must move by QR outside of the antecedent VP for Parallelism to be satisfied.) I.e., we predict that the QNP must outscope the verb *refuse,* (19)'a, if the elided VP is *refuse to read*. To test whether this is in fact the case, one has to access fairly complicated judgments. However, the results are pretty clear when an appropriate experiment is conducted.

So consider the following:

affected if these other positions are taken to be the landing sites for QR. This better be the case since in section 3 we will argue that VP is the common landing for a QNP.

(21) a. John refused to read EVERY book that we thought he HAD <read>. He was willing to read only some of them.

b. John refused to read EVERY book that we thought he would <refused to read>. #He was willing to read only some of them.

Both (21)a and (21)b consist of two sentences uttered in succession. The second sentence ensures that the first sentence is understood with narrow scope for the QNP. (If there are books that John was willing to read than surely it's not the case that every book is such that he refused to read it.) It is a clear judgment that the two sentences contradict each other in (21)b but not in (21)a. This indicates that the latter is restricted to the wide scope interpretation for the QNP, just as the theory predicts.

We started this subsection with a condition on the licensing of VP deletion, Parallelism. This condition was taken as a structure diagnostic, and as such it indicated that apparent cases of antecedent containment cannot be real cases of antecedent containment. If they really involved antecedent containment, there would be no way for Parallelism to be satisfied. This observation supported QR, which eliminates antecedent containment and provides structural descriptions that satisfy parallelism. Furthermore, QR predicted a non-trivial yet correct correlation between scope and the size of the elided VP. This suggested that the conflict between Parallelism and antecedent containment is a real one, and that it is resolved by the mechanism that accounts for scope.[24]

2.3. Binding Theory as a "structure diagnostic"

In the previous sub-section we have seen that Parallelism (taken as a structure diagnostic) indicates that the structures derived by QR are real. In this sub-section, we will see that

[24] This conclusion is shared by some researchers who do not agree that QR is the relevant mechanism. Thus Jacobson's (1994) mechanism for ACD resolution is not movement but a mechanism that has consequences for scope (function composition). Furthermore, it is a mechanism she uses (following many others) to account for cases of overt displacement. (Jacobson's position, if translated to a theory that postulates movement – a translation that she would strongly object to –, would be the position that ACD involves overt movement to the right. In this sense her position is rather close to that of Kayne (1998).) Similarly, for Brody ACD resolution, although not dependent on movement, is dependent on a pattern of "scope marking" that derives the correlation with scope.

Binding Theory appears to provide the opposite indication. This conflict will serve as the basis for the discussion in the following sections.

Consider the sentences in (22).

(22) a. ??John and Bill said that [Mary bought every picture of each other/themselves].

b. *He$_1$ likes every picture that John$_1$ saw.

The fact that (22)a is unacceptable should follow from the locality conditions that restrict the distribution of anaphoric expressions such as *himself* and *each other*. Similarly, the unacceptability of (22)b, on the intended interpretation, should follow from restrictions on co-reference between names and pronouns. This result is achieved by the following two conditions of Binding Theory:

(23) a. Condition A: An anaphoric expression (*himself*, *each other*) must co-refer with (or be bound by) a locally c-commanding antecedent.[25]

b. Condition C: A name and a pronoun cannot co-refer if the pronoun c-commands the name.

The unacceptability of (22)a follows if the matrix subject is outside the local domain for the binding of the anaphor, and the unacceptability of (22)b follows under the assumption that the subject c-commands the object.

The problem, however, is that the two assumptions are far from obvious the moment QR is postulated. To see this, consider the cases of overt movement in (24).

(24) a. John and Bill know [[which picture of each other/themselves] Mary bought t].

b. (Guess) [which picture that John$_1$ saw] he$_1$ likes t.

---

[25] I wish to ignore the question of whether Condition A is satisfied by co-reference or requires variable binding as argued for in Reinhart (1983). If Reinhart is correct, we will need to be a little more careful in the way we re-state Belletti and Rizzi's idea (see p. 30).

Once the *wh*-phrase moves over the embedded subject in (24)a, the anaphor is close enough to its antecedent for Condition A to be satisfied. Similarly, in (24)b, Condition C is satisfied since *wh*-movement removes the name out of the c-command domain of the pronoun.

A theory that postulates QR yields the structures in (22)' as possible LF representations of the sentences in (22). Consequently, it predicts incorrectly that these sentences, just like the sentences in (24), will escape violations of Conditions A and C.[26]

(22)' a. J & B said that [[every picture of each other/themselves] Mary bought t].
b. [Every picture that John_1 saw] He_1 likes t.

3. Binding Theory and QR – some recent research

In section 1 we have seen that QR together with the rule in (11) provides an explanation for various semantic properties of sentences that contain QNPs. Furthermore, we have seen that the explanation is supported by Parallelism and various locality conditions on movement (2.1., 2.2). However, we have seen in 2.3. that Binding Theory poses a challenge for the explanation.

In order to deal with this challenge, Chomsky (1981) suggested that Binding Theory applies at a level of representation at which the effects of QR are not visible (Surface Structure). However, in later work (Chomsky 1993) he has criticized this suggestion on various grounds. I will not attempt to summarize the criticism. Rather, I would like to point out that given our limited perspective (namely the study of the syntactic position of QNPs), the appeal to Surface Structure is a copout (see note 26). There is conflicting evidence with respect to the syntactic position of QNPs, and the postulation of two syntactic representation each corresponding to a subset of the data is not much more than a restatement of the problem.[27] We would like to have one syntactic representation that

[26] Williams (1986) and more recently Brody (1995) pointed out, correctly in my view, that the contrast between (24) and (22) argues against the postulation of QR. However, the facts discussed in the remainder of the paper might reverse the argument.

[27] This is one of the points that Chomsky (1993) makes. But he argues that the appeal to Surface Structure is a copout from a broader perspective as well. He also provides some empirical evidence that Binding Theory must apply at LF. This evidence has been supported by the discovery of a correlation between

captures all of the data. Following Chomsky's lead, various researchers have attempted to articulate a theory that would derive such a syntactic representation. In what follows, I will focus on a few suggestions I have made in recent work (Fox 2000, in press).[28] If there is something to these suggestions, it will turn out that (contrary to initial appearances) Binding Theory is a pretty reliable diagnostic of LF structures and as such it provides further evidence for a theory that incorporates QR.

3.1.Condition A and the position of QNPs

Let's go back to (22)a, repeated below.

(22)a. ??John and BIll said that [Mary bought every picture of each other/themselves].

The unacceptability of this sentence is problematic for a theory that incorporates QR since under such a theory the embedded object QNP should be allowed to move covertly yielding a representation identical (in all relevant respects) to the representation of (24)a. Since in the latter representation Condition A is not violated, a theory that incorporates QR is not able to account for the unacceptability of (22)a. This conclusion follows under the assumption that the distribution of QR is identical to the distribution of *wh*-movement. This assumption in turn has been taken to be the null hypothesis, and has served as the basis for some of the arguments in favor of QR. However, it is not a necessary assumption (see p. 16 and note 29). It is possible that there are some constraints that apply to QR and do not apply to *wh*-movement. If that can be demonstrated on independent grounds, there might be a principled account for the unacceptability of (22)a.

In Fox (2000) I have argued that such a constraint exists. More interestingly (for this context), I have tried to show that in environments where the constraint allows QR to

---

scope reconstruction and condition C (Heycock 1995, Lebeaux 1990). For discussion, extensions, and various additional references see Fox (2000:chapter 5), Romero (1997), and Sportiche (2001).

It is also important to note that the claim that Biding theory applies at Surface Structure is an incoherent position if the approach to Condition C advocated in Reinhart (1983) is adopted.

[28] The latter draws heavily on Fox and Nissenbaum (1999). For alternative proposals, see Lasnik (1993) and Brody (1995).

apply, the covert operation has the expected consequences for Condition A (Fox 2000: 198-199). If these arguments are successful, it will turn out that (contrary to initial impressions) Condition A can serve to further motivate a theory that incorporates QR.[29]

The evidence in favor of a constraint that applies specifically to QR is based on a variety of considerations. Here I will present just one piece of data concerning the scope of QNPs in constructions that involve VP ellipsis. Consider the pair of sentences in (25).

(25) A boy admires every teacher. A girl does, too <admire every teacher>.[30]

These sentences show a scopal ambiguity that we are by now very much familiar with. What I would like to focus on, though, is that the relative scope of the object and the subject is identical in the two sentences: if the object outscopes the subject in one of the sentences the subject cannot outscope the object in the other sentence. The way Parallelism was defined in section 2.2. is not sufficient to account for this demand for parallel interpretation, but it is pretty clear that parallel scope must be a consequence of the theory of ellipsis (See Fox 2000 chapter 3 for a discussion of the necessary properties of the theory of ellipsis.)

Now consider the pair of sentences in (26).

(26) A boy admires every teacher. Mary does, too <admire every teacher>

The first of these sentences is restricted to subject-wide-scope. (It can only express the proposition that there is a boy who admires all of the teachers; i.e., it is false if each of the

[29] It is important to point out that other researchers have provided evidence that QR is more restricted than *wh*-movement. In particular, Johnson and Tomioka 1997 and Hornstein 1995 have suggested that a QNP can never move over the surface position of the subject. These theories can account for the paradigm I present below (under certain assumptions about the Local Domain for Condition A) if the constraint in (27) is viewed as a constraint on the reconstruction of the subject. Hornstein, in fact suggested such a constraint on independent grounds (related to the theory of Diesing 1992). For a few challenges to the claim that QR depends on reconstruction, see Fox (2000: 44-46, 51-62).
[30] In Sag (1976) and Williams (1977) it was assumed that a QNP cannot take scope outside of an elided VP. However, this assumption was based on examples such as (26), and has been challenged by ellipsis sentences in Hirschbühler (1982) similar to (25):

(i) An American flag is hanging from every window, and a Canadian flag is, too.

teachers is admired by a different boy.) This restriction would follow from Parallelism if there were an independent condition that blocked an application of QR in the second sentence in a way that would reverse the c-command relation between the subject and the object.

I will, therefore, assume that the independent condition exists. For current purposes it is not important to provide the correct characterization of the condition, and I will therefore keep to a statement of what must its consequences be if the facts in (26) are to follow from Parallelism:[31]

(27) An object QNP can move over the subject when the subject is an indefinite but not when the subject is a name.

If (27) holds, (22)a is no longer a problem. The object QNP must move by QR to a position where it can be interpreted, but given (27) this position must be below the embedded subject. The two requirements can be satisfied given VPI, and, assuming an appropriate definition of locality, Condition A of the binding theory would still be violated; The embedded subject still intervenes between the anaphor and its antecedent, thereby violating the specified subject condition or other characterizations of the relevant constraint.

(22)'a. ??John and BIll said that $[_{IP}$ Mary$_1$ $[_{VP}$ [every picture of each other/themselves]$_2$ [t_1 bought t_2]]].

But more interestingly, we predict that the violation of Condition A that we've seen in (22)a could be circumvented if the embedded subject, the name *Mary*, were replaced with an indefinite. This seems to be the case as the following contrast indicates:

For discussion and further references, see Fox (2000).

[31] In Fox (2000) I argued that the condition is more general favoring the shortest instance of QR needed for a given semantic interpretation. I also tried to make sense of the fact that the condition applies to QR and not to overt movement. But see Tomioka (2001).

(28) a. ??The two rivals hoped that Bill would hurt (every one of) each-other's operations.

b. The two rivals hoped that someone would hurt (every one of) each-other's operations. *$\exists > \forall$ $\forall > \exists$

(28)a, just like (22)a, is a violation of Condition A resulting from an LF representation in which the name *Bill* intervenes between the anaphor and its antecedent. This intervention cannot be circumvented by QR given the constraint in (27). However, in (28)b the situation is different since the intervening subject is an indefinite, which QR is allowed to cross.

An additional interesting prediction is that (28)b would be acceptable only under the interpretation resulting from Inverse Scope in the embedded sentence. To understand the nature of the prediction, we start with the observation that (28)b is a proposition about the hopes that the two the rivals have. Under the Inverse Scope interpretation of the embedded clause these hopes would be satisfied if every one of the rivals operations are hurt. Under the Surface Scope interpretation more needs to happen for the hopes to be satisfied, namely, the same person must hurt every one of the relevant operations.

Speakers I've consulted seem to feel that only the former interpretation is available. But this is a very delicate judgment (for reasons discussed in Abush 1994, among others). David Pesetsky (personal communication) proposed the following as a way to verify the prediction.

(29) a. The two friends hoped that someone would buy each-other's pictures of Mary.

b. *The two friends hoped that someone$_i$ would buy each-other's pictures of himself$_i$.

c. *The two friends hoped that someone$_i$ would buy each-other's pictures of his$_i$ mother.

(29)a, like (28)b, is a case where QR can circumvent a violation of Condition A. In (29)b and (29)c, however, the embedded subject binds a variable dominated by the embedded object. Since (for obvious reasons) a QNP can only bind a variable within its scope, the intended variable binding would be impossible if the object were to QR over the embedded subject. It follows that in (29)b and (29)c QR cannot circumvent a violation of Condition A.

From all of this it seems reasonable to conclude that Condition A supports a theory that incorporates QR. More specifically, it turns out that Condition A indicates that QR applies in exactly those circumstances that we expect it to (exactly those in which an independently motivated structure diagnostics such as Parallelism indicates that it does).[32] Furthermore, there is a predicted correlation between scope and Condition A, which seems to hold (if the paradigm in (29) is indicative).[33]

3.2.The problem with Condition C

We have seen that Condition A, which at first appears to threaten the hypothesis that QR plays a role in deriving LF representations, turns out, under careful scrutiny, to support the hypothesis. What about Condition C? We mentioned above that the hypothesis might

---

[32] As mentioned there are other pieces of evidence converging on the same results discussed in Fox (2000), e.g., the CSC. Note, however, that there is a potential problem here. The following ACD construction doesn't seem as bad as expected given the unacceptability of (22)a:

(i) John said that you were on every committee that Bill did <say that you were on>. (Wilder 1997)

It is interesting to see what happens in constructions that combine ACD and Condition A. The judgments are at the moment too subtle to support any conclusion:

(ii) (??) John and Bill said that I was in every picture of each other that you did.

If turns out that (ii) is unacceptable, one might consider the possibility that the universal quantifier can QR to the matrix clause but that the relevant constraint forces the embedded subject to move to a higher position yielding a violation of Condition A.

[33] The contrast in (i) is similar to the contrast between (22)a and (24)a, and it raises (more or less) the same theoretical dilemma. It has led some (e.g., Chomsky 1981) to assume that Condition A applies at Surface Structure and others to assume that there is no covert *wh*-movement (e.g., Chomsky 1993 and Brody 1995). But Nissenbaum (2001) provides an intriguing third possibility. He provides various arguments that the contrast follows from a locality condition on a second instance of *wh*-movement (argued for in Richards 2001 on independent grounds). Under Nissenbaum's analysis, the paradigm turns out to provide additional support for covert movement, very similar in its nature to the support that is provided by the paradigm in (29).

(i) a. John and Bill know which picture of each other Mary bought.
b. *. John and Bill know who bought which picture of each other.

be threatened by the unacceptability of (22)b (repeated below), which is unexpected if QR could derive the LF representation in (22)’b.

(22) b. *He$_1$ likes every picture that John$_1$ saw.

(22)’b. [Every picture that John$_1$ saw] He$_1$ likes t

But given our constraint in (27), (22)’b is not a possible LF and the problem is eliminated. More specifically, the only LF structure that is available for (22)b is one in which the object QNP adjoins to VP as in (22)’’b. This LF structure still violates Condition C, and the unacceptability of the sentence is thereby explained:

(22)’b. He$_1$ [Every picture that John$_1$ saw]$_2$ t$_1$ likes t$_2$

However, things are not that simple. To see this, consider the sentence in (30)a. The constraint in (27) allows the dative object to move by QR over the subject, and it is thus incorrectly predicted that the sentence would not violate Condition C. One might consider fiddling with the constraint in (27) so that QR would not be allowed, but that doesn’t seem very plausible given the availability of Inverse scope in (30)b.[34]

(30) a. *Someone introduced him$_1$ to every friend of John’s$_1$.
b. Someone introduced John$_1$ to every friend of his$_1$.

So this is a genuine puzzle. It is clear that QR can move the dative object over the pronoun in (30)a, yielding the structure in (30)’a. Yet this movement has no consequences for Condition C.

(30)’a. * [every friend of John’s$_1$]$_2$ Someone introduced him$_1$ to t$_2$

How might we resolve this puzzle? We might want to appeal to Chomsky's (1993) suggestion that movement is a copying operation (the copy theory of movement), and is thereby never capable of circumventing Condition C. More specifically, the output of QR, under the copy theory of movement, would be the structure in (30)''a, in which Condition C is violated due to the copy of the name *John*, which is not eliminated from the base position.[35, 36]

(30)''a.* [every friend of John's$_1$]$_2$ Someone introduced him$_1$ to [every friend of John's$_1$]$_2$

But this suggestion seems incompatible with some of the conclusions we have reached up to this point. First, consider our conclusion that movement (both overt and covert) is capable of obviating a violation of Condition A (see (22)a, (24)a, and (29)a). We would be unable to understand this conclusion if a copy of an anaphor were to yield a violation of Condition A in the same way that a copy of a name yields a violation of Condition C. This is a problem that can be dealt with by an adaptation (with minor adjustments) of a proposal made in Belletti and Rizzi (1988).

Belletti and Rizzi argue that Binding Theory is sensitive to the structure of a sentence at every level of representation, and that our problem can be eliminated by a particular statement of the relevant type of sensitivity: Condition A requires an anaphor to be "bound" appropriately at *some* level of representation, whereas Condition C demands that a name be "free" at *every* level of representation.[37] Assuming the copy theory of movement, we can adopt this idea while eliminating reference to levels of representation

---

[34] A possible way out of this puzzle, which I will not discuss, invokes the hypothesis that the LF position of the pronoun is higher than the position of the two QNPs. This hypothesis is challenged by the arguments presented in section 4 (see the discussion of (39) below).

[35] Section 4.2. Discusses various ways in which the semantic component could treat the syntactic objects derived under the copy theory of movement. This discussion, I believe, is relevant for Condition C. See pp. 45-47.

[36] Of course, only one copy is realized phonologically. In the case of overt movement it's the copy at the higher position; in the case of covert movement it's the lower copy. The principles that determine which copy is pronounced (i.e., whether the movement is overt or covert) are not relevant at the moment, but they will be. Very shortly.

[37] It is sometimes suggested that that distinction would be more natural if we distinguish between "positive" and "negative" conditions. Under such a distinction, we might say that Condition A is a *positive* condition which is *satisfied* by the existence of a level of representation that meets the *required* structural configuration. By contrast, Condition C is a *negative* condition which is *violated* by the existence of a level of representation that meets the *prohibited* structural configuration. See also Lebeaux (1995).

other than LF (see Brody 1995). Specifically, we will say that *one* copy of an anaphor must meet the locality condition in (23)a, while *every* copy of a *name* must meet the anti-co-reference condition in (23)b. Movement has effects on Condition A because it puts *one* copy of an anaphor in a local relation with what would otherwise be a non-local antecedent. By contrast, movement has no affects on condition C since (given the copy theory of movement) it does not put *every* copy of a name outside the c-command domain of a pronoun (which dominates the base position).

A more direct challenge comes from the observation made in (24)b that overt movement is capable of circumventing a violation of Condition C:

(24)b (Guess) [which picture that $John_1$ saw] He_1 likes t.

Under the copy theory of movement, the representation of (24)b should be that in (24)'b, and the sentence is therefore expected, incorrectly, to violate Condition C.

(24)'b (Guess) [which picture that $John_1$ saw] He_1 likes [which picture that $John_1$ saw].

In order to deal with this challenge, Chomsky (1993) adopts Lebeaux's (1989) suggestion that relative clauses (and more generally adjuncts) can be merged ("counter-cyclically") with an NP after a DP that dominates the NP has undergone movement (late merger). (See also van Riemsdijk and Williams 1981 and Freidin 1987.) (24)b can thus have the following derivation, which results in a representation that does not violate Condition C.

(31) a. He_1 likes [which picture] --*wh*-movement→
b. [which picture] He_1 likes [which picture] –adjunct merger→
c. [which picture that $John_1$ saw] He_1 likes [which picture]

So this type of derivation can account for the acceptability of (24)b. But might it disrupt the account that the copy theory of movement provides for the unacceptability of (30)a?

See Chomsky (1993) for an alternative approache.

(30)a *Someone introduced him_1 to every friend of $John's_1$.

There are two potential reasons that we might appeal to in order to maintain the account. One reason would rely on certain assumptions about the architecture of the system that block late merger to an NP after the NP has been moved covertly. Under these assumptions, covert movement applies after "spell-out" and consequently after the merger of all constituents that need to be pronounced. It follows that covert movement cannot feed late merger of a constituent that dominates a name, and subsequently that a derivation similar to the one in (31) would be impossible if it involved QR (instead of *wh*-movement). The second reason would rely on Lebeaux's (1989) hypothesis that late merger is restricted to adjuncts.[38] In (30)a (repeated below) there is no adjunct within the dative QNP that might be a candidate for late merger. Hence, even if covert movement could feed late merger, there would be no conceivable derivation for the sentence that would circumvent the violation of Condition C.

These two lines of reasoning can explain why late merger, which accounts for the acceptability of (24)b, cannot extend to covert movement in a way that would threaten the account provided for the unacceptability of (30)a.

There are various questions that one should address in order to see which, if any, of these two lines of reasoning might be correct. We will get to some of this in section 4.1., where I will argue that only the latter is. But first I would like to mention a third problem that arises once we adopt the account provided by the copy theory of movement for the unacceptability of (30)a. The problem is that under the copy theory of movement, it is far from clear how QR could provide a representation for ACD sentences that would be compatible with the theory of VP deletion.

---

[38] See Lebeaux (1989), Chomsky (1993), and Fox (in press) where some theoretical justification is provided for the hypothesis.

Lebeaux's empirical argument relies on the following contrast:

(i) a. Guess which argument that supports $John's_1$ theory he_1 adopted.
b. ??Guess which argument that $John's_1$ theory is correct he_1 adopted

For some skepticism regarding the significance of this contrast, see Kuno (199?) and Lasnik (199?). For some discussion and counter-arguments see Safir (1998). See also (43) below. Additional data that are not covered by the proposal made here is discussed in Munn (1994).

To the best of my knowledge, the problem was first noted in Lasnik (1993), and Hornstein (1994,1995) (cf. Schmitt (1995)). To appreciate it, consider again a typical ACD sentence such as (32) (= (17)a). In this sentence, the antecedent and the elided VP don't appear to be identical, yet somehow the condition on VP deletion (Parallelism) is met.

(32) I [$_{\text{VPantecedent}}$ read [a book [that you did <read t>]]].

QR solves this problem by providing an LF representation in which the two VPs are identical:

(32)' [a book [that you did <read t>]
I [$_{\text{VPantecedent}}$ read t].

However, if the copy theory of movement is adopted, it is not obvious that this solution can be maintained. The LF representation of (32) would be that in (32)'' and here again the antecedent and the elided VP are not identical.

(32)'' [a book [that you did <read t>]]
I [$_{\text{VPantecedent}}$ read [a book [that you did <read t>]]].

So we are left with a real problem. We have fairly good evidence for QR. This operation is needed if we want to keep to a simple theory of the interpretation of QNPs and at the same time to account for the locality of scope (2.1.), for the properties of ACD (2.2), and for various Condition A effects (3.1.). However, QR appears to get the wrong results for Condition C. This problem can be eliminated if we adopt the copy theory of movement, but then we loose our account for ACD. In the next section, I would like to present a solution to this conflict for which I have argued in recent work (Fox in press).[39]

4. Late Merger and Antecedent Contained Deletion

[39] For alternatives, see Lasnik (1993), Hornstein (1994), and Schmitt (1995).

We ended the previous section with a conflict between Condition C and the conditions on ellipsis (Parallelism). QR seems to have effects on Parallelism but not on Condition C. The facts pertaining to Condition C can be captured by the copy theory of movement. But then it is no longer obvious that QR should circumvent a violation of Parallelism in ACD constructions. The solution I would like to present for this puzzle is based on the idea that, contrary to initial appearances, QR on its own cannot circumvent a violation of Parallelism. An ACD construction is possible due to a combination of QR and the Late Merger of a relative clause.

In order to understand what I am trying to do, it is useful to return to (24)b.

(24)b (Guess) [which picture that $John_1$ saw] He_1 likes t.

The initial response to this sentence is of course to suggest that Condition C can be obviated by *wh*-movement. But we have seen an alternative proposal consistent with the hypothesis that *wh*-movement cannot circumvent a violation of condition C, namely that (24)b is acceptable due to a combination of *wh*-movement and Late Merger. The proposal presented below for ACD is essentially identical: QR alone cannot circumvent a violation of Parallelism, and ACD is possible due to a combination of QR and late merger. But the proposal relies on a few auxiliary assumptions that need to be introduced. One assumption (defended in Fox and Nissenbaum (1999)) is that Late Merger can follow covert operations, contrary to common assumptions about architecture (pp. 32). The second assumption concerns the interpretation of structures derived by movement, and the third concerns the representation of relative clauses.

4.1. QR and Late Merger

Why is QR covert? A commonly assumed answer is that it applies after a point of the derivation at which the phonological properties of a sentence are determined (Spell Out). If this assumption is correct, nothing (at least nothing overt) can be merged after QR. It follows that there can be no derivation parallel to (31) that involved QR instead of overt *wh*-movement.

But let's consider the possibility that this consequence is false. Of course, if this turns out to be the case, something else needs to be postulated in order to derive the fact that QR is covert. Let's put aside the question of what this "something" might be,[40] and ask what a derivation parallel to (31) would look like if it involved QR instead of overt *wh*-movement. So let's start with (31) (slightly modified in (33) so as not to distract our attention with Condition C and the way it is satisfied):

(33) a. Mary likes [which picture] --*wh*-movement→
b. [which picture] Mary likes [which picture] –adjunct merger→
c. [which picture that John saw] Mary likes [which picture]

(33)a is the structure before the application of QR. QR yields the structure in (33)b, in which the *wh*-phrase, *which picture*, is in two positions. (The higher position — the position that according to notational convention is placed to the left — is the head of the chain; the lower position is the tail.) If the derivation ends at this point, the rules of phonology target the head of the chain, and the pronunciation is that of *which picture Mary likes*. If the derivations precedes, with late merger creating the structure in (33)c, the relative clause is pronounced where it was merged and we get *which picture that John saw Mary likes*.

Now let's consider the same derivation with QR.

(34) a. Mary likes [every picture]--*QR*→
b. [every picture] Mary likes [every picture] –adjunct merger→

[40] Two types of proposals have been made in the literature. One is the proposal that principles of phonology sometimes target the tail of the chain for pronunciation (cf. Brody (1997), Bobaljik (1995, in press), Groat and O'Neil (1996) and Pesetsky (1998)). Another proposal involves multiples points of spell-out, which are followed by some covert operations. (Such a proposal was made in Chomsky (2000) and was developed with particular attention to late merger in Nissenbaum (2001). See also Chomsky (2001).) By now, various arguments have been presented against the traditional architecture, which is committed to a "covert component". (Some of the arguments focus on the commitment to "multiple cycles" advocating a single cycle grammar.) To the best of my knowledge, the initial arguments are presented in Brody (1995), who argues for a theory that dispenses with derivations.

c. [every picture that John saw] Mary likes [every picture]

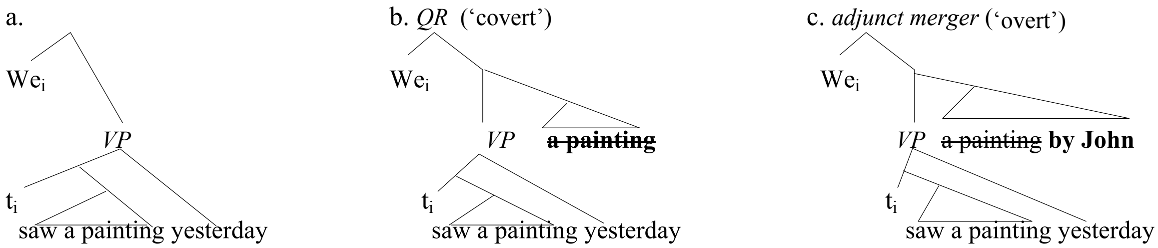
(34)a is a the structure before movement. After QR, the QNP, *every picture*, is in two positions, (34)b. But if the derivation ends at this point, the pronunciation is quite different from what we get from (33)b. QR is covert, which means that the sentence is pronounced as if it did not apply. I.e., the tail of the chain is pronounced, not the head: *Mary likes every picture*.[41] If the derivations precedes, with late merger creating the structure in (33)c, the relative clause should be pronounced where it was merged and we might expect to get the following: *that John saw(,) Mary likes every picture*. Of course, we never get this pronunciation, which might suggest that late merger cannot follow QR.

But in Fox and Nissenbaum (1999), we argue for a different conclusion. Specifically, we argue that QR applies to the right and that this is the explanation for the missing pronunciation. Furthermore, we argue that the derivation in (34) is attested yielding constructions traditionally analyzed as involving extraposition from NP (with the minor correction that QR normally targets VP, see p. 26).

To understand the nature of the claim, consider the sentence in (35). We argue that this sentence is derived by covert QR (to the right) followed by late merger; the DP *a painting* (the DP from which, under traditional accounts, the adjunct has extraposed, henceforth, the source DP) undergoes QR to VP yielding (35)a, and then the NP, *painting*, is merged with the adjunct *by John*, resulting in (35)b. (For related proposals see Gueron and May 1984, and Reinhart 1991.) The pronunciation is expected under the assumption that QR is a rightward covert operation (see note 42). The fact that the movement is covert is indicated by the ~~strikethrough~~ format.

[41] The claim that the tail is pronounced is none other than the claim that QR is covert; it, therefore, does not presuppose a particular account of covert movement. See note 40.

(35) We saw a painting yesterday by John.



We present various arguments in favor of this proposal.[42] Here, I will limit myself to evidence pertaining to scope and Condition C. As for scope, there is a very obvious, though non-trivial, prediction. Because extraposition involves covert movement of the source DP to the position where the adjunct is merged, it is predicted that the scope of the source DP would be at least as high as the extraposition site.[43] That this prediction is correct was argued in Williams (1974; chapter 4) based on slightly different constructions, and it is stated below as Williams' generalization.[44] If the generalization is correct, it provides important evidence in favor of the account.

---

[42] But various questions remain unanswered. One question is what accounts for the (apparently necessary) assumption that QR applies to the right. (See Chomsky 2001 for some criticism and for an alternative proposal.) Another question relates to cases of extraposition for which the analysis does not extend straightforwardly (such as extraposition from subjects and overtly moved *wh*-phrases).

[43] One might question this prediction under the assumption that QR can reconstruct. But If an adjunct is present only at the head of a chain, and if reconstruction results from interpreting only the tail of a chain (i.e., the head of the chain is deleted at LF), then late merger will block reconstruction; the adjunct would not be interpretable as a modifier of the source NP. Exactly these considerations are needed independently (as pointed out by Fox 2000: chapter 5) to account for correlations between scope reconstruction and Condition C. Notice that in both cases there is an argument against "semantic reconstruction". If semantic reconstruction (via type shifting operations or higher type traces) were possible (see Jacobson 1999, Sternefeld 2001 and references therein), it would not conflict with late merger of an adjunct.

[44] Williams, in contrast to Fox and Nissenbaum, focused on comparative- and result-extraposition, and not on extraposition from NP. For an account of the former in terms of late merger see Bhatt and Panceva (2001).

(36) Williams' generalization: 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).[45]

As a first step towards constructing an example that would test whether the generalization is correct, consider the ambiguity of (37).[46]

(37) John must teach no book about evolution in order to please the school committee.

One reading of this sentence asserts that if John is to please the school committee, it must be the case that he teaches no book about evolution. Under this reading the sentence might be used to convey something about the school committee’s intolerance. The other reading of the sentence (which might be facilitated by replacing ‘no book’ with ‘no *particular* book’) asserts that there is no book such that pleasing the school committee requires that John teach *that particular* book. Under this reading, the sentence does not impute intolerance to the school committee, but rather indifference. The former reading results from the DP *no book about evolution* taking narrow scope with respect to the modal operator *must,* which the rationale clause restricts. The latter interpretation results from the reverse scope relation.

If Williams’ generalization is correct then extraposition of an adjunct to a position dominating the rational clause should disambiguate sentences similar to (37) towards the interpretation in which the source DP outscopes the modal. This disambiguation is expected since the rationale clause restricts the modal (see Williams 1974). That the expectation is borne out is suggested by the contrast in (38).

[45] Of course, if Fox and Nissenbaum’s proposal is correct, it is (strictly speaking) wrong to refer to the attachment site of EC (since under the proposal, EC is adjoined to *NP* in the raised position rather than raised to VP). The intention is to refer to what is taken to be the attachment site of EC under traditional analyses. A statement of Williams’ generalization consistent with Fox and Nissenbaum proposal would be the following: When an adjunct β is “extraposed” from a “source DP” α, Every node that c-commands the base position of α and does not c-command β will be interpreted in the scope of α.

(38) a. John must miss no assignment that is required by his math teacher in order to stay in school.

b. #John must miss no assignment in order to stay in school that is required by his math teacher.

c. John must hand in no assignment in order to stay in school that is required by his math teacher.

(38)a is—in principle—ambiguous in the same way that (37) is. However, in this case the interpretation in which the DP outscopes the modal is pragmatically odd. In particular, given our assumptions about the function of schools, we do not expect students to be required to miss assignments. Therefore it is not particularly informative to assert that there are no assignments (of a particular sort) that John must miss in order to stay in school. The fact that (38)b has only this odd (non-informative) interpretation suggests that extraposition to the right of the rationale clause forces wide scope for the source DP, evidently an instance of Williams' generalization. This is reinforced by the relative naturalness of (38)c. This example, in contrast to (38)b, makes sense (only) under wide scope for the source DP, again given our assumptions about the function of schools, and thus is expected to be natural if Williams' generalization is the correct description.

Let us now move to the predictions that the analysis of extraposition has for Condition C. Because extraposition involves post QR merger of adjuncts, it is predicted to have effects on Condition C of the binding theory (identical to the effects that were observed in the case of *wh*-movement, (24)b). Evidence that this prediction is correct is provided by the following example from Taraldsen (1981) (compare with (30)a):

(39) I gave him$_i$ a book yesterday that John$_i$ liked.

The acceptability of this sentence is expected under Fox and Nissenbaum's approach to extraposition. In fact the account is identical to the account provided by

[46] The discussion of this ambiguity, inspired by Williams (1974), is taken (more or less verbatim) from Fox and Nissenbaum (in progress).

Lebeaux and Chomsky for (24)b above: Condition C is obviated because the adjunct *that John liked* is merged with the source DP at the landing site of QR, yielding a structure in which the name *John* is not c-commanded by the coindexed pronoun:

(40) I [[VP t gave him$_i$ [a book] yesterday] [~~a book~~ that John$_i$ liked]

The explanation that Taraldsen offers for the status of (39) is different; it is based on the assumption that extraposition is a movement operation and as such is capable of bleeding Condition C. Evidence against this explanation and in favor of an explanation based on late merger is provided by the contrast in (41).

(41) a. ? I told him$_i$ about your new argument the other day that supports John's$_i$ theory.

b. * I told you about his$_i$ new argument the other day that supports John's$_i$ theory.

In (41)a extraposition obviates a violation of Condition C.[47] This follows from the standard explanation as well as from the one based on late merger. However, the fact that in (41)b extraposition does not obviate a violation of Condition C is extremely puzzling under Taraldsen's explanation; the pronoun *him*$_i$ in (41)a is in a higher (surface) position than the pronoun *his*$_i$ in (41)b; therefore, if extraposition is a movement operation capable of bleeding Condition C in (41)a, it should also be able to do so in (41)b.

On the late merger account, the contrast in (41) is expected. If the effects on Condition C are due to late merger, there are very specific predictions: a violation of the condition should be obviated only if the problematic pronoun (*him*$_i$ in (41)a and *his*$_i$ in

[47] The sentence is slightly marginal, but this is unrelated to Condition C: replacing the R-expression with a pronoun does not improve the status of the sentence. In any event, the contrast between (a) and (b) is pretty sharp.

(41)b) is not contained within the source DP. This is seen by the representations of the sentences in (41) that are derived by post QR merger:[48]

(42) a. I [[$_{VP}$ t told him$_i$ about [your new argument] the other day] [~~your new argument~~ that supports John's$_i$ theory]].

b. I [[$_{VP}$ t told you about [his$_i$ new argument] the other day] [~~his$_i$ new argument~~ that supports John's$_i$ theory].

Additional evidence in favor of Fox and Nissenbaum's account comes from a consideration of Lebeaux's hypothesis that late merger is possible for adjunctts but not for complements (see p.32). This hypothesis leads to the prediction that complement extraposition will differ from adjunct extraposition in failing to bleed Condition C. The contrast in (43) suggests that the prediction is correct.

(43) a. I gave him$_i$ an argument yesterday **that supports John's$_i$ theory**.

b. ??/*I gave him$_i$ an argument yesterday **that John's$_i$ theory is correct**.

(43)a is derived by post QR merger of the relative clause (in bold face), an operation which is expected to obviate Condition C. (43)b, by contrast, cannot be derived by late merger of the bold faced phrase since this constituent is a complement. Instead, it is derived by rightward movement of a complement, which (given the copy theory of movement) does not obviate Condition C.[49]

---

[48] The relations that are relevant for Condition C might be more transparent when we consider the way the structures are interpreted, section 4.2. Thus we can characterize the interpretation of (42)a, with the structure in (i), in which the violation of Condition C is fairly transparent.

(i) [his$_1$ new argument that supports John's$_1$ theory] λx. [I told you about his$_1$ new argument, x]

[49] A derivation of adjunct extraposition via rightward movement needs to be blocked; otherwise the explanation Williams' generalization would be lost. This blocking is achieved in Fox and Nissenbaum (1999) with appeal to the observation that in general movement out of NPs is restricted to complements. See Fox and Nissenbaum (1999, in progress) for a variety of arguments that complement extraposition and adjunct extraposition have a very different analysis.

Finally, the account in terms of late merger predicts a correlation between the ability of extraposition to circumvent Condition C and the scope of the source DP. To understand this prediction, consider the contrast in (44).

(44) a.? I wanted him$_i$ not to talk to a (certain) girl yesterday that John$_i$ has known for years.[50]

b. * I wanted him$_i$ not to talk to any girl yesterday that John$_i$ has known for years.

c. I wanted John$_i$ not to talk to any girl yesterday that he$_i$ has known for years.

In (44)a, extraposition to the matrix clause is expected to yield the following structure in which Condition C is not violated:

(44)'a I [[$_{VP}$ t wanted him not to talk to [a (certain)] girl yesterday] [~~a (certain) girl~~ that John has known for years]

However, an obviation of Condition C by extraposition has predicted consequences for the scope of the source DP; it must be at least as high as the extraposition site. In the case of (44)a, the indefinite *a (certain) girl that John has known for years* must outscope both negation and the intensional verb *want*. This predicted consequence seems to be correct (as far as speakers can access the judgments). More importantly, the unacceptability of (44)b verifies the prediction: the source DP is an NPI, which is not allowed to outscope negation, and thus brings about a conflict with the requirement imposed by Condition C. (44)c is acceptable since Condition C imposes no requirements and extraposition to the embedded VP is available.

What have we learned from this discussion? We saw that if we adopt a model of grammar in which overt and covert operations can be interspersed, we might expect covert movement to be followed by late merger of an adjunct. Furthermore, if we assume that covert movement applies to the right, extraposition from NP cries out for an analysis

[50] (44)a is marginal for some speakers. However, it is definitely better than (44)b. See Fox (in press).

in these terms. Finally, we saw that there are various non-trivial predictions for scope and for condition C, as well as for their interaction, and that there is some evidence that these predictions are borne out.

Can this help us in resolving the conflict between ACD and the copy theory of movement? Given the copy theory, QR cannot circumvent the violation of Parallelism that ought to arise in ACD constructions. This, as we noted, is similar to the fact that movement cannot circumvent a violation of Condition C. But could the acceptability of ACD constructions rely on QR + late merger (extraposition) along the logic of our account for (24)b?

So consider a simple ACD construction such as (32), repeated below.

(45) I [$_{VPantecedent}$ read [a book [that you did <read t>]]].

Suppose that this sentence has the following derivation:

(45)' a. I [$_{\mathbf{VPantecedent}}$ t read a book] –QR→
b. I [[$_{\mathbf{VPantecedent}}$ t read a book] **~~a book~~**] –late merger→
c. I [[$_{\mathbf{VPantecedent}}$ t read a book] ~~a book~~ **[that you did <read t>]**].

The elided VP and the antecedent look a little closer to each other in (45)'c then they do in an LF that involves QR without late merger (see (32)''). But it is still not obvious that Parallelism is satisfied. Parallelism requires that the trace/copy formed by QR be identical to the trace/copy inside the relative clause. But we haven't said anything yet about the form that the latter has under the copy theory of movement. And, also, we haven't talked about the way that structures derived be movement might be interpreted once the copy theory is adopted. Both issues will be relevant for the question of Parallelism, as we will see. I will begin with the latter.

4.2. Movement and interpretation

Chomsky (1995) points out that the copy theory of movement simplifies syntax in two ways. First, the theory eliminates the need to postulate new objects (traces) beyond run-of-the-mill lexical items. In this sense, it brings us closer to a view of syntax as a recursive procedure that does not access anything but the lexicon. Second, the copy theory turns movement into a simpler operation; it is practically identical to the elementary structure-building operation (*merge*), differing only in that it takes as input an object that has served as input for an earlier merger (*move* is *re-merge*).

How do the copy theory of movement and the traditional alternative (with traces) compare with respect to semantic interpretation? We spelled-out a semantic mechanism that interprets structures with traces in (11), repeated below with a minor though insignificant modification.

(46) In a structure formed by DP movement, $DP_n[_{\varphi}\ldots t_n\ldots]$, the derived sister of DP, φ, is interpreted as a function that maps an individual, x, to the meaning of φ[x/n].

φ[x/n] is the result of substituting every constituent with the index n in φ with him_x, a pronoun that denotes the individual x.

It is fairly obvious that this rule does not depend on a theory that postulates traces instead of copies. The rule refers to the interpretation of structures derived by substituting indexed elements with pronouns of a certain sort. It, therefore, doesn't matter what properties the indexed elements have. In particular, it doesn't matter whether these indexed elements are traces or copies. So the rule might as well be stated as in (47), and it, therefore, cannot serve to distinguish between trace- and copy-theory.

(47) In a structure formed by DP movement, $DP_n[_{\varphi}\ldots DP_n\ldots]$, the derived sister of DP, φ, is interpreted as a function that maps an individual, x, to the meaning of φ[x/n].

φ[x/n] is the result of substituting every constituent with the index n in φ with him_x, a pronoun that denotes the individual x.

We conclude that syntax is simplified by copy-theory and that semantics is not affected. So we have a decent conceptual argument in favor of the theory. But notice that the rule in (47) ignores every property of the copy besides it index. In this sense, it might be viewed as mysterious that the content of the copy has relevance for Condition C of the binding theory. I.e., we don't understand why Condition C is different from the rule in (47) in that only the former is sensitive to the fact that a name is dominated by a copy. Both the rule in (47) and in (46) treat traces the way that we treat variables in logic, and it is therefore somewhat surprising that Condition C "sees" something which is richer than a variable.

As a response to this, I have postulated a syntactic rule that converts lower-copies/traces into structures that contain variables. The rule was designed to convert a trace to a definite description, eventually yielding interpretations similar to those of the following paraphrases:[51]

(48) Which boy Mary visited which boy?
Paraphrase: Which is the boy, *x*, such that Mary visited **the boy x**?

(49) every boy A girl talked to every boy.
Paraphrase: For every boy, *x*, there is a girl who talked to **the boy x**.

The boy x in the paraphrase is modeled on definite descriptions in natural language such as *the man John*. These definite descriptions, I assume, are interpreted by predicate modification of ⟦man⟧ and $\lambda x(x=[\![John]\!])$ with the resulting predicate serving as argument of the determiner. The postulated rule, thus, involved two operations:

(50) **Trace Conversion:**

a. Variable Insertion: (Det) Pred → (Det) [Pred $\lambda y(y=him_n)$] [52]

[51] The suggestion was based on a proposal made in a different context by Rullmann and Beck (1998).

[52] Where n is the index of the moved QNP.

b. Determiner Replacement:(Det) [Pred $\lambda y(y= him_n)$] → the [Pred $\lambda y(y= him_n)$][53]

Furthermore, I assumed that the output of this syntactic rule is interpreted by the semantic rule in (46). To see how this would work, consider the derived sister of *every boy* in (49), *A girl talked to every boy*. This constituent is converted by Trace Conversion to *A girl talked to the boy (identical to) him$_n$*, which is interpreted by the relevant semantic rule as *λx. A girl talked to the boy, x*.[54]

Once Trace Conversion is part of grammar (and the rule in (46) is maintained) it is pretty clear why QR cannot affect condition C. A sentence such as (30)a, repeated below as (51)a, is unacceptable since the LF structure is one in which the name *John* is *semantically interpreted* in a position that is not licensed by Condition C. To make this clear (i.e, to make the interpretation clear together with the ramifications for Condition C), I will use the lambda notation to indicate a position in the structure where the rule in (46) applies. So the structure of (51)a derived by a combination of QR and trace conversion is (51)b, and it is clear that Condition C is ruled out in this structure.[55]

(51) a. A boy introduced him_1 to every friend of $John's_1$.

b. [every friend of $John's_1$]

λx. A boy introduced him_1 to [the friend of $John's_1$ identical to x].

Trace Conversion, I will argue, can also help us in accounting for Antecedent Contained Deletion. But before we get to this, I would like to make a short conceptual

---

53 For the sake of accessibility the rule is stated in somewhat informally using objects of the semantic theory (the "meta language") as if they were syntactic objects (in the "object language"). This informal practice will continue throughout this paper. Furthermore, sometimes the output of trace conversion will be stated as *the Pred x* (as in (48)), and sometimes to aid the reader I will add the words "identical to" between *Pred* and *x* (as in (51)b).

54 The lamda-expression is a partial function defined only for boys. (It maps a boy to TRUE iff Mary likes that boy). There are no semantic problems with this rule since natural language determiners denote conservative two-place predicates ($D(A,B)=D(A, A\cap B)$). My hope is that this rule could play an interesting role in accounting for the conservativity.

55 So, if Reinhart's (1983) perspective on Condition C is right, the sentence could be ruled out on the grounds that the derived predicate (the derived sister of the raised QNP) would be the same if *John* is replaced by a variable bound by the pronoun *him*. The same story could not be told if the name *John* does not contribute to semantic interpretation, as in the rule in (47).

diversion. If Trace Conversion is viewed as part of syntax, there still needs to be a semantic rule that tells us how the structure is interpreted, e.g., the rule in (46). But if this is how things work, it seems to me that Chomsky's conceptual argument in favor of the copy theory of movement is undermined. It is true that a copying operation (*re-merge*) is simpler than an operation that moves an object and replaces it with a trace/variable. But if the copy has to be converted to an object that contains a variable, questions of simplicity are much harder to address. For this reason, I would like to mention the possibility that Trace Conversion is an artifact of the semantic rule that interprets the derived sister of a moved constituent, and that this rule should therefore be modified (see Elbourne, in press):

(52) In a structure formed by DP movement, $DP_n[_{\varphi}\ldots DP_n\ldots]$, the derived sister of DP, φ, is interpreted as a function that maps an individual, x, to the meaning of $\varphi[x/n]$.

$\varphi[x/n]$ is the result of replacing the head of every constituent with the index n in φ with the head the_x, whose interpretation, $[\![the_x]\!]$, is, $\lambda P.\ [\![the]\!](P \cap \lambda y.y = x)$.

This semantic rule interprets the derived sister of a moved constituent directly (without Trace Conversion). Nevertheless, the interpretation is identical to the one resulting from the standard rule, (46), when it applies to the output of Trace Coversion. This raises the possibility that ultimately Trace Conversion can be eliminated with no significant modification to the ideas outlined below. But at the moment, I will continue to employ Trace Conversion as a syntactic rule.

4.3. Relative Clauses

I will assume (following Kayne 1974, Cinque 1982 and Sauerland 1998) that relative clauses are both head external and head internal (what Sauerland calls the "matching analysis", see also Cresti 1999 and Kennedy 2000). More specifically, I will assume that the derivation of relative clauses involves "movement to Comp" of a CP internal NP, which is deleted under identity with a CP external NP:

(53) Every boy [CP ~~boy~~ Mary likes ~~boy~~].

Furthermore, I will assume (along with many) that the NP in [spec, CP] is not interpreted but that movement turns the relative clause into a predicate that combines with the CP external NP by predicate modification. More specifically, given Trace Conversion, we get the following structure:

(54) Every [boy λx. Mary likes the boy x][56]

4.4. Explaining Antecedent Contained Deletion

If (54) is the interpreted structure of a DP containing a relative clause, and if the relative clause can be merged with NP after the QNP has undergone QR, then a sentence involving ACD can have the following derivation.

(55) [$_{VP}$ John likes every boy] –*QR*--->
[[$_{VP}$ John likes] every boy] **every boy** –*adjunct merger*-->
[[$_{VP}$ John likes every boy] every boy **that Mary does <likes boy>**]

Trace Conversion transforms the structure to (56), which satisfies Parallelism.

(56) [every boy λx Mary does < likes the boy x>]
λy John likes the boy y]

So, here is what we've learned. Given the copy theory, ACD constructions do not escape a violation of Parallelism by virtue of QR alone. However, we have seen (in 4.1.) evidence in favor of the proposal that QR can be followed by later merger of an adjunct

[56] The meaning of this QNP is the following: $\lambda P \forall x((boy(x) \& \text{Mary likes the boy } x) \rightarrow (P(x)=1))$.

(extraposition).[57] Next, we asked whether Parallelism would be satisfied by a combination of QR and late merger. We couldn't address this question before we spelled out some specific assumptions about the structure of relative clauses and the methods that might allow chains to be interpreted. We saw that Trace Conversion together with a (version of a) head internal analysis for relative clauses (the matching analysis) provided a positive answer. But it is clear that quite a few assumptions are involved. We would therefore like some additional evidence.

In Fox (in press), I've attempted to provide various sources of evidence. Some of the evidence consisted in arguments (partially taken from previous literature) that extraposition is indeed needed for ACD resolution. There was also important evidence from work by Chris Kennedy and (in particular) Uli Sauerland that the effects of the copy theory of movement are visible in ACD constructions. In particular, there were specific environments in which the copy theory of movement predicted that late merger would be insufficient to obviate a violation of Parallelism, and there was evidence that the predictions come out the right way. Unfortunately, I will have to stop here.

5. Conclusions

We ended section 3 of this paper in a somewhat peculiar situation. We had a good deal of evidence in favor of LF structures in which QNPs are in their scope position (I.e., they are sisters of their argument/scope). The LF structures were needed for a simple semantic procedure to be possible (section 1). Furthermore, there was evidence from locality considerations (section 2.1), from Parallelism (section 2.2), and from Condition A of the Binding Theory (section 3.1). The nagging problem was Condition C. This condition contrasted with all the other evidence in showing no indication that the postulated LF structures were real. The Copy Theory of Movement provided a way out, which was consistent with our discussion of Parallelism, as long as we adopted a model of grammar

[57] The idea that extraposition is necessary for ACD has been proposed in Baltin (1987) and has resurfaced in various forms (See Lasnik 1995, Wilder 1995, Abe and Hoshi 1998). For some comments on the differences and similarities between the proposals, see Fox (in press).

in which an overt operation of late merger could follow the covert operation of Quantifier Raising (section 4).

References

Abe, Jun and Hiroto Hoshi. 1999. A generalized rightward movement analysis of Antecedent Contained Deletion, *Journal of Linguistics* 35, 451-487.

Aoun, Joseph and Audrey Li. 1993. *The Syntax of Scope*, MIT Press.

Abush, Dorit. 1994. The Scope of Indefinites. *Natural Language Semantics* 2:83-135.

Aguero-Bautista, Calixto. 2001. *Cyclicity and the Scope of WH-Phrases*, Doctoral dissertation, MIT.

Barker 2001. Continuations and the Nature of Quantification. Unpublished ms., UC San Diego.

Baltin, Mark. 1987. Do Antecedent Contained Deletions Exist? *Linguistic Inquiry* 18:579-595.

Beghelli, Filippo and Tim Stowell 1997. Distributivity and Negation. In *Ways of Scope Taking*, Anna Szabolcsi, ed, 71-107. Dordrecht, the Netherlands: Kluwer.

Belletti, Adriana and Luigi Rizzi. 1988. Psych-Verbs and Theta Theory. *Natural Language and Linguistic Theory* 6:291-353.

Bhatt, Rajesh and Roumyana Pancheva. 2001. Variable Height of Adjunction and the LF Licensing of Degree Complements. Unpublished manuscript. University of Texas at Austin.

Bobaljik, Jonathan 1995 *Morphosyntax: the Syntax of Verbal Inflection* Doctoral dissertation, MIT, MITWPL.

Bobaljik, Jonathan. in press. A-chains at the Interface: Copies, Agreement and 'Covert Movement', to appear in *Natural Language and Linguistic Theory*.

Brody, Michael 1995 *Lexico-Logical Form,* MIT Press.

Brody, Michael 1997 Perfect Chains, in Haegeman, Liliane ed. *Elements of Grammar: Handbook in Generative Syntax Kluwer International Handbooks of Linguistics,Vol 1*, Kluwer.

Bruening, Benjamin. 2001. QR Obeys Superiority: ACD and Frozen Scope. *Linguistic Inquiry* 32:233-273.

Cooper 1979. Variable Binding and Relative Clauses. In *Formal Semantics and Pragmatics for Natural Languages*, F. Guenthner and S. J. Schmidt, eds, Dordrecht, the Netherlands: Reidel.

Chomsky, Noam. 1976. Conditions on Rules of Grammar. Linguistic Analysis 2:303-351.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht, the Netherlands: Foris.

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory, in *The View From Building 20*, Cambridge, MA: MIT Press re-printed as chapter 3 of Chomsky 1995

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

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Roger Martin, David Michaels, and Juan Uriagereka, eds, 89-155, Cambridge, Mass: MIT Press

Chomsky, Noam. 2001. Beyond Explanatory Adequacy. Manuscript. MIT.

Cinque, Guglielmo. 1981-1982. On the Theory of Relative Clauses and Markedness, *The Linguistic Review* 1, 247-294.

Cresti, Diana. 1999. Ellipsis and Reconstruction in Relative Clauses, in *Proceedings of NELS 30*, GLSA Publications, Amherst.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Dummett, Michael. 1977. *Frege: Philosophy of Language*. Harvard University Press.

Fox, Danny. 2000 *Economy and Semantic Interpretation*, MIT Press.

Fox, Danny. 2002. Antecedent Contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry*. 33(1): 63-96.

Fox, Danny and Jon Nissenbaum. 1999. Extraposition and Scope: A case for overt QR, *Proceedings of the West Coast Conference on Formal Linguistics 18*.

Friedin, Robert 1986 Fundamental Issues in the Theory of Binding, in Barbara Lust, ed. *Studies in the Acquistion of Anaphora*, Dordrecht: Reidel.

Géuron Jacqueline and Robert May 1984. Extraposition and Logical Form, *Linguistics Inquiry*, 73, 429-462.

Groat, Erich, and John O'Neil 1996 "Spellout at the LF interface," In Werner Abraham, Samuel David Epstein, Hɣskuldur Thrℸinsson & C. Jan-Wouter Zwart eds.. Minimal Ideas: Syntactic Studies in the Minimalist Framework. John Benjamins Publishing Company, Amsterdam/Philadelphia, 113-139.

Heim, Irene. 1997. Predicates or Formulas? Evidence from Ellipsis, In *Proceedings of SALT 7*, Aaron Lawon and

Eund Cho, eds, 197-221, Ithaca, New York: CLC Publications.
Heim, Irene and Angelika Kratzer 1998, *Semantics in Generative Grammar*, Blackwell, Oxford.
Hendriks, Herman. 1993. Studied Flexibility. Amsterdam: ILLC Dissertation Series.
Hirscbuhler, Paul. 1982. VP Deletion and Across the Board Quantifier Scope. In Proceedings of NELS 12.
Hornstein, Norbert. 1994. An Argument for Minimalism. The Case of Antecedent Contained Deletion. *Linguistic Inquiry* 25:662-688.
Hornstien, Norbert. 1995. *Logical Form: From GB to Minimalism.* Oxford: Blackwell.
Hoji, Hajime. 1985. Logical form constraints and configuration structures in Japanese. Doctoral dissertation, University of Washington.
Huang, James 1982 Logical Relations in Chinese and the Theory of Grammar, Doctoral Dissertation, MIT.
Jacobson, Pauline 1992 Comment Flexible Categorial Grammar, Questions and Prospects, in Robert Levine ed., Formal Grammar, Theory and Interpretation, Oxford University Press.
Jacobson, Pauline. 1999. Toward a Variable Free Semantics. Linguistics and Philososphy 22:117-184.
Jacobson, Pauline. In press. The (dis) organization of Grammar: 25 years, to appear in *Linguistics and Philosophy*.
Johnson, Kyle and Satoshi Tomioka. 1997. Lowering and Mid-sized Clauses. In Proceedings of the Tübingen Workshop on Reconstruction.
Johnson, Kyle. 2000. How Far Will Quantifiers Go? In *Step by Step*, Roger Martin, David Michaels, and Juan Uriageka, eds, 187-210. Cambridge, Mass.: MIT Press.
Kayne, Richard S. 1974. French Relative que, In Maria Lujan and Frita Hensey eds. *Current Studies in Romance* Linguistics. Washington D.C.: Georgetown University Press.
Kayne, Richard S. 1998. Overt vs. Covert Movements, *Syntax* 1:128-191.
Kayne, Richard S. 1981. Two Notes on the NIC, In A. Belletti, L. Brandi , and L. Rizzi eds., *Theory of Markedness in Generative Grammar, Proceedings of the 1979 Glow Conference*, 475-516. Scuola Normale Superiore, Pisa.
Kennedy, Christopher, 1994. Argument contained Ellipsis, Linguistics Research Center Report LKC-94-03, University of California, Santa Cruz.

Kennedy, Christopher, 1997. Antecedent Contained Deletion and the Syntax of Quantifiers, *linguistic Inquiry*.
Kennedy, Christopher, 2000. "Comparative Deletion and Optimality in Syntax," Unpublished ms. Northwestern University.
Kiss, Katalin. 1991. Logical Structure in Syntactic Structure. In *Logical Structure in Linguistic Structure: Cross-Linguistic Perspectives*, James Huang and Robert May, eds, 111-147, Dordrecht, the Netherlands: Kluwer.
Klima, Edward. 1964. Negation in English. In *The Structure of Language*, Jerry Fodor and Jerrold Katz, ed, 246-323. New York: Prentice-Hall.
Kuno, Susumu. 1997. Binding Theory and the Minimalist Program, Ms. Harvard University.
Kuroda, S.-Y. 1971. Remarks on the notion of subject with reference to words like also, even, or only, Part II. *Annual Bulletin* 4, 127-152.
Lasnik, Howard 1993. Lectures on Minimalist Syntax. Uconn Working Papers in Linguistics: Occasional Papers Issue 1. [reprinted in Howard Lasnik, *Minimalist Analysis*, Blackwell, 1999.]
Lasnik, Howard 1995. A Note on Psuedogapping, In *Papers on Minimalist syntax,* Rob Pensalfini and Hiroyuki Ura, eds., 143-164. MIT Working Papers in Linguistics 27. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, MA.
Lasnik, Howard 1998. Reconstruction Riddles, In *UPenn Working Papers in Linguisitics*, vol 5.1..
Larson, Richard and Robert May 1990. Antecedent Containment or Vacuous Movement: Reply to Baltin, *Linguistic Inquiry*, 21, 103-122.
Lebeaux, David 1988. *Language Acquisition and the Form of the Grammar*, Doctoral dissertation, University of Massachusetts, Amherst.
Lebeaux, David. 1991. Relative Clauses, licensing, and the nature of the Derivation. Syntax and Semantics 25: Perspectives on Phrase Structure: Heads and Licensing, Susan Rothstein, ed, 209-239. Orlando, Florida: Academic Press.
Lebeaux, David. 1995. Where Does the Binding Theory Apply? In *University of Maryland Working Papers in Linguistics*, 3:63-88. University of Maryland, College Park.
May, Robert. 1977. The Grammar of Quantificiation. Doctoral dissertation. MIT.
May, Robert 1985. *Logical Form*. Cambridge, Mass.: MIT

Press.
May, Robert. 1988. Ambiguities of Quantification, *Linguistic Inquiry* 19:118-135.
McCawley, James: 1970, “Where do Noun Phrases Come from?”, in R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar.* Waltham, MA: Ginn & Co.
McCawley, James. 1988. *The Syntactic Phenomena of English*. Chicago University Press.
Munn, Alan. 1994. A Minimalist Account of Reconstruction Asymetries. In *Proceedings of NELS 24*:397-410. ed. Merce Gonzalez GLSA, University of Massachusetts, Amherst.
Nissenbaum, Jonathan. 2001. *Investigations of Covert Phrase Movement*. Doctoral dissertation. MIT, Cambridge, Mass.
Pesetsky, David. 1987. WH-in Situ: Movement and Unselective Binding. In The Representation of (In)definiteness, Eric Reuland and Alice ter Meulen, eds, 98-167, Cambridge, Mass.: MIT Press.
Pesetsky, David. 1998. Some Optimality Principles of Sentence Pronunciation, in *Is the Best Good Enough? Optimality and Competition in Syntax?,* P. Barbosa, et.al, eds., Cambridge, Mass.: MIT Press.
Pesetsky, David. 2000. *Phrasal Movement and its Kin*, Cambridge, Mass.: MIT Press.
Postal, Paul: 1974, “On Certain Ambiguities”, *Linguistic Inquiry V,* 367-424.

Reinhart, Tanya. 1983.*Anaphora and Semantic Interpretation.* London: Croom Helm.
Reinhart, Tanya. 1991. Non-Quantificational LF, In A. Kasher ed. *The Chomskian turn*,
Reinhart, Tanya. 1997. Quantifier-Scope: How labor is divided between QR and choice functions, *Linguistics and Philosophy*, 20:335-397, 1997.
Riemsdijk, Henk van, and Edwin Williams 1981. NP-Structure, *The Linguistic Review* 1:171-217.
Richards, Norvin. 2001. *Movement in Language: Interactions and Architecture.* Oxford University Press.
Rodman, 1976. Scope Phenomena, Movement Transformations, and Relative Clauses. In *Montague Grammar*,Barbara Partee, ed, 156-176. New York: Academic Press.
Romero, 1998, "The Correlation between Scope Reconstruction and Connectivity Effects", in E. Curtis, J. Lyle and G. Webster (eds.), *Proceedings of the XVI West Coast Conference in Formal Linguistics* (Seattle, 28 February - 2 March 1997). Standford: CSLI Publications. Pp: 351-366.

Rooth, Mats 1992 Ellipsis reduandancy and Reduction Redundancy, In *Proceedings of Stutgart Ellipsis Workshop.* Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29, S. Berman and A. Hestvik, eds., IBM Germany, Heidelberg.

Ruys, Edy. 1992. *The Scope of Indefinites*. Doctoral dissertation. Utrecht University.

Rullmann, Hotze and Sigrid Beck 1997. Reconstruction and the Interpretation of *Which*-Phrases, in Grahahm Katz, Shin-Sook Kim and Heike Winhart (eds.) *Reconstruction: Proceedings of the 1997 Tübingen Workshop,* Universtat Stuttgart, 223-256.

Safir, Ken. 1999. Vehicle Change and Reconstruction of A-bar-Chains, *Linguistic Inquiry*, 304: 587-620.

Sag, Ivan. 1976. *Deletion and Logical Form*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.

Sauerland, Uli. 1998. *The Making and Meaning of Chains*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.

Sauerland, Uli and Armin von Steckow. 2000. The syntax-semantics Interface. In *International Encyclopedia of the Social and Behaviorial Sciences* N. Smelser and P. Bates, ed, Amsterdam: Elsevier.

Schmitt, Christina, 1995, "Antecedent Contained Deletion meets the Copy Theory." In Beckmann, G. (ed.) *Proceedings of NELS 25, Amherst, GLSA.*

Sportiche, Dominique 2001. Reconstruction, Binding and Scope. Unpublished manuscript. UCLA.

Sternefeld, Wolfgang. 2001. Semantic vs. Syntactic Reconstruction. *Linguistic Form and its Computation*, H. Kamp, A. Rossdeutscher, C. Rhorer eds, 145-182, Stanford: CSLI Publications.

Szabolcsi, Anna. 1998. Ways of Scope Taking. Dordrecht, the Netherlands: Kluwer.

Stroik, Thomas 1992. Adverbs and Antecedent Contained Deletion, *Linguistics* 30:375-.

Taraldsen, Tarald. 1981. "The Theoretical Interpretation of a Class of Marked Extractions, In *Theory of Markedness in Generative Grammar, Proceedings of the 1979 Glow Conference,* A. Belletic, L. Brandi , and L. Rizzi eds., , 475-516. Scuola Normale Superiore, Pisa.

Tiedeman, Robyne. 1995. Some Remarks on Antecedent Contained Deletion, in Shosuke Haraguchi and Michio Funaki, *Minimalism and Linguistic Theory*, Hituzi Linguistics Workshope Series No. 4, Japan: Hituz Syobo.

Vergnaud, Jean-Roger. 1974. French Relative Clauses. Doctoral dissertation. MIT, Cambridge, Mass.

Wilder, Chris 1995. "Antecedent Containment and Ellipsis," FAS Papers in Linguistics 4, 132-165.
Wilder, Chris 1997. Phrasal Movement in LF: de re readings, VP Ellipsis and Binding, *Proceeding of NELS 27*, 425-440.
Williams, Edwin 1974. *Rule Ordering in Syntax*, , Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
Williams, Edwin 1977 "Discourse and logical form," *Linguistic Inquiry* 8.101-139.
William Edwin. 1986 "A Reassignment of the Functions of LF," *Linguistic Inquiry*, 17, 265-299.
Wyngaerd, Vanden G. and C.J.W. Zwart. 1991. Reconstruction and Vehicle Change. In Linguistics in the Netherlands 1991, F. Drijkoningen and A. van Kemenade, eds. John Benjamins, Amsterdam, 151-160.