

## CHAPTER 1

### Introduction

This dissertation is primarily concerned with the syntax of sentential coordination, that is, coordination of TP constituents, although the case of coordination of CPs will become relevant in a few sections as well. A simple case of TP-Coordination is given in (1.1) below:

(1.1)

$_{TP}$ [John bought this book] and  $_{TP}$ [Fred read that magazine]

The immediate question that arises is how the two TP conjuncts [John bought the book] and [Fred read that magazine] are concatenated in the syntax. The traditional answer is that these two TPs are coordinated at “the same level” of syntactic structure, creating a ternary branching phrase structure tree. This analysis has been abandoned in the Principle and Parameters tradition, and there have been several attempts to accommodate coordination to the X'-schema of phrase structure, which disallows ternary branching. Here I will assume one of those X'-Theory compatible

analyses (Munn's 1993 Adjunction Analysis), which I will modify in relevant respects in chapter 2.

One of the most interesting syntactic constraints that applies to coordinate structures is that they are islands for movement; no element can be extracted from a position inside a coordinate structure to a position outside it. Examples of this violation are given in (1.2):

(1.2)

- a. \*Which magazine did John buy this book and Fred read \_?
- b. \*Which book did John buy \_ and Fred read that magazine?

This illustrates the Coordinate Structure Constraint, discovered by Ross (1967). In this dissertation I provide further evidence that this constraint is not a syntactic constraint, but should be understood as a semantic restriction (see Munn 1993 and Hornstein and Nunes 2002). Moreover, I am going to argue that extraction from the first conjunct and extraction from a non-first conjunct should not receive a unified treatment but are banned by distinct principles. Extraction from a non-first conjunct is ruled out in the syntax by the Condition on Extraction Domains, while extraction from the first conjunct is ruled out by a parallelism requirement at the PF and LF interfaces.

A famous apparent exception to the Coordinate Structure Constraint is the phenomenon of Across-the-Board extraction, in which one single overt element

seems to be interpreted in two different positions, one inside each of the conjuncts.

An example of this phenomenon is given in (1.3):

(1.3)

Which book has John bought \_ and Mary read \_?

Another phenomenon that has this property is that of Parasitic Gap constructions, like the one in (1.4):

(1.4)

Which book did John file \_ without reading \_?

Given this similarity, there have been a number of attempts in the literature to assimilate one to the other and provide a unified syntactic analysis for them (see for example Munn 1993, Nunes 2004, and others). As will be seen, the motivation for a unified analysis for these constructions follows trivially from the analysis of coordination that I propose here in this dissertation.

One topic that I will not discuss in this dissertation is the semantics of coordinate structures. As I explain in chapter 2, my analysis of the syntactic structure of coordination is based on Munn's 1993 approach, and I assume his approach to the semantics of coordination.

This dissertation is not intended to exhaustively investigate *all* syntactic properties of coordination or of all the other related phenomena that I discuss. Rather, it is an attempt to understand how particular phenomena relate to one another. Thus, one important goal of this dissertation is to scrutinize traditional notions like “coordination” and “subordination”, and also more recent (yet still construction-specific) notions like the Coordinate Structure Constraint, Across-the-Board extraction, and Parasitic Gap constructions, and argue for an analysis that will provide us with new insights into these phenomena based on two independently motivated theoretical constructs: adjunction (see e.g. May 1985) and Sideward Movement (Nunes 1995, 2001, 2004). Following the strong Minimalist thesis (Chomsky 1995), an overall goal of this dissertation is to reduce the properties of these constructions without resorting to construction-specific mechanisms in the narrow syntax. Rather, I attempt to deduce construction specificity from specific properties of lexical items interacting with the PF and LF interfaces, noting possible difficulties for such an approach.

## CHAPTER 2

### **The syntactic analysis of Coordination**

In this chapter I briefly summarize certain aspects of the three main analyses that have been proposed for coordinate structures. I will argue that Munn's (1993) adjunction analysis is the most adequate one, although I will modify it slightly in section 2.3. In section 2.4 I discuss the consequences of this modified approach for the traditional coordination/subordination distinction and in section 2.5 I lay out the consequences of this approach for our understanding of the Coordinate Structure Constraint and for the derivation of Across-the-Board and Parasitic Gap constructions. Exploring these consequences in detail is the central purpose of chapters 3 and 4 of this dissertation.

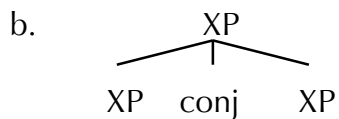
#### **2.1. Ternary branching, X', or adjunction?**

Three main analyses have been proposed for coordinate structures that I will briefly discuss here: the (traditional) ternary branching analysis (see e.g. Chomsky 1965), the X' analysis (Kayne 1994 and Johannessen 1998, among others), and the Boolean Phrase/adjunction analysis (Munn 1993).

In the Standard Theory framework of Chomsky 1965, coordinate structures were thought to be the result of applying the Phrase Structure Rule in (2.1a), which yields the Phrase Structure Tree in (2.1b):

(2.1)

a.  $XP \Rightarrow XP \text{ conj } XP$



This Phrase Structure Rule predicts that a coordinate structure generally has the same categorial status as the syntactic objects that are coordinated. Thus, any syntactic position where, say, a DP is generable also allows for an infinite number of coordinated DPs, as shown in (2.2):

(2.2)

a. John saw <sub>DP</sub>[**Mary**]

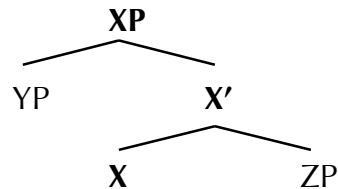
b. John saw <sub>DP</sub>[**Mary and Peter**]

c. John saw <sub>DP</sub>[**Mary and Peter and Harry**]

With the introduction of the X'-Theory of phrase structure (Jackendoff 1977), this analysis was abandoned. (Actually X'-Theory meant the abandonment of the whole Phrase Structure Rule formalism.) X'-Theory imposes binary branching and requires

that the category of a phrase be the result of projecting a head of the same category, as sketched for the category X in (2.3):

(2.3)



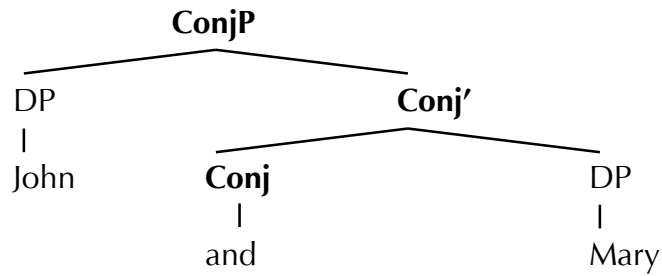
As can be seen, this general template for phrase structure is incompatible with the structure in (2.1b) above, since (2.1b) involves ternary branching, and the category of the whole phrase is determined by the category of the phrasal elements, not by its head.

Another problem for this approach is that there are acceptable examples of coordination where the two conjuncts do not belong to the same category (see (2.6) below for example). All these cases would be ungenerable under the ternary branching approach.

Two main analysis have been proposed for coordinate structures that are compatible with X'-Theory: Kayne's (1994) X' analysis and Munn's (1993) adjunction analysis.

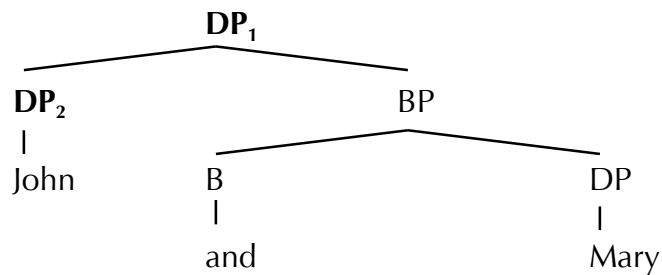
Kayne takes the coordinate conjunction to be the head of the whole coordinate structure. Under this analysis, a phrase like *John and Mary* would have the structure in (2.4) below:

(2.4)



Munn argues that it is the first conjunct that determines the category of the whole phrase, and also argues that the conjunction is a Boolean operator (B) that heads a Boolean Phrase (BP). This Boolean Phrase is adjoined to the first conjunct, as shown in (2.5):

(2.5)



One piece of evidence in favor of this analysis over Kayne's comes from cases of coordination in which the two conjuncts are not categorically identical. As the paradigm in (2.6) below shows, when this is the case, the category of first conjunct is the one that sets the selectional restrictions for the whole coordinate structure:



(2.6)

- a. You can depend on [my assistant]
- b. \*You can depend on [that John will be here on time]
- c. You can depend on [my assistant [and that John will be here on time]]
- d. \*You can depend on [that John will be here on time [and my assistant]]

Even though the X' analysis in principle does not preclude two conjuncts of different syntactic category, it is not clear how it can account for this asymmetry in the conjuncts. The explanation that has been proposed for cases like these (see e.g. Johanessen 1998) is that there is a Spec-Head relation that transfers the selectional requirement of the Specifier (a DP in the cases in (2.6c)) to the conjunction. While it is not unreasonable to appeal to a Spec-Head relation to explain these facts (although this kind of syntactic relation has been undermined lately (see Chomsky 2001)), it is not clear at all why a Head-Complement relation could not do the same (notice the ungrammaticality of (2.6d)). For more arguments against the idea that the whole coordinate structure is a ConjP, see e.g. Borsley (2005).

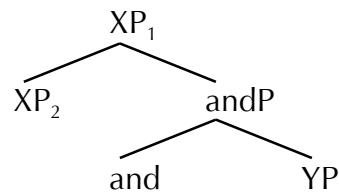
These facts, however, follow naturally under the adjunction analysis (see Munn 1993 for more arguments in favor of this approach). I will be assuming the main idea of Munn's analysis of coordination in this dissertation and will also explore in detail its implications for the formulation of the Coordinate Structure Constraint. In section 2.3 I argue, however, that Munn's analysis should be modified, at least in the case of sentential (TP-)coordination, given evidence that I

provide from Binding effects which indicate that the subject of the first conjunct c-commands into the second conjunct.

## 2.2. Some notes on the adjunction formalism

As pointed out in the previous section, there are good reasons to adopt Munn's (1993) analysis of coordination, under which the second conjunct (an andP) is adjoined to the first one, as shown again in (2.7) below. For simplicity, I am now labeling the Boolean Phrase simply as "andP" (and will continue to do so in the remainder of this dissertation):

(2.7)



Before I discuss TP-coordination under this analysis, I would like to discuss issues that the analysis raises regarding linear order.

One issue that Munn 1993 does not address is how an adjoined element (andP in (2.7)) can follow the terms of the structure it adjoins to. The structure in (2.7) (coupled with a resulting order <XP, and, YP>) is incompatible with Kayne's (1994) Linear Correspondence Axiom, as I show below. (It should be noted, however, that Munn's proposal predates Kayne's theory). As is well known, Kayne

proposes that linear order is predictable from hierarchical structure, and argues for the following principle:

(2.8) Kayne's (1994) Linear Correspondence Axiom

- a. If  $\alpha$  c-commands  $\beta$  then  $\alpha$  precedes  $\beta$ ,
- and*
- b. If  $\alpha$  precedes  $\beta$  then the terms of  $\alpha$  precede the terms of  $\beta$

Following May's (1985) adjunction formalism, Kayne claims that an adjoined element c-commands (hence it and its terms precede) the structure to which they are adjoined (see also the discussion on (2.11) in the following section). If Munn is correct in that non-first conjuncts are adjuncts, then by the Linear Correspondence Axiom they should precede the category they adjoin to, contrary to fact.

In previous work (Fernández-Salgueiro 2003, 2007) I argued for a linearization algorithm that provides a linear order based on when a phrase or a lexical item is inserted in the derivation, as follows:

(2.9) (adapted from Fernández-Salgueiro 2007)

The phonological features of lexical items are interpreted in the opposite order in which they were inserted in the derivation.

As will be seen in more detail in chapter 4 (section 4.2), the Sideward Movement formalism that I will be assuming (following Nunes 2004) to derive Across-the-Board and Parasitic Gap constructions inherently assumes that the adjunct is built *before* the syntactic object it adjoins to. Thus, even though Munn's analysis seems to be incompatible with Kayne's Linear Correspondence Axiom, it is compatible with my linearization proposal; if the andP is built before the XP it adjoins to in (2.7), then the andP will follow the XP in the linear order computation, following (2.9) (see Fernández-Salgueiro 2007 for further discussion on these issues).

### **2.3. Lowering the adjunction site of non-first conjuncts**

A new empirical problem arises, however, for *all* three of the analyses that I discussed in section 2.1 in the light of Condition C effects. Consider the sentences in (2.10) below:

(2.10)

- a. \*He<sub>i</sub> bought the book and John<sub>i</sub> liked it very much.
- b. His<sub>i</sub> best friend bought the book and John<sub>i</sub> liked it very much.

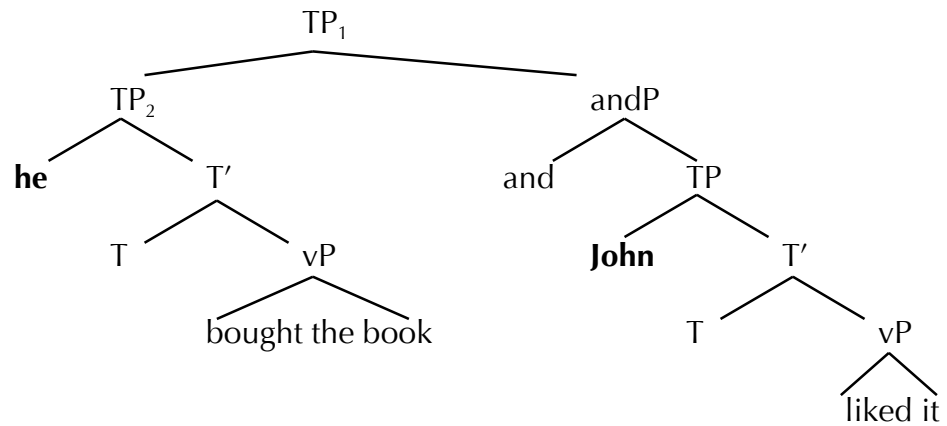
As is well known, R-expressions like *John* cannot be A-bound, i.e., co-indexed with a c-commanding nominal expression, by Condition C of the Binding Theory (Chomsky 1981). Given the contrast illustrated in (2.10), it appears that the subject of the first conjunct does c-command into the second conjunct, yielding a

Condition C violation when both subjects are co-indexed (2.10a). (2.10b) shows that if it is a *non-c-commanding* pronoun that is co-indexed with the R-expression subject of the second conjunct the sentence is grammatical.

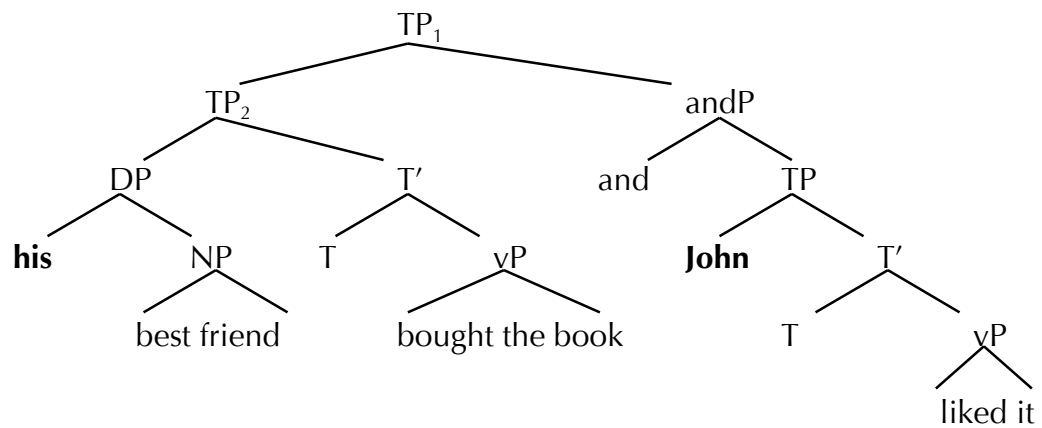
The trees in (2.11) below illustrate the fact that under Munn's analysis, c-command between the pronoun and the R-expression does not obtain in both (2.10a) or (2.10b) (irrelevant structure omitted):

(2.11)

a.



b.



At first sight, it looks as if *he* in (2.11a) would c-command into the second conjunct assuming the segment/category distinction for the adjunction site and that c-command only applies to categories and not to segments (May 1985). However, even under these assumptions *he* in (2.11a) does not c-command into the second conjunct. Let's assume the definition of c-command in (2.12):

(2.12)

$\alpha$  c-commands  $\beta$  if the first category dominating  $\alpha$  dominates  $\beta$

Let  $\alpha$ =*he* and  $\beta$ =*John* in (2.11a). The first category dominating *he* is the two-segment category  $TP_1$ - $TP_2$ , not just  $TP_1$ . Crucially, it is not true that the category  $TP_1$ - $TP_2$  dominates *John* (since only one of its segments,  $TP_1$ , does). Hence, *he* does not c-command *John* in (2.11a).

One way to resolve this incompatibility between Munn's analysis and Condition C of the Binding Theory would be to say that in the case of TP coordination, the second conjunct adjoins to a position lower than the TP. I return to the details of this proposed modification after considering the case of subordination.

Interestingly, the same condition C effects that we can see in (2.10) are present in the case of adverbial (subordinate) clauses, which are assumed to adjoin to  $vP$ , presumably  $VP$  in the case of unaccusatives and passives, as illustrated in (2.13):

(2.13)

- a. \*He<sub>i</sub> bought the book after John<sub>i</sub> saw it.
- b. His<sub>i</sub> best friend bought the book after John<sub>i</sub> saw it.

The fact that coordinate and subordinate structures display the same Condition C effects implies that the same c-command relations obtain between the subject of the first conjunct and the subject of the second conjunct/adverbial clause. In the light of this evidence, I am going to assume that coordinate and subordinate structures actually display the same phrase structure geometry, and hypothesize (2.14) below:

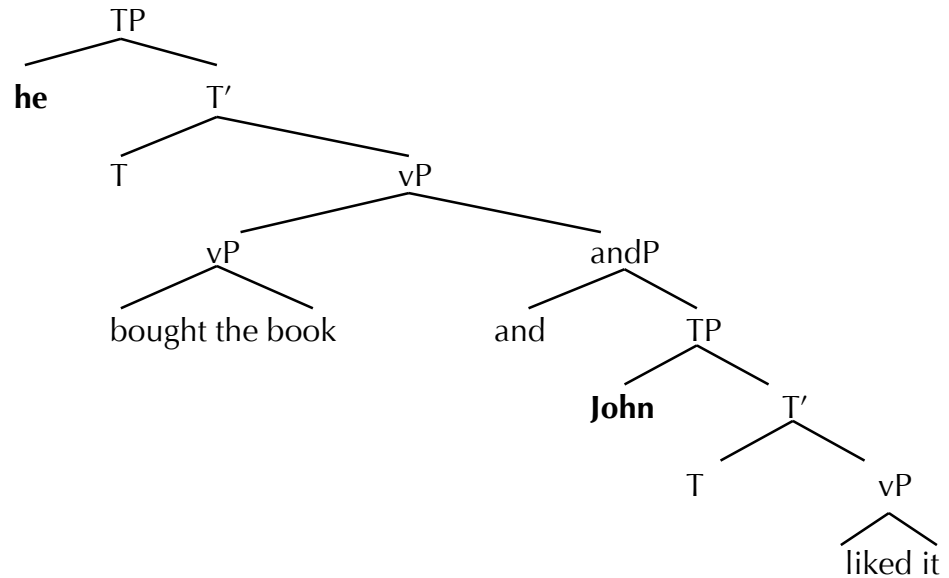
(2.14)

Non-first conjoined TPs adjoin to the same syntactic category as adverbial clauses do, namely vP/VP.

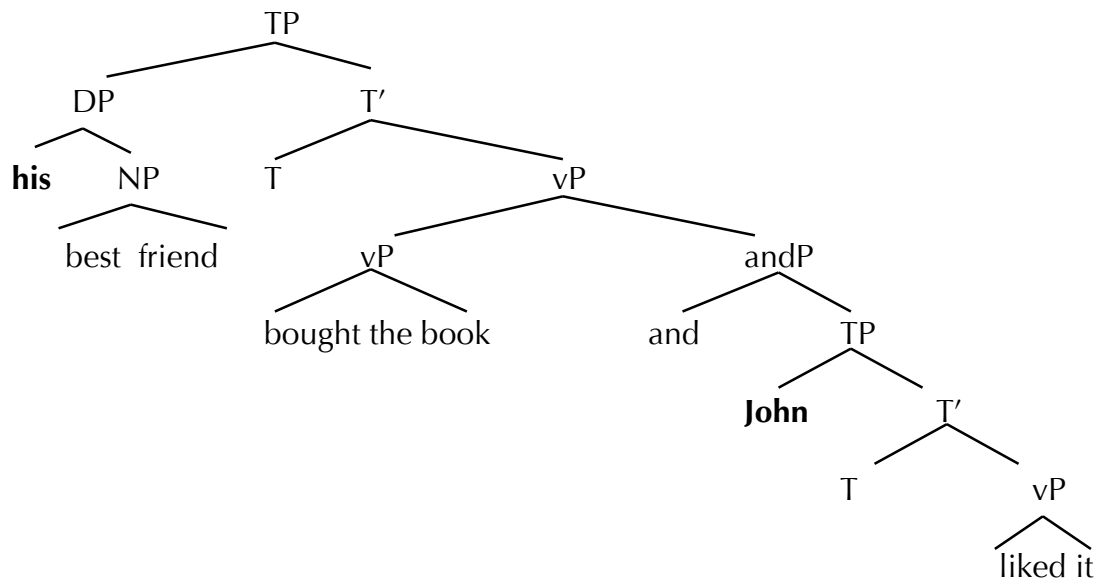
The trees in (2.15) below show the proposed structures for (2.10a) and (2.10b) (irrelevant details omitted). Notice that the subject of the first conjunct *does* c-command into the second conjunct under the proposal in (2.14):

(2.15)

a.



b.



A question that arises under this analysis is the status of a “first conjunct” with respect to constituency. As can be seen in (2.15a), *he bought the book* does not seem to form a constituent. Conjuncts, however, are traditionally considered to be



constituents (cf. *John met Mary and Peter met Sylvia*; *the former really bothered me, though the latter didn't*). There are a couple of ways out of this paradox. One would be to say that *the former* refers back to the event that has been computed by the semantics (i.e., John meeting Mary in the past), not to the actual syntactic structure. Another solution would be to say that *the former* targets the syntactic structure before adjunction takes place, assuming that adjunction can apply acyclically (see sections 3.2, 4.2, and 4.4).

It should also be noted that at this point I do not have a principled explanation for why andP adjuncts should attach to vP/VP. While adverbial clauses, being event modifiers, have to be attached close to the event bearer, non-first conjuncts cannot really be treated in the same way. In this dissertation I provide evidence that non-first conjuncts attach to vP/VP, with the acknowledgment that further research is needed to understand why this is the case.

Further evidence supporting the claim that the subject of the first conjunct c-commands into the second conjunct comes from the binding properties of the non-local anaphor *he himself* in English, as recently analyzed in Fernández-Salgueiro and Marlo (2006). They argue that when *he himself* is interpreted as anaphoric, and not merely emphatic, it co-refers with the closest c-commanding antecedent with which it agrees in  $\phi$ -features. Thus, *he himself* in (2.16a) below has to co-refer with *Bill's brother*, but in (2.16b) it has to co-refer with *John* (it cannot co-refer with *Bill's mother* because of  $\phi$ -feature mismatch and it cannot co-refer with *Bill* because there is no c-command relation between the two elements).

(2.16)

- a. [Bill<sub>i</sub>'s brother]<sub>j</sub> is sure that [he himself]<sub>\*i,j</sub> is smart
- b. John<sub>i</sub> is sure that [Bill<sub>j</sub>'s mother]<sub>k</sub> believes that [he himself]<sub>i,\*j,\*k</sub> is smart

In this respect, consider (2.17):

(2.17)

John<sub>i</sub> is sure [that Bill<sub>j</sub> bought a car and [he himself]<sub>\*i,j</sub> fixed it]

In this example, under the interpretation in which the whole complex CP *that Bill bought a car and that he himself fixed it* is the complement of the adjective *sure*, *he himself* has to co-refer with *Bill*, and cannot co-refer with *John*. If *Bill* did not c-command *he himself*, we would expect *John* to be the antecedent, since *John* is the closest matching c-commanding DP. However, the fact that *he himself* actually has to co-refer with *Bill* provides evidence that *Bill* c-commands *he himself* in (2.17), as (2.14) predicts.

The analysis that I am proposing here, however, is in principle challenged by contrasts concerning bound variable anaphora like the following:

(2.18)

- a. [No man]<sub>i</sub> bought the book after he<sub>i</sub> read it
- b. \*[No man]<sub>i</sub> bought the book and he<sub>i</sub> read it

As can be seen in these examples, a pronoun can be interpreted as a bound variable in a subordinate construction (2.18a) but not in a coordinate construction (2.18b). Assuming, following Higginbotham (1980), that pronouns must be bound by a quantifier in order to render the bound variable interpretation possible, (2.18b) seems to constitute counterevidence for my analysis because under (2.14) a quantifier in subject position of the first conjunct c-commands the pronoun in the second conjunct, incorrectly predicting that the binding is possible.

Interestingly, however, other examples with subordinating conjunctions can be constructed which behave like (2.18b) with respect to the availability of a bound variable reading for the pronoun. (2.19) below is one such example:

(2.19)

\*[No man]<sub>i</sub> read the book although he<sub>i</sub> knew the story was interesting

*Although* is traditionally treated as a subordinating conjunction, yet it patterns with the behavior of the coordinating conjunction *and* in (2.18b). Conversely, we can also find examples with coordinating conjunctions that actually allow the bound variable reading for the pronoun (cf. (2.18a)), like (2.20):

(2.20)

[No student]<sub>i</sub> should ignore those notes or he<sub>i</sub> will fail the exam

Here we have the opposite case, a conjunction traditionally labeled as a “coordinating conjunction” patterns with the behavior of the subordinating conjunction that we see in (2.18a).

This indicates that the availability of the bound variable reading for the pronoun is not a reliable diagnostic to test the geometry of the phrase structure of conjoined vs. subordinate TPs. In order to allow inter-sentential bound variable interpretations at all in the grammar, I will keep on assuming that c-command *has* to hold between a quantifier and a pronoun to render a bound variable interpretation possible. However, intrinsic lexical properties of different conjunctions may block bound variable interpretations in the semantic component that are in principle allowed by the syntactic configuration.

Another possible challenge to the approach I pursue here comes from the distribution of Negative Polarity Items (NPIs). As is well known, elements like *anybody*, *anything*, *anywhere*, and so on require a licenser in a c-commanding position (abstracting away now from examples like *nobody’s articles got published anywhere*). Thus, NPIs have been used to test whether there is c-command between two elements. Consider, in this respect, the sentences in (2.21):

(2.21)

- a. Nobody could have possibly read the book [before anyone wrote it]
- b. \*Nobody has read the book [and anyone wrote it]

Again, we see that NPIs seem to be licensed in the subject position of the second clause by a negative element in subject position in the first clause in the case of subordination, but not in the case of coordination.

However, it is easy to find examples of subordination in which NPIs are not licensed either (that is, patterning with (2.21b)), like the ones in (2.22):

(2.22)

- a. \*Nobody read the book [because anyone left]
- b. \*Nobody ignored my ideas [although anyone said they were good]
- c. \*No student attended the meeting [since anyone announced it]

These examples show that the environments where NPIs are or are not licensed do not correspond to the traditional coordinate vs. subordinate distinction.

From these data I draw a similar conclusion to the one above. I will be assuming, as is standard, that in order for an NPI to be licensed, it must be c-commanded by an appropriate licenser. Again, however, intrinsic lexical properties of different conjunctions may block NPI licensing.

As can be seen, the bound variable reading and the NPI tests give different results depending on the conjunctions that are used, while the Condition C effects seem to remain constant at least in the case of adverbial clauses. Ellen Thompson (1995), however, has argued that (temporal) PP adverbials may attach to different positions, yielding different Condition C effects.

If all of this is on the right track, we can conclude that the traditional distinction between coordinate and subordinate “structures” is not so easy to formally define in syntactic terms. Rather, we need to analyze the intrinsic lexical properties of the different coordinating and subordinating conjunctions and how they affect the syntax of the structure they head for diagnostics involving c-command. As in any other area of syntactic theory, we first need to provide a set of criteria in terms of grammatical behavior, like the ones I have used here:

(2.23)

- a. Condition C effects
- b. Allowing/Blocking a bound variable reading
- c. Licensing of NPIs
- d. Selectional properties (this section)
- e. Availability of fronting (section 2.4)

Then, we will subject cases with each conjunction to these tests and this is how we will formulate hypotheses about their lexical properties. Ultimately, we will come to the non-surprising conclusion that their different grammatical behavior will be the result of the interaction between general properties of the syntactic module and intrinsic lexical properties of the lexical items this module manipulates, rather than the result of obeying construction-specific notions such as “coordination” and “subordination”.

One difference that does seem to support the traditional coordination/subordination distinction is the different conjunctions' selectional properties; a coordinating conjunction can conjoin two clauses headed by a subordinating conjunction but a subordinating conjunction cannot link two coordinated clauses, as shown in (2.24):

(2.24)

- a. That happened [[after John left] **and** [before Peter arrived]]
- b. \*That happened [[and John left] **before** [and Peter arrived]]

The contrast in (2.24) is explained if subordinating conjunctions are actually Complementizers (and thus generally take TPs as complements), whereas coordinating ones can occupy a position higher than Comp in the tree and actually can conjoin any two (or more) syntactic objects. If this is on the right track, this difference between coordination and subordination reduces to lexical properties of different lexical items; some lexical items like *and* and *or* can take virtually any phrase as complement (including CP, as in (2.24a) above), while other lexical items like *after*, *as* and *because* for example take TPs as complements.

#### **2.4. Where is the coordination/subordination distinction to be found, then?**

One consequence of the unified syntactic analysis of subordination and coordination that I am proposing here then is that any differences in the syntactic behavior of conjoined TPs vs. adverbial clauses (i.e., coordination vs. subordination) cannot be derived from their phrase structure geometry. Munn (1993) also notes this as a consequence of his approach to coordination, although given his analysis there could still be a crucial purely syntactic difference, namely, attachment site. Under Munn's approach non-first sentential conjuncts would adjoin to TP, whereas subordinate clauses would adjoin to VP/vP. Under my proposed analysis of TP-coordination, however, there is nothing at all in the syntax that distinguishes coordination from subordination, not even in terms of adjunction site.

It seems then that at least some of the differences between coordination and subordination are actually due to lexical properties of coordinating vs. subordinating conjunctions, and actually more fine-grained lexical differences within these conjunction classes.

One example of this difference in syntactic behavior would be the availability of fronting. An important traditional distinction between coordination and subordination is that an adverbial clause can be fronted, while a second conjunct cannot:



(2.25)

- a. [Because we went home]<sub>i</sub> the party ended really early  $t_i$
- b. \*[And we went home]<sub>i</sub> the party ended really early  $t_i$

This is not universally true, however. In Spanish and Galician, for example, the subordinating conjunctions *porque* and *que* are synonymous (*porque* being more formal than *que*), yet only the former can be fronted (the Spanish examples below involve clefting). A similar contrast also applies in French (Chris Collins, personal communication).

(2.26)

- a. Me voy        porque        estoy cansado  
leave.1sg    because        am    tired
- b. Me voy        que                estoy cansado  
leave.1sg    because        am    tired  
  
'I'm leaving because I'm tired'
- c. Es        porque        estoy cansado        que    me voy  
is        because        am    tired                that    leave.1sg
- d. \*Es        que                estoy cansado        que    me voy  
Is        because        am    tired                that    leave.1sg  
  
'It's because I'm tired that I'm leaving'

Availability of fronting, then, does not seem to be a reliable cross-linguistic diagnostic for determining coordination vs. subordination status either.

For the purposes of the present study, the important conclusion is that Condition C effects obtain in both coordinate and subordinate structures, a uniformity that I have argued here comes from the fact that all TPs that are introduced by conjunctions invariably adjoin to vP, i.e., they are all vP adjuncts.

Given the evidence presented above related to Condition C effects, the non-local anaphor *he himself*, availability of bound variable readings and licensing of NPIs, an interesting future research path would be to analyze the lexical properties of different conjunctions and try to understand why it is that certain syntactic relations, which are in principle allowed by a syntactic configuration in which c-command obtains, cannot take place when those conjunctions are used, despite the argued satisfaction of the c-command requirement.

## **2.5. Consequences of the “lowered adjunction” analysis**

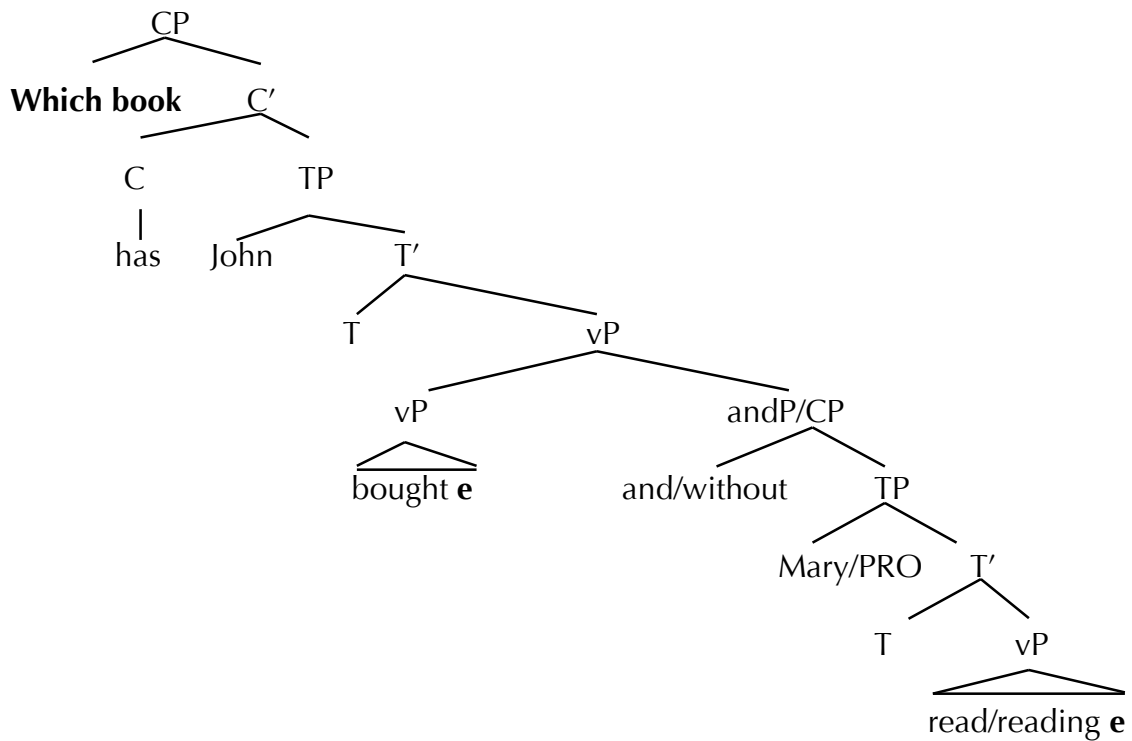
One important consequence of the adjunction analysis of coordination is that extraction from a non-first conjunct is now redundantly ruled out by two independent principles of the narrow syntax, the Condition on Extraction Domains (Huang 1982) and the Coordinate Structure Constraint (Ross 1967). In chapter 3 of this dissertation, I examine this redundancy in detail. I will also argue that extraction from the first conjunct and extraction from a non-first conjunct are banned by very different principles.

Another consequence of the analysis that I have proposed here is that Across-the-Board (ATB) and Parasitic Gap (PG) constructions display the same syntactic structure. In both cases, there is an element in Spec-CP that is related to two gaps, one inside the main clause and another one inside the clause that is adjoined to vP. This is illustrated in the tree in (2.28) for the sentences in (2.27):

(2.27) (e stands for empty category)

- a. Which book has John bought e and Mary read e? (ATB)
- b. Which book has John bought e without PRO reading e? (PG)

(2.28)



Interestingly, there have been two previous attempts to unify the analysis of ATB and PG constructions, the Null Operator analysis (Munn 1993), and the Sideward Movement analysis (Nunes 2004). An interesting aspect of the proposal that I presented in this chapter is that it substantiates this intuition regarding unification in the theory; the analysis of these constructions should be the same because the syntactic configuration is identical (only the lexical items are different, as shown in (2.28)). In chapter 4 of this dissertation I evaluate these two approaches and argue that the Sideward Movement analysis can provide a more straightforward explanation of the crossover phenomena that we find in ATB and PG constructions.

## CHAPTER 3

### Revisiting the Coordinate Structure Constraint

In this chapter I examine the consequences of the analysis that I proposed in the previous section for our understanding of the Coordinate Structure Constraint (CSC). I argue, following Munn (1993), that no such constraint exists *in the syntax* and I also argue that extraction from a first conjunct and extraction from a non-first conjunct are banned by principles of a different nature.

#### 3.1. Exploring the redundancy between the CSC and the Condition on Extraction Domains

The CSC was first proposed by Ross (1967). Ross observed that a coordinate structure is an island for movement, with the exception of ATB extraction (which I deal with from section 3.2 on). One important consequence of *any* version of the adjunction approach to coordination is that extraction from a non-first conjunct is ruled out redundantly by two different principles: the CSC and the Condition on Extraction Domains (CED). For concreteness, let's assume the following definitions of these two principles:

(3.1) Ross's (1967) Coordinate Structure Constraint (CSC)

In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

(3.2) Condition on Extraction Domains (CED) (adapted from Huang 1982)

- (i) Adjunct Condition: No extraction is possible out of an adjunct.
- (ii) Subject Condition: No extraction is possible out of a subject.

Consider now the sentences in (3.3). Once the non-first conjunct *and Peter bought which book* in (3.3a) is (re-)analyzed as an adjunct, the sentence would already be ruled out by the same constraint (3.2i) that disallows extraction of *which book* in (3.3b):

(3.3)

- a. \***[Which book]** has Mary read a magazine **[and Peter bought \_ ]**?
- b. \***[Which book]** has Mary read a magazine **[after PRO buying \_ ]**?

The question arises now whether we should keep both principles in our theory. If we just focus on extraction from the second conjunct, a natural route to take would be to reduce one of the principles to the other, in the spirit of Occam's razor. If we choose to maintain the CSC and dispense with the CED, we need an alternative explanation for why extraction out of subjects and adjuncts is blocked. Conversely,

if we choose to maintain the CED and dispense with the CSC, we would need an alternative explanation for why extraction out of a first conjunct only is not possible.

A third option to explore would be to maintain *both* principles in the syntax. Moreover, a natural extension of this claim would be that extraction from a non-first conjunct should generally be significantly worse than extraction from a first conjunct (T. Daniel Seely, personal communication). A potential problem with this idea is that it is hard to relate different syntactic violations to degrees of ungrammaticality, on two levels. On the one hand, the same violation can sound better or worse depending on other factors like choice of lexical items or the semantics of the sentence. On the other hand, even if we controlled for such variables, we might be unable to compare degrees of ungrammaticality due to violations of principles that apply at different levels of representation or in different components of the grammar. In the case at hand, I will explore the idea that the CSC could be more a semantic restriction than it is a syntactic one (see Munn 1993 and the discussion below). If this is correct, it amounts to saying (in more Minimalist terms) that this restriction applies at LF, rather than in the syntax. As I will argue in the following section, this approach also makes the prediction that extraction from the first conjunct does not lead to ungrammaticality if the semantic restriction can be obeyed and the violation circumvented.

As Minimalism seeks to explore syntactic relations and the constraints applying to them without resorting to construction specific constraints, maintaining

the CED in the syntax should be the preferred option, in my view. Before I discuss how the CED is a more natural explanation, in the sense that it follows from the way syntactic objects are concatenated in a derivation, let's see how the CSC is actually unformulable as a syntactic principle under the adjunction analysis of coordination.

As Munn (1993:97) already notes, the adjunction analysis of coordination removes the notion of 'coordinate structure' and therefore it also invalidates any attempt to construct a syntactic principle that refers to such a notion. He then redefines the CSC as a purely semantic constraint (see section 3.2 below). The structure of the sentence in (3.3a) above, for example, would contain a TP and an andP, as shown in (3.4):

(3.4)

\*[Which book] has <sub>TP</sub>[Mary read a magazine <sub>andP</sub>[and Peter bought \_ ]]?

We obviously do not want to say that extraction out of a TP is impossible, since this would involve massive undergeneration. The other possibility then is to say that it is extraction from the andP that is blocked (by the CED). As can be seen, under this approach it is hard to treat the whole syntactic object *Mary has read a magazine and Peter bought which book* as an island for movement, which would be the claim made by the CSC, since the syntactic label for this constituent is just TP, and we know that TP does not block extraction.

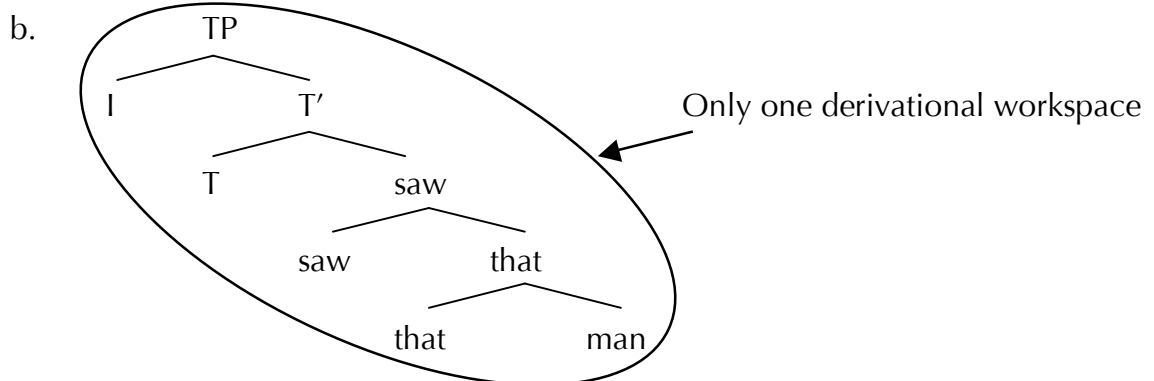


Conversely, the CED seems to be different from (and more theoretically appealing than) the CSC in that it does not make reference to a specific syntactic category. It does stipulate that adjuncts and subjects (specifiers) are different from, say, complements, but this distinction can be related to how these different elements emerge in a derivation.

A main difference between complements on the one hand and adjuncts and specifiers on the other is that complements do not require a separate derivational workspace to be built when they're complex i.e., when they contain more than one lexical item. This is shown in (3.5-3.7) for sentences with phrasal complements, phrasal specifiers, and phrasal adjuncts, respectively (irrelevant details omitted):

(3.5)

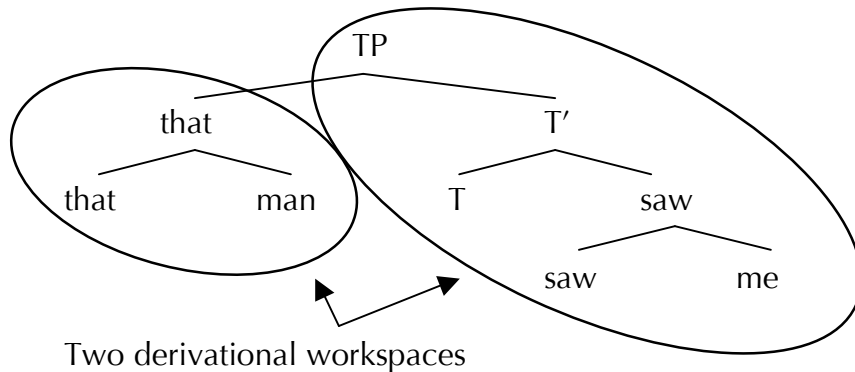
a. I saw **[that man]**



(3.6)

a. **[That man]** saw me

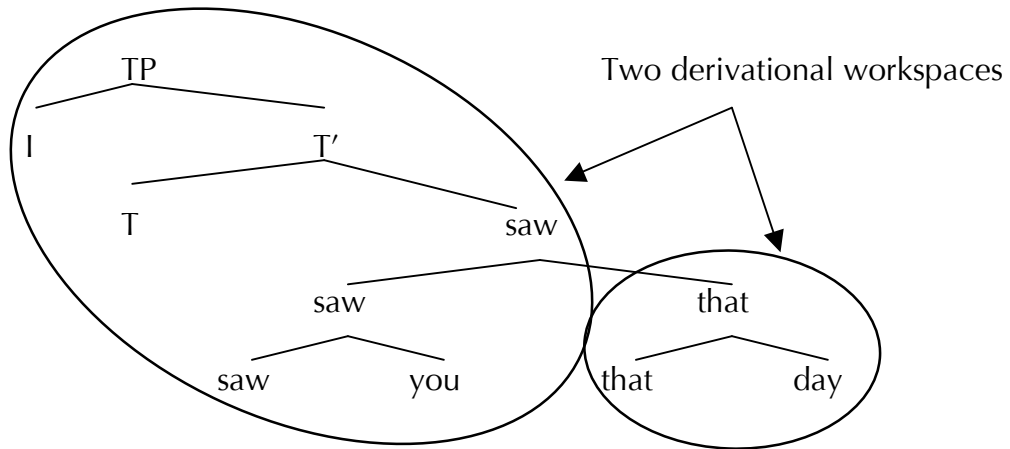
b.



(3.7)

a. I saw you **[that day]**

b.



This relationship between derivational workspaces and the CED has been exploited to account for these extraction asymmetries. Uriagereka (1999), for example, claims that separate workspaces require a separate application of Spell-Out. The main theoretical motivation for this approach is that it eliminates the second part of

Kayne's Linear Correspondence Axiom (cf. (2.8b) in section 2.2 above). If this is on the right track, it makes sense that extraction out of an adjunct or specifier is not possible; if they are spelled out independently, their internal structure cannot be available for syntactic operations inside the main tree.

Even without adopting this radical view of Spell-Out, there is a sense in which extraction islands can be related to how syntactic objects are built in a derivation. This makes the CED a more natural principle than the CSC, since it refers to construction-independent derivational notions that are readily available in the narrow syntax, rather to a specific construction that is not available in the syntactic structure once the adjunction analysis of coordination is adopted.

Since we now have to account for why extraction from a first conjunct is also disallowed, one possibility would be to hypothesize something like (3.8):

(3.8)

Extraction out of a syntactic object is not possible after another object has been adjoined to it.

In other words, we would be claiming that even though a first conjunct is not an island for movement, it becomes an island when the second conjunct is adjoined to it. Although stipulative, this explanation would indeed account for why (3.8a) is allowed but (3.8b) is not; once the adjunct *and John read a magazine yesterday* is attached, no extraction of *which book* would be possible:

(3.9)

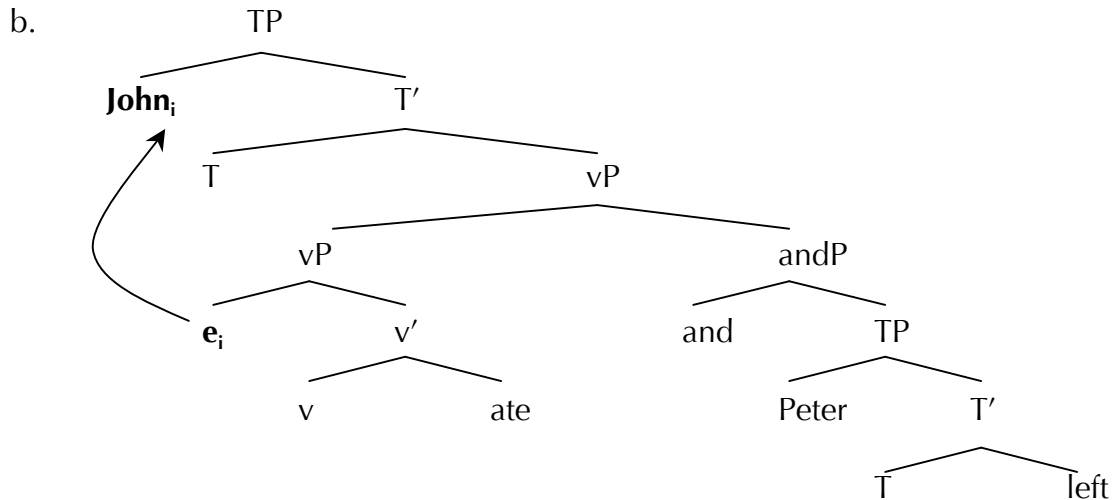
- a. Which book has Peter bought e?
- b. \*Which book has Peter bought e and John read a magazine?

This account would face two problems, however. The first problem is that, if Lebeaux (1988) is right, adjuncts do not have to be cyclically inserted in the derivation (see also section 3.2 below). If this is correct, the movement of *which book* in (3.9b) could occur prior to the adjunction attachment operation, incorrectly predicting (3.9b) to be grammatical.

A second potential problem with the hypothesis in (3.8) is that, since the adjunction operation targets the vP, we would predict that A-movement of the subject to Spec-TP is impossible, contrary to fact. This is shown in (3.10b) below for a simple sentence like (3.10a), where the subject does undergo A-movement to Spec-TP:

(3.10)

a. John ate and Peter left



It should be noted, however, that this is only a problem if adjunction is cyclic and thus occurs before A-Movement of the subject. In other words, if (3.8) is the right approach, we would have to claim that all andPs are acyclically adjoined, which seems like a strong claim.

We still need to find an explanation then for why extraction from a first conjunct is blocked. This is the focus of the sections that follow.

### 3.2. On extraction from a “first conjunct”? The parallelism requirement

In this section I argue that extraction from the first conjunct can be accounted for if we adopt a version of Hornstein and Nunes’ (2002) account of the main differences between Across-the-Board (ATB) and Parasitic Gap (PG) constructions. Hornstein

and Nunes argue that ATB constructions, unlike PG constructions, obey a parallelism requirement. This is illustrated in the examples in (3.10):

(3.10)

- a. [Which paper]<sub>i</sub> have you read e<sub>i</sub> and Mary recommended e<sub>j</sub>/\*[this book]?
- b. [Which paper]<sub>i</sub> did you read e<sub>i</sub> before filing e<sub>j</sub>/[this book]?

As can be seen, (3.10a) requires an object gap in the second conjunct, while in (3.10b) the object position can be a gap or an overt phrase. Hornstein and Nunes adopt the hypothesis (which I will revise later) that ATB constructions impose some kind of parallelism restriction, which has been formalized in different ways in the literature. Williams (1978) formalizes this restriction in the ATB formalism itself, while Munn (1993:97) claims that it is a semantic restriction:

“[...] there are good reasons to believe that the CSC is not a constraint on movement, [but (GF-S)] is much more a semantic constraint which imposes semantic identity on elements that may be conjoined.”

Hornstein and Nunes adapt this idea to the Minimalist framework and formalize this parallelism in terms of a legibility condition at the LF Interface.

If this approach is correct, we can claim that extraction from the first conjunct is not blocked in the syntax, but at LF, and only in derivations in which the parallel syntactic element has not been extracted in the non-first conjunct(s).

As far as the narrow syntax is concerned then, there is nothing wrong in the syntax with extracting an element out of the first conjunct under the analysis entertained here: the first conjunct, unlike the second one, *is not an island*. What is wrong, adopting Hornstein and Nunes' proposal, is that the resulting LF-representation would violate the parallelism requirement, which can be defined as in (3.11):

(3.11) Parallelism requirement (adapted from Hornstein and Nunes 2002)

- a. All conjuncts must be parallel with respect to presence or absence of a gap,  
*and*
- b. The gaps (if any) have to correspond to the same grammatical function in all conjuncts.

If this is on the right track, we reach the (non-trivial) conclusions in (3.12):

(3.12)

- a. The ban on extraction from a non-first conjunct is derivational, i.e., it is a constraint on the movement operation (in the narrow syntax).
- b. The ban on extraction from a first conjunct is representational, i.e., it is a constraint on the output of the derivation (at LF).

Once the adjunction analysis of coordination is adopted, one is forced to view the constraint ruling out extraction from a first conjunct as a representational constraint, given that grammars display late adjunction. Consider the sentences in (3.13) (see Lebeaux 1988):

(3.13)

- a. [[**Which claim**]<sub>i</sub> [that John<sub>i</sub> made]] was he<sub>i</sub> willing to discuss e<sub>j</sub>?
- b. \*[[**Which claim [that John<sub>i</sub> was asleep]**]<sub>i</sub> was he<sub>i</sub> willing to discuss e<sub>j</sub>?

As can be seen, only (3.13a) in which the bracketed element is an adjunct allows a reading in which *John* and *he* are co-referent. Lebeaux (1988) claims that the adjunct *that John made* can be connected to *which claim* after *which claim* moves to Spec-CP, avoiding a syntactic configuration in which *he* c-commands *John* at some point on the derivation yielding a Condition C violation. Conversely, the complement *that John was asleep* cannot be “late merged” but it is already the complement of *which claim* in its base position. Since this would mean that *he* c-commands *John* before movement, (3.13b) is correctly predicted to be a Condition C violation.

If non-first conjuncts, being actually adjuncts, can be inserted acyclically in the derivation, there is no way to rule out extraction out of the first conjunct derivationally. Movement of, say, a wh-element to Spec-CP could apply before the adjunct is attached, making it impossible to explain at that point in the derivation



why that movement is not allowed. Thus, the only way to rule out such a derivation is to wait until LF (and check whether parallelism is obeyed). This is illustrated in the derivation in (3.14b), for the ungrammatical sentence in (3.14a):

(3.14)

- a. \*Which book has John bought [and Mary read a magazine]?
- b. (i) Adjunct is built: [and Mary read a magazine]
- (ii) CP is built:  ${}_{CP}[C\ {}_{TP}[\text{John has } {}_{vP}[\text{bought [which book]}]]]$
- (iii) [which book] moves to Spec-CP, and T-C movement applies:  
 ${}_{CP}[[\text{which book}]_j \text{ has}_i {}_{TP}[\text{John } e_i {}_{vP}[\text{bought } e_j]]]$
- (iv) Adjunct attaches to vP:  
 ${}_{CP}[[\text{which book}]_j \text{ has}_i {}_{TP}[\text{John } e_i {}_{vP}[\text{bought } e_j]$   
 $\text{andP}[\text{and Mary read a magazine}]]]$

One could argue that under the approach that I develop here, extraction from a non-first conjunct only is also redundantly ruled out by the CED and the parallelism requirement. However, since the CED is a derivational principle, the claim would be that once the CED is violated, the derivation is canceled, and so no LF representation is ever derived.

A natural extension of (3.11) would be to assume that resumptive pronouns are also allowed instead of gaps. If this is true, the approach that I propose here makes an interesting prediction regarding resumptive pronouns: they should be

allowed in non-first conjuncts, but not in the first conjunct. Since extraction from the first conjunct is allowed in the narrow syntax, we would expect examples in which extraction from the first conjunct applies to be grammatical, as long as there is a resumptive pronoun in the corresponding position in the second conjunct, thus ensuring parallelism.

This prediction is indeed borne out, and such asymmetry has been noted several times in the literature (Sells 1984, Munn 1993, 1999, Citko 2005, among others). Hebrew, for example, displays such an asymmetry (see Munn 1999):

(3.15)

- a. Hais<sub>i</sub> se Rina roca e<sub>i</sub> ve ohevet oto<sub>i</sub> yoter mikulam  
the-man that Rina wants and loves him more-than anyone
- b. \*Hais<sub>i</sub> se Rina roca oto<sub>i</sub> ve ohevet e<sub>i</sub> yoter mikulam  
the-man that Rina wants him and loves more-than anyone  
‘The man that Rina wants and loves more than anyone’

As can be seen, extraction of *hais* in (3.15a) does not lead to ungrammaticality, precisely because the parallelism between both conjuncts still holds. Conversely, (3.15b) is ungrammatical, because extraction from the second conjunct is disallowed in the narrow syntax.

Interestingly, in Romance languages having clitics co-indexed with the extracted phrase are much more acceptable in the second conjunct than in the first

conjunct, mirroring the data from Hebrew just discussed (using Spanish now to illustrate this):

(3.16)

- a. ¿[Qué libro]<sub>i</sub> ha comprado e<sub>i</sub> Juan y querrá leer e<sub>i</sub> María?  
 Which book has bought Juan and want.fut read María
- b. ?¿[Qué libro]<sub>i</sub> ha comprado Juan e<sub>i</sub> y querrá leerlo<sub>i</sub> María?  
 Which book has bought Juan<sub>i</sub> and want.fut read-it María
- c. \*¿Qué libro lo<sub>i</sub> ha comprado Juan y querrá leer e<sub>i</sub> María?  
 Which book it has bought Juan and want.fut read María  
 ‘Which is the x, x a book, such that Juan has bought x and María will read x?’

The most interesting example here is (3.16b), in which extraction of the wh-element from *only* the first conjunct has applied, but parallelism still obtains at the LF interface provided the clitic *lo* ‘it’ in the second conjunct is interpreted as a bound variable. Interestingly, if the clitic in (3.16b) is interpreted pronominally, the sentence is completely ungrammatical.

In conclusion, as evidenced from the differences between ATB and PG constructions and the contrasts and asymmetries discussed above, the parallelism requirement seems to be unavoidable. In section 3.4, however, I will present data from Spanish that forces a revision of this parallelism requirement. Before that, I argue in section 3.3 below that the parallelism requirement also applies at the PF

interface, as can be seen in languages with more variable word order than English, like Spanish, for example.

### 3.3. A parallelism requirement at PF: word order restrictions in ATB extraction in Spanish

In this section I discuss some issues that arise regarding the scope of the parallelism requirement. Interestingly, there is evidence from languages with more options regarding word order that the parallelism requirement in ATB constructions goes beyond legibility at the LF interface. Consider the Spanish data in (3.17) below (similar effects can be observed in other Romance languages, such as Catalan or Galician):

(3.17)

- a. ¿A quién pudo **invitar la secretaria** y **saludar el jefe**?  
Who could invite the secretary and greet the boss
- b. ¿A quién pudo **la secretaria invitar** y **el jefe saludar**?  
Who could the secretary invite and the boss greet
- c. ??¿A quién pudo **invitar la secretaria** y **el jefe saludar**?  
Who could invite the secretary and the boss greet
- d. ??¿A quién pudo **la secretaria invitar** y **saludar el jefe**?  
Who could the secretary invite and greet the boss  
'Who could the secretary invite and the boss greet?'

As can be seen, (3.17a) and (3.17b), in which the relative order of verb and subject is the same in the first and second conjuncts, are fully grammatical, while (3.17c) and (3.17d), in which that relative order has not been preserved, are degraded.

Further evidence for the claim that verb and subject has to display the same order comes from cases involving null subjects:

(3.18)

a. ¿A quién pudieron **pro**<sub>i</sub> **invitar** y **los** **jefes**<sub>j</sub> **saludar**?

Whom could.3pl invite and the bosses greet

b. \*¿A quién pudieron **pro**<sub>i</sub> **invitar** y **saludar** **los** **jefes**<sub>j</sub>?

Whom could.3pl invite and greet the bosses

‘Who could they invite and the bosses greet?’

As can be seen in (3.18a) and (3.18b), when the subject of the first conjunct is null, the subject of the second conjunct has to be preverbal. Assuming that null subjects in finite clauses (*pro*) are licensed only in Spec-TP (see Fernández-Salgueiro 2005 and references therein), it is also the case that grammaticality in these cases obtains only when the relative order of subject and verb is preserved.

The data above suggest that the parallelism requirement on the two conjuncts is not only of a semantic, LF interface-related nature, but is also operative at the PF interface, imposing certain word order restrictions, which we can detect in Spanish for example, a language with more variable word order. An interesting

research path to pursue then would be to investigate more languages with variable word order and see to what extent these PF parallelism restrictions apply universally.

### 3.4. ATB extraction in English vs. Romance

In this section I compare the restrictions applying to ATB constructions in English vs. Romance (I will be focusing on Spanish here) and discuss their implications for our understanding of the parallelism requirement at LF.

The first observation that should be made is that ATB extraction in Romance is less restrictive than in English, with respect to two parameters. On the one hand, ATB extraction with matrix questions in Spanish allows for the realization of independent Tense in non-first conjuncts, while in English the Tense of the non-first conjunct is obligatorily dependent on the Tense of the first conjunct. This is illustrated in (3.19) and (3.20) below:

(3.19) Spanish

- a.    ¿Qué libro ha comprado Susana y leído María?  
      Whichbook has bought Susana and read María  
      ‘Which book has Susana bought and María read?’

- b. ¿Qué libro ha comprado Susana y leyó María?  
 Whichbook has bought Susana and read.past María  
 'Which is the x, x a book, such that Susana has bought x and María read x?'

(3.20) English

- a. Which book has Sue bought and Mary read? (read = past participle)  
 b. \*Which book has Sue bought and Mary read? (read = past tense)

This contrast between Spanish and English vanishes when we consider ATB extraction in *embedded* wh-questions. In these cases both languages allow for the realization of independent Tense in non-first conjuncts, as shown in (3.21) and (3.22) below:

(3.21) Spanish

- a. Me pregunto qué libro compró Susana y leyó María  
 Wonder.1sg which book bought Susana and read María  
 'I wonder which book Susana bought and Mary read'
- b. Me pregunto qué libro compró Susana y leerá María  
 Wonder.1sg which book bought Susana and read.fut María  
 'I wonder which book Susana bought and Mary is reading'

(3.22) English

- a. I wonder which book Susan bought and Mary read (read = past participle)
- b. I wonder which book Susan bought and Mary read (read = past tense)

As of now, I do not have an account of this parametric difference between English and Spanish, an issue that I hope to address in future research. As to the question of why English allows independent Tense in non-first conjuncts only in the case of embedded wh-questions, it seems clear that there is an interaction with T-C movement. When T-to-C movement applies (matrix questions) then all conjuncts have to have the same Tense interpretation, which obviously corresponds to whatever Tense head undergoes T-to-C movement.

This is actually one of the arguments Nunes (2001, 2004) uses to argue for a Sideward Movement approach to ATB extraction and against a Null Operator analysis, as I will discuss in chapter 4. The question that still remains (even assuming ATB structures are derived through Sideward Movement) then is why it is possible in Spanish, but not in English, to do Sideward Movement of an argument to Spec-CP without also triggering ATB T-to-C movement.

Another important respect in which Spanish differs from English is that the two gaps in Spanish need not be parallel (in either matrix or indirect wh-questions), as shown below:



(3.23) Spanish

- a. ¿Qué libro ha leído Juan y sorprendido a María?  
Which book has read Juan and surprised María  
'Which is the x, x a book, such that Juan has read x and x has surprised María?'
- b. ¿Qué libro ha sorprendido a Juan y leído María?  
Which book has surprised Juan and read María  
'Which is the x, x a book, such that x has surprised Juan and María has read x?'
- c. Me pregunto qué libro ha leído Juan y sorprendido a María  
Wonder.1sg which book has read Juan and surprised María  
'I wonder which is the x, x a book, such that Juan has read x and x has surprised María'
- d. Me pregunto qué libro ha sorprendido a Juan y leído María  
Wonder.1sg which book has surprised Juan and read María  
'I wonder which is the x, x a book, such that x has surprised Juan and Mary has read x'

All these examples are acceptable in Spanish but their analogs are unacceptable in English. Besides the interesting parametric difference here, these cases provide evidence against the parallelism requirement as we defined it in section 3.2 above. The examples in (3.23) all involve subject-object gap mismatches that clearly

violate the parallelism requirement at the LF interface, yet they are nonetheless grammatical.

Before trying to account for this difference, it is interesting to notice the following contrasts in Spanish (I am illustrating this contrast in direct wh-questions, even though it also applies in embedded wh-question contexts):

(3.24) Spanish

a. ¿Qué libro pudo **sorprender** **a Juan** y **leer** **María**?

Whichbook could surprise Juan and read Mary

b. ??¿Qué libro pudo **sorprender** **a Juan** y **María leer**?

Which book could surprise Juan and Mary read

'Which is the  $x$ ,  $x$  a book, such that  $x$  could surprise Juan and María could read  $x$ ?'

c. ¿Qué libro pudo **leer** **Juan** y **sorprender** **a María**?

Whichbook could read Juan and surprise Mary

d. ??¿Qué libro pudo **Juan leer** y **sorprender** **a María**?

Which book could Juan read and surprise Mary

'Which is the  $x$ ,  $x$  a book, such that Juan could read  $x$  and  $x$  could surprise Mary?'

The data above show that even though Spanish allows the two gaps to correspond to different grammatical functions, the order verb+DP is still enforced at the PF interface, regardless of whether the DP is subject or object.

If this is on the right track, the parallelism requirement needs to be redefined: it is not really the case that the two gaps have to correspond to the same grammatical function; if it were, the Spanish data in (3.23) and (3.24) would remain unexplained. What seems to be enforced in Spanish is actually that the non-extracted elements in the two conjuncts display the same order in the PF representation.

Notice that we are still claiming that conjuncts need to be parallel with respect to presence or absence of a gap. If this is correct, this revision of the parallelism requirement also accounts for the observed pattern in the English cases. The idea that conjuncts needed to be parallel with respect to their gap position was then epiphenomenal. The reason that in English the gaps look like they need to correspond to the same grammatical function is that this is the only way that non-extracted elements can display the same order in the PF representation, given the rigidity of word order in the language. Again, once another language is examined in which this rigidity does not apply, the parallelism effects at PF are more apparent.

How can this be formalized in the PF component? We need to find a point in the derivation to PF at which both categorial and linear order information are

relevant. A straightforward possibility here then is to evaluate this requirement when the linearization algorithm applies.

There is evidence that categorial information matters when linearization applies. Richards (2001), for example, argues for a linearization algorithm that takes into account only the categorial label of the syntactic objects to be linearized. His approach is motivated by cases in which linearization fails when two syntactic objects with the same category would appear adjacent *if* linearization could apply. To briefly illustrate this, consider the examples in (3.25) below:

(3.25) (from Richards 2001)

- a. The singing<sub>PP</sub>[of the children]
- b. The singing<sub>PP</sub>[of songs]
- c. \*The singing<sub>PP</sub>[of the children]<sub>PP</sub>[of songs]

(3.25a) and (3.25b) show that gerunds allow one argument (either logical subject/Agent or logical object/Theme) as its surface complement. However, when more than one argument surfaces (3.25c), the result is ungrammatical. It is hard to explain the ungrammaticality of (3.25c) on semantic grounds, since the verb on which the gerund *singing* is based is a two-place predicate (cf. *the children sing songs*). It is also hard to rule it out in the syntax, since this is a case of recursive PP merger. However, if it is true that all that the linearization algorithm sees is <PP, PP> these data are accounted for.

Let's claim then that when conjuncts are linearized, PF must receive parallel linear order instructions for them. To illustrate how this approach would work, consider the ungrammatical case in (3.26) below:

(3.26)

\*[Which book]<sub>j</sub> has<sub>i</sub> Sue *t<sub>i</sub>* bought *t<sub>j</sub>* and *t<sub>j</sub>* surprised Mary?

When linearization applies to determine the linear order of the syntactic objects inside the two TPs [Sue (has) bought (which book)] and [(which book) surprised Mary], the linear order in (3.27) below is computed. As can be seen, V and DP display opposite orders in this PF representation:

(3.27)

<DP, V, and, V, DP>

Suppose now that the language allows one of the conjuncts to display the opposite order, as was the case in the Spanish examples in (3.24) above. Then a linear order of either form <V, DP, and, V, DP> or <DP, V, and, DP, V> would be computed, thus obeying parallelism in the PF representation.

I thus reformulate the parallelism requirement as follows:

(3.28)

- (i) Conjuncts need to be parallel with respect to presence or absence of a gap (or have a variable co-indexed with the gap),  
*and*
- (ii) At PF, non-extracted elements have to display the same relative order in all conjuncts.

The condition in (ii) raises the question of how PF knows which elements have or have not been extracted, since PF presumably has no access to information regarding movement operations that apply in the narrow syntax. Abstracting away now from the precise mechanism that is responsible for linearization, some elements are spelled out with their full phonological make-up, while some others are spelled out as traces, a kind of information that PF definitely has access to (see section 4.2 for some details regarding Nunes' approach to chain linearization).

### **3.5. Some remaining issues**

In this section I will discuss three main issues that arise once we adopt the approach that I have argued for in this chapter.

The first two issues concern the nature of the parallelism requirement at PF. On the one hand, it is somewhat surprising that this requirement applies once extraction (or more accurately A'-Movement) has occurred. Interestingly, when two clauses are coordinated in which no A'-movement has applied, no such

requirement needs to be met. An example of this can be found in the Spanish sentence in (3.29), in which the first conjunct displays the order <V, DP, PP>, and the second conjunct displays the order <DP, V>:

(3.29)

Estábamos esperando todos en la cola y Pedro se piró  
Were.1pl waiting all in the line and Pedro left.3sg  
'We were all waiting in line and Pedro left'

It is not clear then, why it is only when extraction occurs that linear order parallelism is enforced between two syntactic objects that are actually built independently (in different derivational workspaces). However, as will be seen in the next chapter, under a Sideward Movement approach to ATB extraction syntactic objects can be copied from one workspace and merged in another workspace. Though still far from explaining why parallelism is enforced only when extraction applies, at least this approach predicts that there is some syntactic relation between two different workspaces. A future research question, one that lies beyond the scope of this dissertation, is whether this can be exploited to account for when the parallelism requirement applies or does not apply.

Another important question that arises concerns the use of construction-specific mechanisms in the grammar. As is evident from my re-formulation of the parallelism requirement in (3.28) above, the question still remains of why sentences

containing certain adjuncts headed by *and*, *but*, or *or* (i.e., “non-first conjuncts”) require any kind of parallelism while adjuncts headed by *without*, *after*, or *before* (i.e., “adverbial clauses”) do not, as shown in (3.30) (modified from (3.10) above):

(3.30)

- a. [Which paper]<sub>i</sub> have you read t<sub>i</sub> and/but Mary recommended t<sub>i</sub>/\*[this book]?
- b. [Which paper]<sub>i</sub> did you read t<sub>i</sub> without/before/after filing t<sub>i</sub>/[this book]?

The discussion leading to the re-formulation of the parallelism requirement shows how hard it is to get rid of construction-specificity in the domain of coordination. Although it seems to be hard to get rid of construction-specific constraints (i.e., “conjuncts” vs. “adverbial clauses”), the approach presented here (which builds on ideas by Munn (1993) and Hornstein and Nunes (2002)) has one advantage in this domain, since no construction-specific principle is assumed to apply in the narrow syntax. The ban on extraction from non-first conjuncts (i.e., the CED) that applies in the syntactic component is based solely on the way different elements are concatenated in the syntax.

The CSC should now then be understood as a semantic constraint, as Munn (1993) proposed. Even though this constraint is still construction-specific, it is reducible to properties of specific lexical items, according to Munn. The claim would be that conjunctions take scope over the conjuncts at LF, requiring semantic identity between the conjuncts. Here I have argued that this identity requirement



should be relaxed, in the sense that the gaps inside the conjuncts need not be parallel with respect to their grammatical function. What is required, if (3.28) is correct is that conjuncts be parallel with respect to presence or absence of a gap (or a variable co-indexed with a gap).

On the other hand, it is also true that trying to push construction-specificity away from the narrow syntax has led to the discovery of a previously unnoted parallelism requirement that applies at the PF interface. Hopefully, in the future we will be able to investigate further why this restriction exists, preferably with the prospect of getting rid of construction specificity altogether in this domain.

A final issue that I would like to discuss briefly is the relationship between the parallelism requirement and ellipsis. It has been claimed in the literature that a violation of the CSC can be salvaged by ellipsis (Chung, Ladusaw, and McCloskey 1995, Merchant 2001, 2008, Fox & Lasnik 2003). Consider the example in (3.31), which involves sluicing:

(3.31) (from Merchant 2008)

- a. Bob ate dinner and saw a movie that night, but he didn't say [which]
- b. \*Bob ate dinner and saw a movie that night, but he didn't say [[which]<sub>i</sub>,  
[Bob ate dinner and saw e<sub>i</sub> that night]]

Abstracting away now from whether examples like (3.31a) involve movement of the wh-phrase and PF-deletion (Merchant 2001) or base generation of the wh-

phrase and LF-copying (Chung, Ladusaw, and McCloskey 1995), the LF representation for both (3.31a) and (3.31b) is the same. This causes a potential problem for the requirement that both conjuncts (*Bob ate dinner* and *saw e*) need to have a gap once extraction applies (see (3.28i) above), since there is no gap in the first conjunct in (3.30b).

A solution to this problem (compatible with both approaches to ellipsis, as far as I can see) would be to assume that what is elided/not pronounced in sluicing cases like these is not the whole two conjunct structure, but only the second conjunct, as shown in (3.32):

(3.32)

[Bob ate dinner] and [e saw a movie], but he didn't say [[which]<sub>i</sub> [e saw e<sub>i</sub>]]

This solution relies on the assumption that in cases of coordination in which the subject is shared by the two conjuncts, what we have is a null subject in the second conjunct ([e saw a movie]), which is one configuration in which null subjects are allowed in English. This is supported by the fact that there are non-null subject languages that do not allow a null subject in this position (see Holmberg 2005 and references therein). The reason why the overt version of (3.32) would be ungrammatical then is that the null subject is not licensed in the overt syntax in that position. This solution seems to be more compatible with the LF-copying approach

than with the PF-deletion approach, although I am not going to discuss this issue further here.

Assuming this solution to be on the right track, we can still maintain the parallelism requirement as formalized in section 3.4.

## CHAPTER 4

### On the unification of ATB and PG constructions

Another important consequence of my version of the adjunction analysis of coordination that I entertain here is that the syntactic configurations in which ATB and Parasitic gaps occur are identical. This means that a unified analysis of the two kinds of constructions not only is desirable but unavoidable. There have been two main attempts in the literature to unify the analysis of these constructions: the Null Operator analysis (Chomsky 1986, Munn 1993, Nissenbaum 2000) and the Sideward Movement analysis (Nunes 2001, 2004, Hornstein 2001). For this reason, I am going to focus on these two approaches in this chapter, with occasional reference to other approaches to ATB extraction.

The basic idea behind the Null Operator Analysis is that a base-generated empty operator moves to the edge of the adjunct and is co-indexed with another syntactic object in the main clause (see Chomsky 1986, Nissenbaum 2000 for PG constructions, extended to ATB constructions by Munn 1993). In section 4.1 I present some empirical problems and incompatibilities of the null operator analysis with Minimalism.

In section 4.2 I will review the Sideward Movement analysis and discuss its implications for syntactic theory in terms of a number of assumptions that need to be made in order for the approach to be tenable. The main idea behind the Sideward Movement analysis is that an element inside the adjunct can be copied and merged in the main clause *before* the adjunct is attached. This is the way this analysis makes sense of the most intriguing feature of ATB and PG constructions, namely, how one overt element can be interpreted in two different thematic positions, giving the impression of being extracted from two syntactic positions and moved to one single syntactic position.

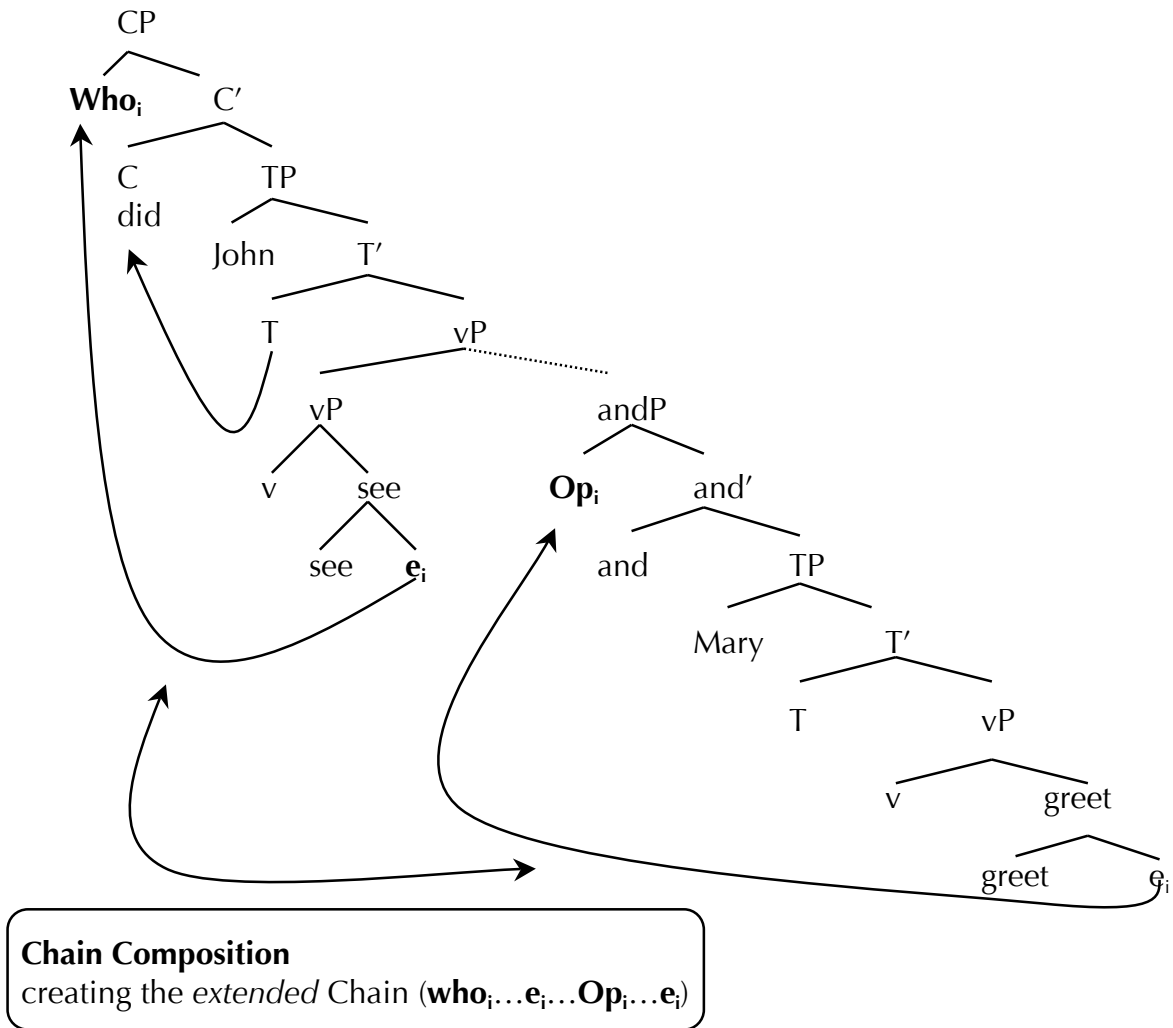
Section 4.3 deals with Crossover phenomena, which I will in turn use to provide further evidence for a Sideward Movement analysis of ATB and PG constructions. I first attempt to formulate a more Minimalist version of the principle of Weak Crossover and then show how Sideward Movement avoids Weak Crossover effects in the adjoined clauses (non-first conjuncts and adverbial clauses), giving rise to what Lasnik and Stowell (1991) call *Weakest Crossover*. I also show how Strong Crossover, in contrast, cannot be avoided, making the right predictions.

Finally, in section 4.4 I extend the Sideward Movement analysis to other constructions that have been claimed to involve movement of a Null Operator, like relative clauses and Tough-Movement. I also show how the Crossover properties of these constructions can be accounted for under the Sideward Movement approach, while they are problematic for a Null Operator analysis.

#### 4.1. The Null Operator analysis

Under the Null Operator analysis (Chomsky 1982, 1986, Munn 1993, Nissenbaum 2000), the gap in the adjoined clause (adverbial clause in PG constructions and non-first conjunct in ATB constructions) is not a trace of the element that appears in the matrix clause, but actually a trace of another empty category, a Null Operator. This operator (a null counterpart of a wh-operator) moves to the edge of the adjunct and the A'-chain that results from that movement is then linked to the A'-chain in the matrix clause (the one that results from movement of the overt wh-phrase) a mechanism called Chain Composition. The structure of a sentence like *who did John see and Mary greet* under this analysis would be the one in (4.1) below. I am attaching the second conjunct to vP, following the modified adjunction approach I proposed in chapter 1. I am also ignoring subject for now raising to Spec-TP and the issue of how the Tense element inside the adjunct is realized. I return to this issue briefly in section 4.2.

(4.1)



A number of problems arise when trying to adapt this approach to Minimalist theory assumptions. First of all, the very use of empty categories like *Op* is called into question in Minimalism, and thus the rich empty category system of Government and Binding has been abandoned. A similar argument against the empty category system has been used in a different domain, namely, minimalist research on null subjects (Alexiadou and Anagnostopoulou 1998, Holmberg

2005, Fernández-Salgueiro, to appear). In Alexiadou and Anagnostopoulou's (1998:531) words, for example:

"In the Minimalist framework, the theory of empty categories as we know it from GB has been largely abandoned. Specifically, traces of movement are viewed as copies of moved elements, and for PRO we need to appeal to null Case."

Interestingly, other kinds of empty categories are indeed assumed in Minimalism, like *v*, *T*, and *C*, which are often dubbed "core functional categories" (Chomsky 2001). The main difference that I see between these categories and the Null Operator is that the former are part of the main structure of all clauses and actually are taken to drive many core syntactic operations. The categories *v* and *T* are understood to be involved in A-Movement and Case operations and are also argued to have semantic import, with *T* carrying Tense information and *v* being involved in aspectual information (see e.g. Ritter and Rosen, 1993), while *C* is the locus of Wh-Movement and determines sentence type (matrix vs. embedded), among other functions. These categories are present in *any* derivation, independently of which kind of "construction" we are dealing with (with the exception of *v*, which might not be present in unaccusatives and passives). Null Operators, on the other hand, are understood to appear only in a restricted class of constructions: relative clauses, ATB and PG constructions, and Tough-Movement constructions. The fact that a covert category is invoked in only these constructions has indeed been questioned in Minimalism and thus different mechanisms have



been proposed to account for their properties without assuming a Null Operator, like Sideward Movement, Parallel Merge (Citko 2005), PF-Deletion (see e.g. An 2007 and references therein). I briefly discuss some aspects of both the Parallel Merge and the PF-Deletion analyses in section 4.4 below.

A second problem concerns the kind of movement operation that a Null Operator undergoes. Even though Null Operators are understood to be a phonologically empty version of an overt operator like a *wh*-phrase, only Null Operators can undergo movement to the specifier position of the *andP*, as shown in (4.1), or movement to the edge of the adverbial clause in the case of PG constructions. The fact that a certain kind of syntactic operation is available for a syntactic object because of its phonological matrix is also problematic. As far as we know, narrow syntactic operations are not concerned with the phonological make-up of the objects it operates on. Such a problem does not arise under the Sideward Movement approach, since gaps in non-first conjuncts and adverbial clauses will be analyzed as copies left by overt movement.

A third problem that arises is the unclarity of the mechanism of co-indexation, another aspect of the theory that has been dispensed with. Indexation is considered a “coding trick”; an *ad-hoc* procedure that gets the facts right without understanding the nature of the syntactic relation(s) that (co-)indexing represents. More specifically, indices clearly violate an important tenet of Minimalism, which is Inclusiveness (Chomsky 1995):

(4.2) Inclusiveness Condition (Chomsky 1995:228)

Any structure formed by the computation is constituted of elements already present in the lexical items selected [...]; no new objects are added in the course of computation, apart from rearrangements of lexical properties (in particular, no indices, bar levels in the sense of X-bar theory, etc.)

A related problem concerns the notion of Chain Composition, which is needed to account for why the two gaps have to be co-indexed. Again, it is not clear why a syntactic operation would apply only when a specific kind of empty category is involved (the Null Operator). Chomsky (1986) argues that Chain Composition is required for Null Operators because they are semantically deficient. This accounts for why a Null Operator is not compatible with an expletive, as shown in (4.3):

(4.3)

- a. John is easy [Op<sub>i</sub> PRO to see e<sub>i</sub>]
- b. It is easy [PRO to see John]
- c. \*It is easy [Op<sub>i</sub> PRO to see e<sub>i</sub>]

As will be seen in the next chapter, under the Sideward Movement approach the expletive *it* would have to originate as the object of *see*, a position where expletives cannot be inserted. At any rate, one of the tenets of the Minimalist

Program is to try to derive the properties of linguistic expressions from more simple and general operations, and Chain Composition, though empirically motivated, seems to be at odds with this view.

A fourth main problem with this approach concerns evidence found in so-called wh-in situ languages. The Null Operator approach makes the prediction that ATB and PG constructions should be possible in languages like Chinese, which display covert wh-movement. While Chinese does indeed display ATB and Parasitic Gaps, these are only possible when the wh-phrase undergoes overt movement (by means of topicalization, for example), as shown in (4.4):

(4.4) (see Lin 2005)

Shei <sub>i</sub>	Laowang	[zai	huijian	e <sub>i</sub>	zhiqian]	jiu	kaichu-le	e <sub>i</sub> ?
Who	Laowang	at	meet	before	already		fire-PERF	

'Who is the person x such that Laowang fired x before meeting x?'

When the wh-phrase stays in situ, the parasitic gap in the adjunct cannot be licensed, as shown in (4.5):

(4.5)

*Laowang	[zai	huijian	e <sub>i</sub>	zhiqian]	jiu	kaichu-le	shei?
Laowang	at	meet	before	already		fire-PERF	who

The ungrammaticality of (4.5) is unexpected under the Null Operator approach. There is good evidence (Huang 1982) that A'-Movement to Spec-CP applies in Chinese, although covertly. If this is true, it is not clear why Operator movement and/or Chain Composition cannot apply in (4.5).

A final problem for this approach, which I discuss in more detail in sections 4.4 and 4.5, is that A'-movement of the Null Operator (unlike A'-movement of an overt operator) does not give rise to Weak Crossover effects. I take this to be the main piece of evidence that there is no such thing as movement of any category to Spec-CP in ATB and PG constructions, exactly as is predicted by the Sideward Movement approach I will adopt here.

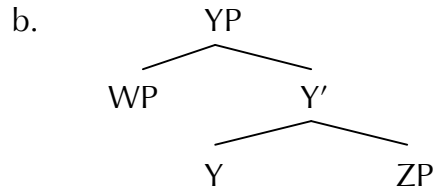
#### **4.2. Nunes' Sideward Movement analysis of ATB and PG constructions**

As mentioned in the previous section, under standard Minimalist assumptions, movement of a syntactic object is always to a c-commanding position, a condition that resembles the so-called Proper Binding Condition on traces in Government and Binding Theory (Fiengo 1977). Bobaljik and Brown (1997) and Nunes (1995, 2001, 2004) and others have argued that under the copy theory of movement nothing in Minimalism prevents movement to a non-c-commanding position, as long as a Chain can be formed between the moving element and the copy in its original position at a later point in the derivation.

To illustrate the idea behind Sideward Movement (Nunes 2001, 2004 version) schematically, consider (4.6):

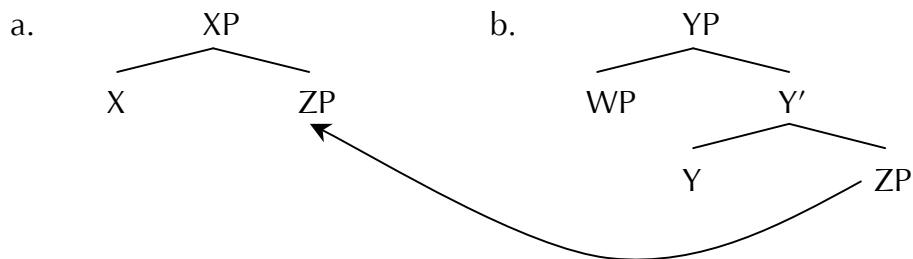
(4.6)

a. X



According to standard Minimalism, an element inside the syntactic object in (4.5b) can never be related to the object in (4.6a), movement being one such banned relation (see below and also e.g. Kitahara (2007)). If, say, ZP were to merge with X, we would get (4.7):

(4.7)

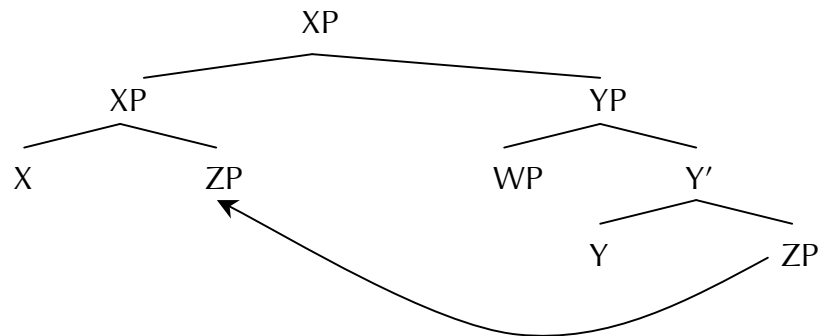


The reason this movement is banned in standard Minimalist theory is that the landing site does not c-command the departure site and therefore violates Fiengo's Proper Binding Condition. This idea has been recast as a Minimal Search condition in recent work (see Chomsky 2001, Kitahara 2007), which I discuss towards the end of this section.

What we have in (4.7) is an instance of Sideward Movement. The key assumption that needs to be made in order to allow this as a possible syntactic

operation is that this movement is permitted as long as a well-formed chain can eventually be created between all copies of a syntactic object. Suppose that the derivation in (4.7) continues as in (4.8), with merger of the two syntactic objects XP and YP:

(4.8)



It is crucial to note that no approach to Sideward Movement (as far as I know) allows (4.8) as the final stage of the derivation. Sideward Movement is always supplemented with a condition that all copies form a Chain. The relevant conditions for our discussion are reproduced in (4.9) (see Nunes 2004:91 for a complete formulation):

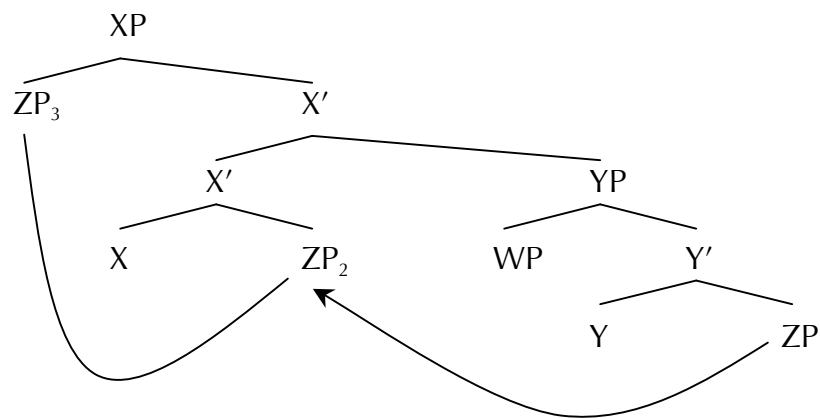
(4.9) Conditions on Form Chain

Two constituents  $\alpha$  and  $\beta$  can form the non-trivial chain  $CH = (\alpha, \beta)$  if

- a.  $\alpha$  is non-distinct from  $\beta$ ,
- and*
- b.  $\alpha$  c-commands  $\beta$ .

It is also crucial to notice that this is not specific to syntactic objects undergoing Sideward Movement only, but all syntactic objects undergoing movement. Thus, there is no special stipulation for Sideward Movement in this respect. Suppose now that the derivation continues as in (4.10):

(4.10)



A crucial difference between (4.8) and (4.10) is that in the latter all copies of ZP can become part of a chain. Assuming the c-command condition, both ZP<sub>1</sub> and ZP<sub>2</sub> can form a (separate) chain with ZP<sub>3</sub>. From a theoretical point of view, an interesting aspect of this approach, then, is that movement and chain formation are not isomorphic.

Bobaljik and Brown (1997) argue for a Sideward Movement approach (which they call an “interarboreal operation”) to elegantly account for the properties of head movement. They claim that a chain can be formed between the

multiple copies of the moved head, since head adjunction will allow for the adjoined head to c-command into the departure site position.

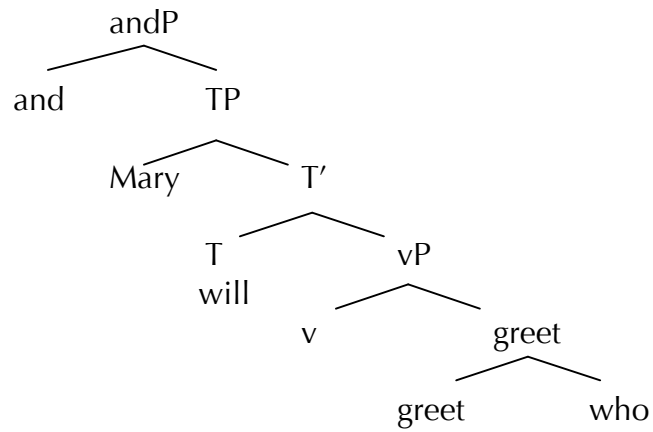
Nunes (2001, 2004) argues that Sideward Movement accounts for the properties of ATB and PG constructions, as well as the different restrictions that apply in each. I now briefly present Nunes' Sideward Movement analysis of ATB constructions, adapted here to my version of Munn's analysis of coordination (Nunes assumes Kayne's X' analysis). For more details about how this approach accounts for the properties of PG constructions as well as the differences between ATB and PG constructions, see section 4.3, and also Nissenbaum (2000), Hornstein and Nunes (2002), and Nunes (2004).

The Sideward Movement analysis of a sentence like *who will John see and Mary greet* would proceed as follows:

First, the andP *and Mary will greet who* is built in a derivational workspace (ignoring subject movement to Spec-TP, for simplicity):

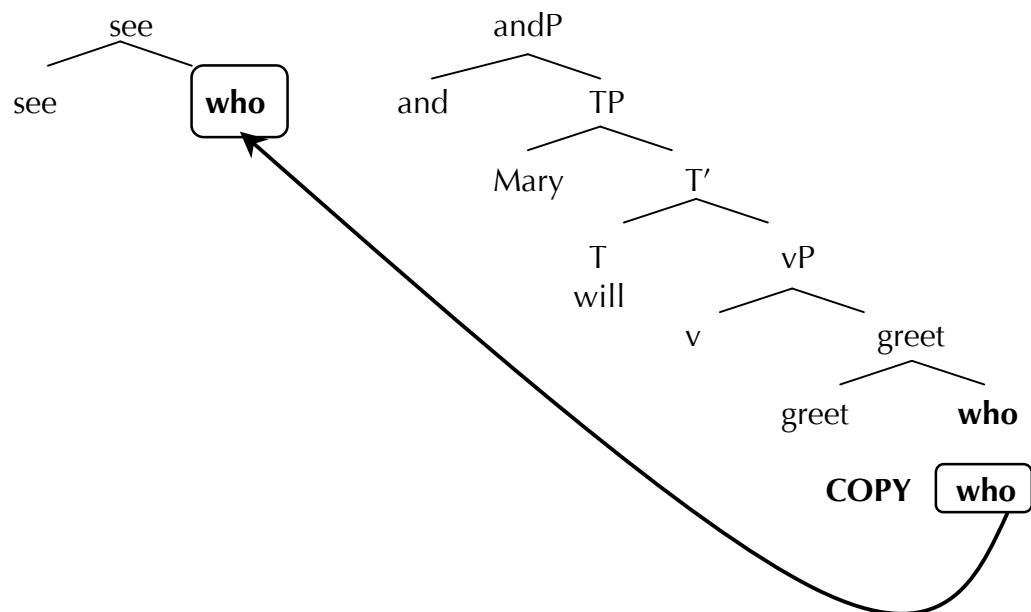


(4.11)



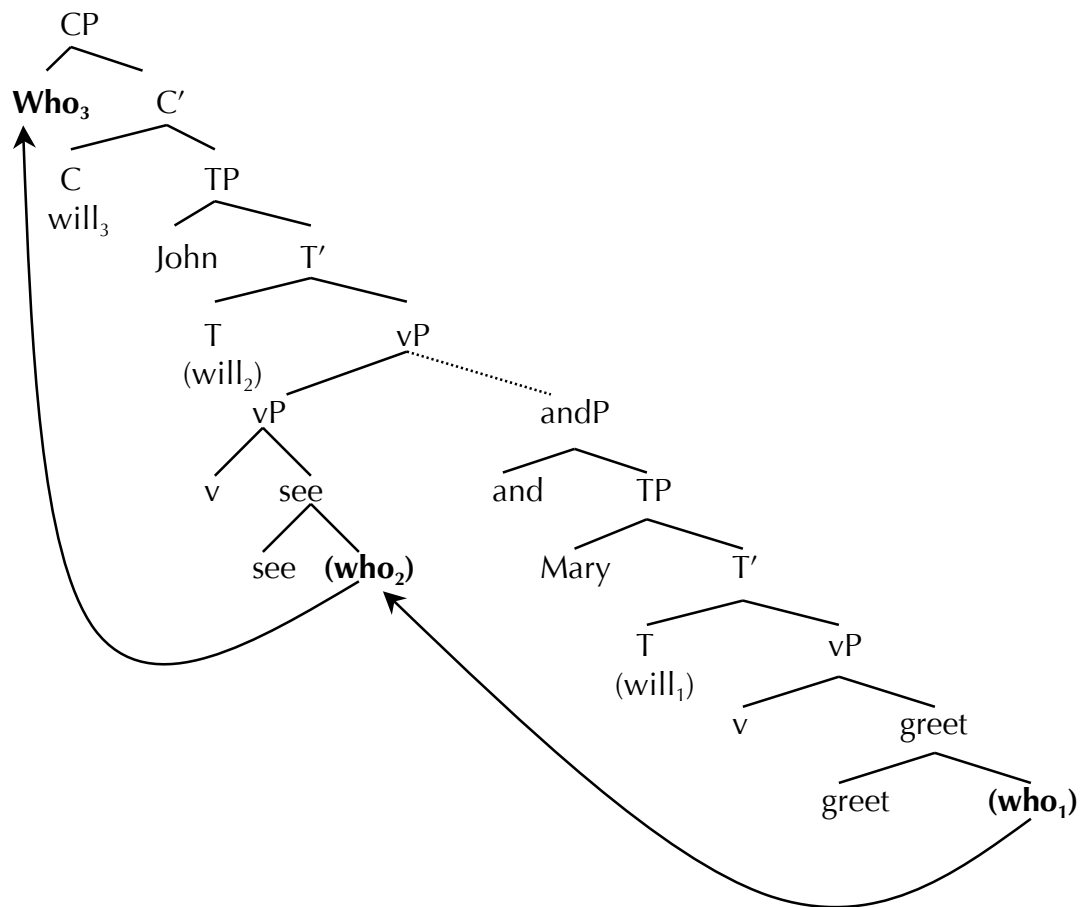
At this point, the verb *see* is selected from the numeration and a copy of *who* is made and merged with *see* in a different workspace in order to satisfy *see*'s selectional ( $\theta$ -)feature, as shown below:

(4.12)



This Sideward Movement of *who* to another syntactic object before the andP is adjoined is what allows *who* to move without violating the Condition on Extraction Domains. When *who* is copied and merged, the andP is *not yet* an island for movement: the andP only becomes an island when it adjoins to vP. Further operations yield the syntactic object in (4.13), after the andP adjoins to vP:

(4.13)

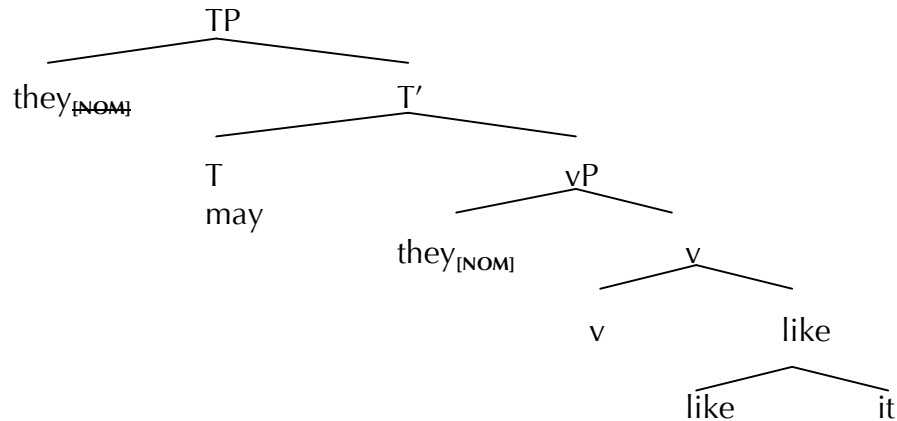


Once this object is created and no more Merge operations apply, the copy of *who* in Spec-CP ( $who_3$ ) can now form two chains, one with  $who_2$  ( $who_3, who_2$ ) and another one with  $who_1$  ( $who_3, who_1$ ).

There is an important reason why all copies have to form a Chain according to Nunes. The main goal of Nunes' approach to chain linearization is to provide an account of why it is generally the top-most copy of a syntactic object that is pronounced (although he also provides an account of cases in which the lower copy is pronounced) by arguing that the other copies undergo a deletion operation. Copies need to be part of a chain in order for this deletion operation to apply. If they are not part of a chain, their unchecked features will cause the derivation to crash at the interface. This is exactly what happens in the Chinese example in (4.5) above; a *wh*-in situ element does not c-command the copy inside the adjunct, therefore the latter cannot form part of any chain. In the grammatical case in (4.4), however, fronting of the *wh*-element allows for the relevant chains to be formed just like in (4.13) above.

Since the theory of Sideward Movement crucially depends on Nunes' general approach to the linearization of Chains, let's discuss briefly how the approach works. In order to understand why Chains need to be formed, consider the structure in (4.14):

(4.14)



First of all, it is important to note that Nunes assumes a Checking-based (not an Agree-based) approach to the deletion of uninterpretable features like Case features. For other approaches that claim Case-checking to occur in a Spec-Head configuration (even assuming Agree) see Rezac (2003), Müller (2004), Fernández-Salgueiro (2005), and Epstein and Seely (2006).

Once this movement applies, then, the chain  $CH=(\text{they}_{[NOM]}, \text{they}_{[NOM]})$  is formed. As can be seen, only the first link of the chain (the upper copy of *they*) has had its features checked, *contra* Chomsky 1995, who claims that all links in a chain have to be identical.

Nunes' question is: why is it only the upper link that is generally pronounced? In the case at hand, why is the linearization output of (4.14) *they may like it* and not *\*they may they like it*, *\*may they like it* or *\*may like it*? Nunes' approach is going to provide a principled answer to this question based on economy considerations and (un)interpretability of features at the PF interface. He

proposes that the minimum amount of deletion operations should be employed provided the result satisfies Full Interpretation at PF (i.e., it feeds *only* phonological features to the PF interface).

The first ungrammatical option, *\*they may they like it*, is ruled out independently by Kayne's Linear Correspondence Axiom, the relevant part of which is reproduced in (4.15):

(4.15) Kayne's (1994) Linear Correspondence Axiom

If  $\alpha$  c-commands  $\beta$  then  $\alpha$  precedes  $\beta$ .

According to Kayne, the linearization algorithm is unable to decide whether *they* precedes or follows *may*, since there is a c-command relation contradiction: *they* c-commands *may* but *they* is also c-commanded by *may* at the same time. In Kayne's terms, there are two sets of ordered pairs that involve a contradiction:  $\langle \text{they}, \text{may} \rangle$  and  $\langle \text{may}, \text{they} \rangle$ . Nunes argues that since a complete linear order cannot be assigned, one of the copies of *they* has to be deleted and hence ignored by the PF interface.

This leaves us with three logical possibilities: if we choose *\*may they like it*, we need two deletion operations, one that deletes the upper copy of *they* and another one which deletes the nominative Case feature of the lower copy of *they* (which has not been deleted in the course of the derivation because it has not entered into a checking relation). If *\*may like it* is chosen, we also need two

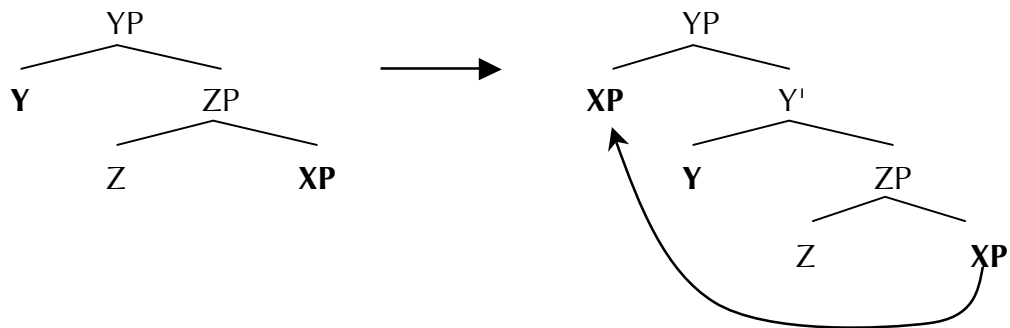
operations, one which deletes the upper copy of *they* and another one which deletes the lower copy of *they*, together with its nominative Case feature. If we choose *they may like it*, though, only one deletion operation is necessary to make the structure linearizable, the one which deletes the lower copy of *they*. In sum, *they may like it* is the option with fewer deletion operations that satisfy Full Interpretation at PF, i.e., it yields a PF representation with no uninterpretable features.

Another interesting advantage of the Sideward Movement approach is that it can also account for the fact that matrix ATB extraction obligatorily yields the same Tense interpretation of both conjuncts (Hornstein and Nunes 2002). This is a property of ATB extraction that is not accounted for under the Null Operator approach. (Neither approach, however, can account for parametric variation in this domain; recall the Spanish data presented in the examples in (3.18) above, section 3.4). The claim is that *T* also undergoes Sideward Movement from the *andP* to the main clause, and eventually undergoes T-to-C movement in the main clause, as already illustrated in the tree in (4.13) above (see *will*<sub>3</sub>, *will*<sub>2</sub>, *will*<sub>1</sub>).

Despite its advantages, Sideward Movement is not considered part of “standard Minimalism.” One of the main criticisms that it faces is that it does not obey the Minimal Search condition (see Chomsky 2001). To illustrate this condition, consider (4.16), from Kitahara (2007) (modified):

(4.16)

**Y** finds an embedded category **XP** by probing into its smallest searchable domain (meaning that embedded categories are accessible only if the probe category finds them):



Since Sideward Movement involves movement to a *different* syntactic sub-tree, this condition is not met.

However, I think that part of this controversy comes from the label Sideward “Movement” itself. If we look at this operation from the point of view of the moving element (for example *who* in (4.13)), it does look like movement because a copy is created and remerged in some other position in the structure. From the point of view of the element that drives the movement (i.e., the target/probe see in (4.13) above), this is not so clear. The only difference between Sideward Movement and Merge in this respect is whether the object that moves comes from the Numeration or from a tree already assembled. If the former is the case, Select+Merge (“pure” or “external” Merge in Chomsky 2001) applies. If the latter, it is Copy+Merge

(“internal” Merge) that applies. As can be seen, computationally speaking, the difference between Sideward Movement and Merge reduces to “select” vs. “copy.”

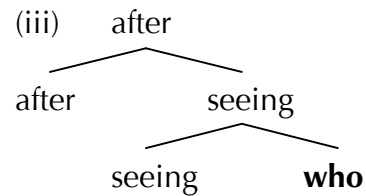
It is important in my opinion to note that the distinction relies on the existence of Numerations, and the *stipulation* that an element merged from the Numeration is “selected”, while a moved element is “copied”. Collins (1993) for example, argues against the existence of Numerations altogether and proposes instead that every Merge operation is preceded by a copying operation, either copied from the lexicon, or from the syntactic object already built (see also Frampton and Gutmann 2002). Even maintaining the notion of Numeration (which arguably reduces computational complexity by restricting the number of lexical items that can be inserted in a derivation) one could argue that lexical items are actually copied from the Numeration. If this is correct, Sideward Movement should maybe be understood as an instance of “external” Merge rather than “internal” Merge (i.e., Merge rather than Move in early Minimalism (e.g. Chomsky 1995) terminology). To illustrate this point more clearly, consider the derivation in (4.17), consisting of (i) a numeration, (ii) a head with a selectional feature (say, a lexical verb demanding that an argument be merged with it), and (iii) an already built syntactic tree:



(4.17)

(i) Num = {..., **John**, ...}

(ii) **greet**



The verb *greet* in (4.15ii) needs to satisfy its internal  $\theta$ -role. In standard Minimalism, only an element inside the Numeration (like *John* in (4.17i) can satisfy it. The reason is that a  $\theta$ -role cannot be satisfied with an element that is “copied”, only with one that is “selected.” Again, I think this distinction is stipulative. If we remove such a stipulation and assume that elements in a numeration are also “copied,” there is no reason to believe that copying *John* from the Numeration is really any computationally more efficient than copying *who* from the other derivational workspace, since neither *John* nor *who* are in the Minimal Search domain of *greet*. If this is correct, it is not clear that Sideward Movement violates Minimal Search, since it is not clear how it is more costly than external Merge once the stipulative distinction between “select” and “copy” is abandoned. Sideward Movement would still be different from “external” Merge in that the source of the syntactic object is different, but it seems that neither operation should be subject to a Minimal Search condition; both seem to involve what Kitahara calls “zero” search.

Kitahara (2007) also argues that Sideward Movement is incompatible with other current assumptions, namely, Phase-based cyclicity and Phase-based application of Transfer. These assumptions rely on the notion of Phase, which was first introduced in Chomsky (2001). The main idea behind this approach is that there are two points in derivations (whenever a *v* or a *C* project) at which the object been built by the syntax is sent to PF (i.e., application of Spell-Out) and LF for interpretation. Any element inside those spelled-out domains is unavailable for further narrow syntax computation.

Sideward Movement was proposed in early Minimalism, before Chomsky's (2001) *Derivation by Phase*. As far as I can see, there is no way to reconcile Sideward Movement with the Derivation by Phase approach to syntax, given this completely different understanding of how derivations proceed.

Even though Sideward Movement is not considered part of "standard" Minimalism, an advantage of the approach is that it can derive the fact that a single element can be interpreted as an argument of two different verbs under certain restricted conditions, thus accounting for this special property of ATB and PG constructions, without relying on the postulation of more empty categories than are needed to account for other (more familiar) instances of movement. Thus, the Sideward Movement analysis faces none of the problems that I pointed out arise for the Null Operator analysis in section 4.1. In sections 4.4 and 4.5 I argue that Sideward Movement can also account for the Crossover properties of these and other constructions. Before analyzing Crossover phenomena, in section 4.3

examine some interesting main clause vs. adverbial clause asymmetries that arise in PG constructions.

### 4.3. Some asymmetries in PG constructions

Most of the discussion on Sideward Movement so far has been based on cases of ATB extraction. In this section I confront some asymmetries that have been claimed to arise in PG constructions but be absent in ATB Extraction. Nissenbaum 2000 uses these for evidence that ATB and PG constructions cannot be derived by the same mechanisms. Let's begin with a Condition A reconstruction asymmetry:

(4.18) Condition A (PG)

- a. Which picture of himself did John sell **e** [before Mary had a chance to look at **e**]?
- b. \*Which picture of himself did Mary sell **e** [before John had a chance to look at **e**]?

(4.19) Condition A (ATB)

- a. \*Which picture of himself did John sell **e** [and Mary buy **e**]?
- b. \*Which picture of himself did Mary sell **e** [and John buy **e**]?

Nissenbaum claims that reconstruction of the wh-phrase containing the anaphor is asymmetric in the case of PG constructions but not in the case of ATB extraction.

However, these data have been disputed. Munn (1992, 2000), Citko (2005) report a contrast in (4.19) as well, mirroring the asymmetry found in (4.18).

The next asymmetry claimed by Nissenbaum concerns Condition C:

(4.20) Condition C (PG)

- a. Which picture of John<sub>i</sub> did Mary buy **e** [without letting him<sub>i</sub> look at **e**]?
- b. \*Which picture of John<sub>i</sub> did he<sub>i</sub> buy **e** [without letting Mary look at **e**]?

(4.21) Condition C (ATB)

- a. \*Which picture of John<sub>i</sub> did Mary buy **e** [and not let him<sub>i</sub> look at **e**]?
- b. \*Which picture of John<sub>i</sub> did he<sub>i</sub> buy **e** [and not let Mary look at **e**]?

Again, these data are disputed. Citko (2005), for example, provides the following judgments:

(4.22)

- a. Which picture of John<sub>i</sub> did Mary like **e** and he<sub>i</sub> dislike **e**?
- b. \*Which picture of John<sub>i</sub> did he<sub>i</sub> like **e** and Mary dislike **e**?

Since most speakers do not find a clear contrast between ATB and PG constructions, I will continue to assume that they are derived by the same

mechanism, leaving the implications of possible parametric variation for further research.

However, the Sideward Movement analysis still has to account for the asymmetries between conjuncts in the case of ATB extraction and between main and adverbial clauses in the case of PG constructions.

As noted by Nunes (2004), one of the advantages of the Copy theory of movement is that it eliminates the operation of LF-Reconstruction (see also Chomsky 1995 for the original arguments). Rather, LF can pick one of the copies of the moved element for computation of Binding Theory Conditions. In (4.23) for example, *himself* co-refers with *John* if LF chooses the lower copy of *which pictures of himself* (a), or with *Bill* if LF chooses the intermediate copy of *which pictures of himself* (b):

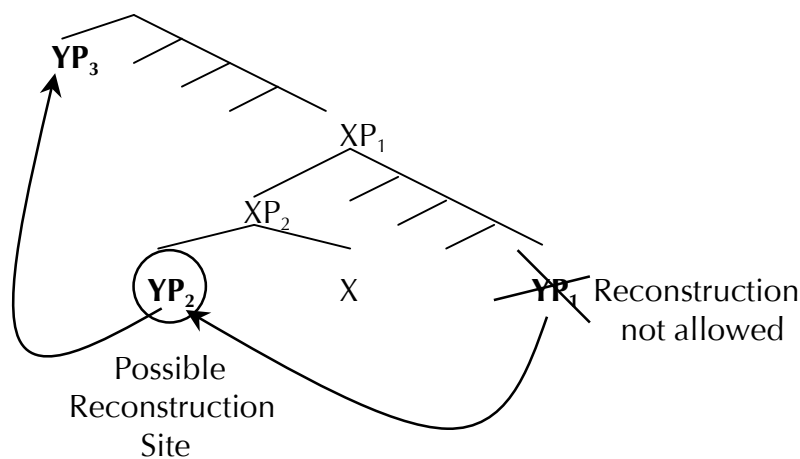
(4.23)

- a. **[Which pictures of himself<sub>i</sub>]** does Bill<sub>j</sub> say [which pictures of himself<sub>i</sub>]  
**John<sub>i</sub>** likes [which pictures of himself<sub>i</sub>]?
- b. **[Which pictures of himself<sub>i</sub>]** does **Bill<sub>i</sub>** say [which pictures of himself<sub>i</sub>]  
John<sub>j</sub> likes [which pictures of himself<sub>i</sub>]?

However, when Sideward Movement applies, a different pattern emerges, in that LF seems to be able to choose the copy inside the first conjunct/main clause *only* (recall (4.18) and (4.20) above).

Descriptively speaking, it seems that LF only considers for purposes of Binding the copy that underwent movement to a c-commanding position but not the one(s) that underwent Sideward Movement, as illustrated schematically in (4.24):

(4.24)



Why should  $YP_2$  be privileged in this respect? Even though I cannot give a principled answer to this question, it is also true that the Sideward Movement approach can indeed distinguish between the two copies ( $YP_1$  and  $YP_2$ ). As pointed out above, under Nunes' approach chains and Movement are not isomorphic. In (4.24) for example, the chain formed between  $YP_3$  and  $YP_1$  is not isomorphic to movement, that is, the positions occupied by these two links of the chain were not related by a movement operation. Conversely, the chain formed between  $YP_3$  and  $YP_2$  is isomorphic to the movement operation. We could push this idea further and assume that the chains  $CH = (YP_3, YP_1)$  and  $CH = (YP_3, YP_2)$  have different

properties in the LF component with respect to the computation of binding theory principles. Assuming Sideward Movement then, we are compelled to formulate the principle in (4.25):

(4.25) Copy Visibility at LF

Only copies that are part of a chain that is isomorphic to a movement operation are visible at LF for computation of binding theory principles

A final asymmetry that arises in PG constructions is a CED-related asymmetry. The PG cannot be separated from its licensor by more than one island. This is illustrated in (4.26):

(4.26) (see Hornstein and Nunes 2002:44 and references therein)

- a. **[Which documents]<sub>i</sub>** did John mention **e<sub>i</sub>** [without checking **e<sub>i</sub>**]?
- b. \***[Which documents]<sub>i</sub>** did John mention **e<sub>i</sub>** [after signing papers  
[without checking **e<sub>i</sub>**]?

Chomsky (1986) uses this kind of contrast as evidence that Parasitic Gaps were generated by movement (of a Null Operator), and not by just licensing of a null category in an A-position. It is not clear, however, why the Null Operator cannot move to the edge of the adjunct [*after signing papers...*] through successive cyclic movement, that is, by first moving to the edge of [*without checking e*]. If it is true

that this Operator is unable to undergo successive cyclic A'-Movement for some reason, this seems to be another problem for the Null Operator approach.

This is explained by Hornstein and Nunes in terms of timing of adjunction. At the point at which the verb *mention* could trigger Sideward Movement of *which documents* the adjunct [*without checking which documents*] has already been attached to [*after signing papers*], making *which documents* unavailable for extraction (see also Nunes and Uriagereka 2000 for a more detailed version of this approach).

However, the fact that adjuncts can be attached acyclically in the derivation is a problem for this explanation. More specifically, if the adjunction of [*without checking which documents*] can take place after Sideward Movement of [*which documents*], then (4.26b) would be allowed. It seems then that we need to supplement Sideward Movement with the condition that it can only apply between two adjacent workspaces.

Let's assume first that the syntax builds only one derivational workspace at a time, and that the order in which they are built is the opposite order in which they are interpreted by PF (see Fernández-Salgueiro 2007 and also section 2.2 above). With these assumptions in mind we can formulate (4.27):

#### (4.27) Adjacent Workspace Condition

Sideward Movement is only possible from a given derivational workspace to the next derivational workspace built by the syntax.



This condition also relates to Hornstein and Nunes' (2002) claim that Sideward Movement always proceeds from more embedded to less embedded domains, since adjuncts are generally built in the syntax before the objects they attach to.

A related problem caused by the fact that the adjunct can be inserted acyclically is that in principle it permits (4.28):

(4.28)

\*[**Which documents**] did you file that book [without reading e]

This ungrammatical example could be derived by Sideward Movement of *which documents* before the adjunct is attached, circumventing a CED violation. Recall, however, that Sideward Movement is licensed by a  $\theta$ -feature, whereas (4.28) is an instance of a Q feature on C driving wh-movement. Since this feature can only be satisfied through Internal Merge, (4.28) can never be derived. If adjunction takes place acyclically, the Q feature on C will not find the wh-element in its search (c-command) domain, so Sideward Movement will not be allowed. If adjunction takes place cyclically then the wh-element, though visible to the Q feature, is inside the island, so it is unavailable for extraction.

#### 4.4. Revisiting Crossover phenomena: Weak and Weakest (and Strong)

In this section I provide an overview of the different Crossover effects that we find in ATB and PG constructions. I am going to argue that (the absence of) Weak Crossover (WCO) effects inside the adjoined elements in these constructions favor a Sideward Movement analysis over the Null Operator analysis, assuming that WCO is a constraint on the syntactic operation of movement (*contra* Koopman and Sportiche's (1982) representational Bijection Principle).

Let's first try to formalize a minimalist version of the principle of WCO. Consider the well-known contrast in (4.29):

(4.29)

- a. Who<sub>i</sub> e<sub>i</sub> invited his<sub>i</sub> friend to the party?
- b. \*Who<sub>i</sub> did his<sub>i</sub> friend invite e<sub>i</sub> to the party?

Ross' (1967) and Postal's (1971) early accounts of these cases claimed that the ungrammaticality of (4.29b) is due to the fact that movement of *who* literally crosses (i.e., moves past) a co-indexed element in the linear order, while that is not the case in (4.29a).

In previous work (Fernández-Salgueiro 2008) I followed the basic idea of this approach, and attempted to state a WCO principle in a way that relates a syntactic operation of movement with the resulting linear order. I will first review this approach and will then point out some problems for it.

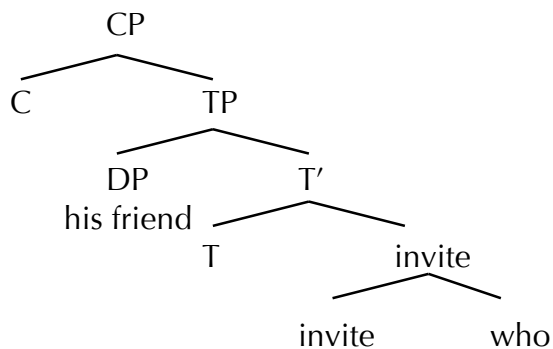
My earlier proposal argues for a Minimalist approach to “crossing” based on PF and LF interface properties, by assuming a strong derivational approach to syntactic relations under which interpretation at the interfaces proceeds derivationally (Epstein et al. 1998, Epstein and Seely 2006; see also Fox and Pesetsky’s (2004) cyclic approach to Linearization), as stated in (4.30):

(4.30)

Each operation of Merge provides a set of instructions to the PF and LF components (including linear order information).

If this is true, at the point in the derivation in (4.31) (i.e., after movement of the subject to Spec-TP, ignored here),

(4.31)



the following linear order is computed by PF:

(4.32)

PF = <C, his, friend, T, invite, who>

Also, the LF component assigns a semantic interpretation to (4.31). Since *his* does not c-command *who*, *his* and *who* can co-refer (cf. his<sub>i</sub> best friend likes John<sub>i</sub>).

If the co-referent semantic interpretation is assigned and *who* were then to move to Spec-CP, the following linear order would be computed by PF,

(4.33)

PF = <**who**, did, **his**, friend, invite, (**who**)>

and *who* would move past *his* in the final PF representation, with which *who* is interpreted as co-referent at LF. The assumption is that this movement of a wh-phrase past a co-referent element is blocked by the grammar.

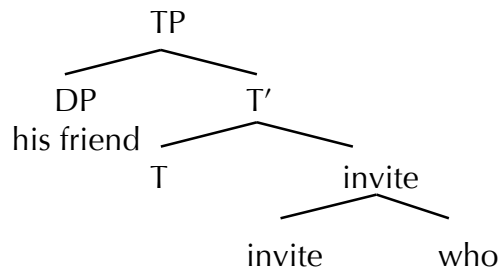
There are at least two main problems with this approach. The first is a conceptual problem. Even though this approach translates Ross' and Postal's idea into Minimalist notions, it is unclear how information about one of the interfaces can be visible to the other. More specifically, the step in (4.33) above is suspicious in this respect, in that that PF representation has no information about the indices for *who* or *his*.

The second problem is an empirical one, and it has to do with the fact that quantifier raising also gives rise to WCO effects, even though no overt movement is

involved (so no information is fed to PF). Again, the problem is the step in (4.33) which simply never takes place in cases involving quantifier raising.

The alternative approach that I propose here is also derivational but does not need to assume that PF information is computed derivationally. The timing of PF interpretation will be irrelevant here. It still assumes, though, that semantic interpretation (including scopal relations) proceeds derivationally. Let's consider the step of the derivation in (4.34):

(4.34)



As argued above, a semantic interpretation at this point can be assigned in which *his* co-refers with *who*. This possibility should not be blocked, as a sentence like *his<sub>i</sub> friend invited who<sub>i</sub>?* is acceptable under an echo question interpretation.

The proposal that I would like to make here is that WCO effects are due to a violation of the following condition:

(4.35) Scope Preservation of co-referential elements

Once co-reference between two syntactic objects  $\alpha$  and  $\beta$  has been assigned,  $\alpha$  and  $\beta$  cannot undergo a movement operation that alters their relative scope.

As can be seen, this approach differs from my previous proposal (and also Ross' 1967 and Postal 1971) in that it does not make reference to linear order. It is also different from Koopman and Sportiche (1982) in that it is a syntactic constraint, rather than a constraint on the LF representation.

Consider again the contrast in (4.29), repeated here as (4.36)

(4.36)

- a. Who<sub>i</sub> e<sub>i</sub> invited his<sub>i</sub> friend to the party?
- b. \*Who<sub>i</sub> did his<sub>i</sub> friend invite e<sub>i</sub> to the party?

In (4.36a), before wh-movement applies, *who* has scope over *his* and *his* does not have scope over *who*, since e<sub>i</sub> c-commands his<sub>i</sub> and his<sub>i</sub> does not c-command *who*. These scopal relations remain after wh-movement applies. Conversely, (4.36b) violates the condition in (4.35). Before wh-movement applies neither *who* nor *his* has scope over the other (since no c-command obtains in either direction). However, once wh-movement applies, *who* has scope over *his*.

As is well known, A-Movement does not give rise to WCO effects, as shown in (4.37):

(4.37)

**Who<sub>i</sub>** seems to his<sub>i</sub> mother **e<sub>i</sub>** to be happy

An interesting way to link the condition in (4.35) to the lack of WCO effects in A-Movement would be to say that scope is assigned after Case checking/valuation takes place, assuming that Case valuation is what drives A-Movement (see Chomsky 1995, Epstein and Seely 2006, among others, *contra* Chomsky 2001). This would mean that (4.35) is evaluated when *who* is already in the matrix Spec-TP position in (4.37). If this is correct, (4.37) is similar to (4.36a) in the relevant aspects, so no WCO effects are expected.

Consider now the data in (4.38) (see also Lasnik & Stowell 1991, Citko 2005), which shows that WCO effects are absent in the second conjunct/adverbial clause (yet still present in the first conjunct/main clause), hence the term “Weakest Crossover” coined by Lasnik and Stowell.

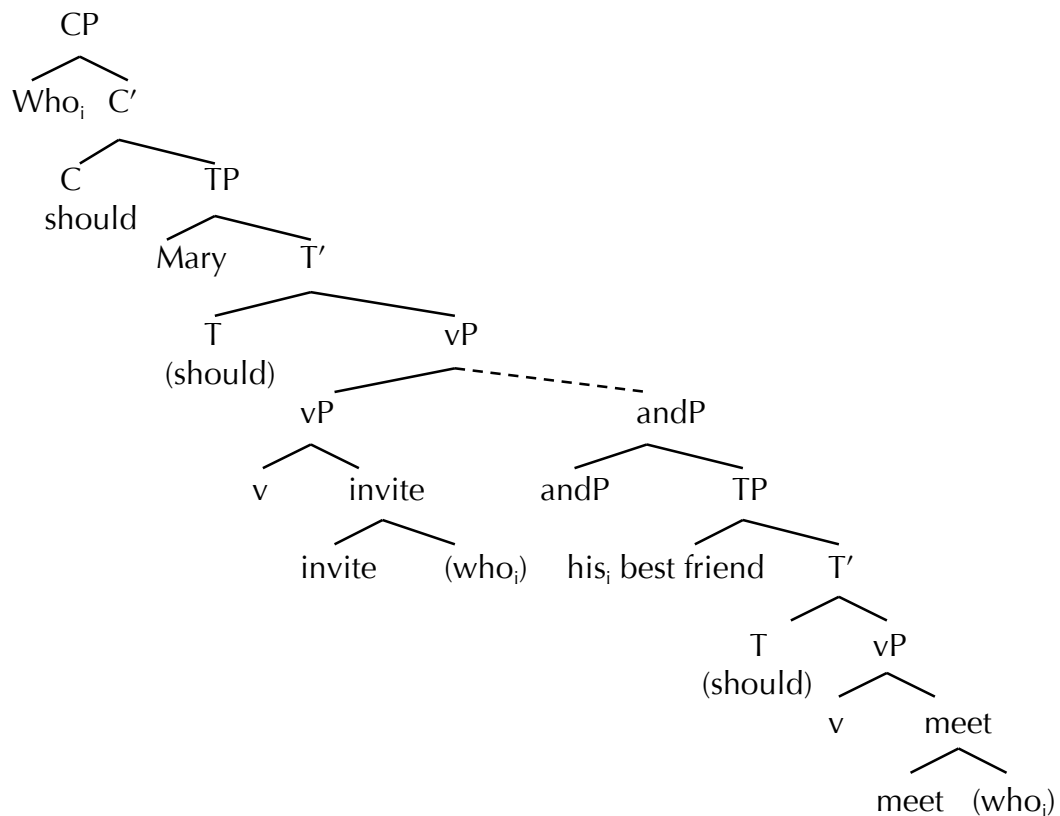
(4.38)

- a. Who<sub>i</sub> should Mary invite e<sub>i</sub> [and his<sub>i</sub> best friend meet e<sub>i</sub>]?
- b. \*Who<sub>i</sub> should his<sub>i</sub> best friend invite e<sub>i</sub> [and Mary meet e<sub>i</sub>]?
- c. Who<sub>i</sub> did you stay with e<sub>i</sub> [before his<sub>i</sub> best friend had spoken to e<sub>i</sub>]?
- d. \*Who<sub>i</sub> did his<sub>i</sub> best friend stay with e<sub>i</sub> [before you had spoken to e<sub>i</sub>]?

It should be noticed that some speakers do not accept a PG construction like (4.38c). For speakers who do, however, there is a clear contrast between (4.38c) and (4.38d).

As we saw in section 4.2, under the Sideward Movement approach, *who* is copied from the object position of *meet* and merged with *invite*, before *andP* adjoins to *vP* (for the sentence in (4.38a)):

(4.39)



This first (sideward) movement of *who* from object of *meet* to object of *invite* does not alter their relative scope; none of them had scope over the other before

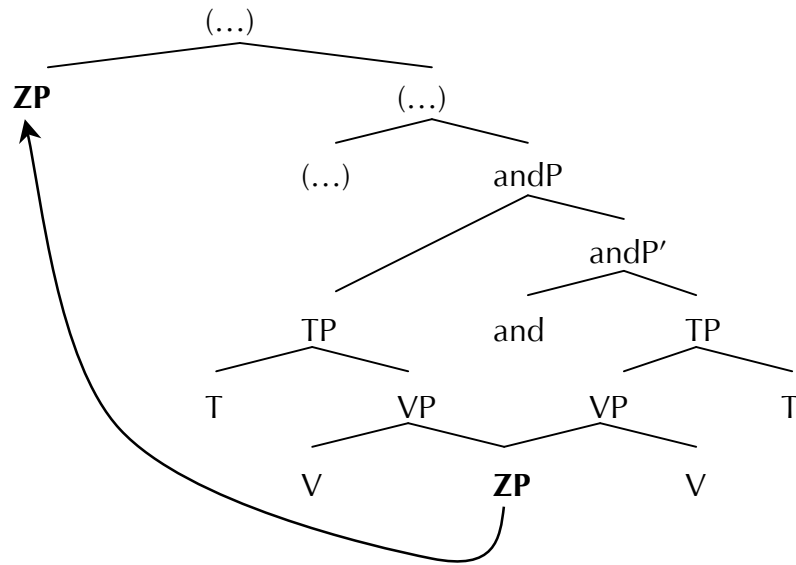


movement and none of them have scope over the other after movement either. Consider now the second movement of *who*, from object position of *invite* to Spec-CP. Does this movement alter the relative scope of *who* and *his*?

If we look at the tree in (4.39) the answer is “yes.” After this movement takes place, *who* now takes scope over *his*. However, as we saw in section 3.2, we know that adjuncts can be inserted late in the derivation (recall the discussion on (3.13) and (3.14)). Suppose that the andP in (4.39) does not attach until after movement of *who* to Spec-CP takes place. In that case, the operation of movement crucially did not alter the relative scope of *who* and *his*. What altered their scope was the operation of adjunction of andP to vP. Therefore, if the tree in (4.39) is derived through late adjunction of the andP, the condition in (4.35) above is satisfied.

The asymmetry in WCO effects between the first conjunct and the second conjunct pointed out above is not easy to explain under Citko’s (2005) multidominance/parallel Merge analysis of ATB constructions (which could in principle be extended to PG constructions). Under her approach, the element that undergoes ATB-extraction originates as daughter of both verbs, the one on the first conjunct and then one on the second conjunct, as shown in (4.40) (Citko assumes the X’ analysis of coordination).

(4.40)



As can be seen, the two TPs are built in a parallel symmetric fashion, which makes any asymmetric behavior in the conjuncts hard to account for (as Citko herself acknowledges).

These WCO effects are also problematic for a deletion approach to ATB constructions (see An 2007 and references therein). Under this approach, the elements in the CP domain of the second conjunct delete under identity with the first conjunct, as sketched in (4.41) below. It is true, however, that one could think of a solution to this problem, in terms of repair by Ellipsis (see section 3.5 above).

(4.41)

[Who does John like] and [~~who does~~ Mary dislike]

Lasnik and Stowell (1991) also report that WCO effects do not arise either in other constructions involving operator movement, like tough-movement and relative clauses.

Interestingly, however, in the constructions where we do not find WCO effects, these effects are also absent with overt phrases. This can be illustrated with relative clauses, in which both overt and null operators are possible, as shown in (4.42):

(4.42)

This is [the employee]<sub>i</sub> who<sub>i</sub>/Op<sub>i</sub> his boss couldn't stand e<sub>i</sub>

This example provides evidence that the absence of WCO effects is not due to special properties of the null operator (or rather the null epithet e<sub>i</sub> as Lasnik and Stowell proposed). Alternatively, the absence of WCO effects could be due to the way that these constructions are derived. One possibility is that what these constructions have in common is that they also involve Sideward Movement, and thus avoid WCO in the same way ATB and PG constructions do under the account presented here. I elaborate on this idea in section 4.5 below.

It is interesting to note, however, that *both* first and non-first conjuncts display *Strong Crossover* (SCO) effects, as shown in (4.43):

(4.43)

- a. \*Who<sub>i</sub> did Mary invite e<sub>i</sub> [and he<sub>i</sub> see e<sub>i</sub>]?
- b. \*Who<sub>i</sub> did he<sub>i</sub> invite e<sub>i</sub> [and Mary see e<sub>i</sub>]?
- c. \*Who<sub>i</sub> did you stay with e<sub>i</sub> [before he<sub>i</sub> had spoken to e<sub>i</sub>]?
- d. \*Who<sub>i</sub> did he<sub>i</sub> stay with e<sub>i</sub> [before you had spoken to e<sub>i</sub>]?

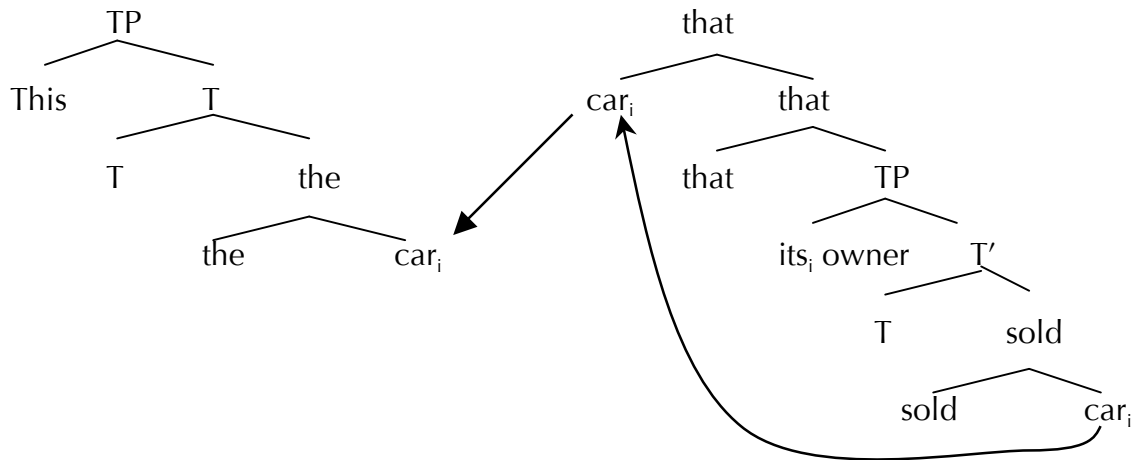
This is expected under the view that the SCO constraint reduces to Condition C of the Binding Theory, which renders the interpretation in which *he* co-refers with a(ny) trace of *who* impossible. Again, if interpretation at LF proceeds derivationally, the interpretation where *he* would co-refer with *who* is disallowed already when the syntactic object *he see who* is interpreted at the LF interface, so movement of an element past its antecedent would never occur in these cases.

If this is correct, SCO and WCO constraints are different in nature; SCO (reducible to Condition C) is representational, whereas WCO is derivational, that is, it is a constraint on the movement operation.

#### **4.5. Extensions of Sideward Movement**

A question that arises then under the approach that I present here is whether all Weakest Crossover configurations can be derived through Sideward Movement. Interestingly, Henderson (2007) provides a sideward movement analysis of Relative clauses, on independent grounds. Under his approach, sketched in (4.44), *car* moves from its base position to Spec,CP, and is copied and merged with *the*:

(4.44)



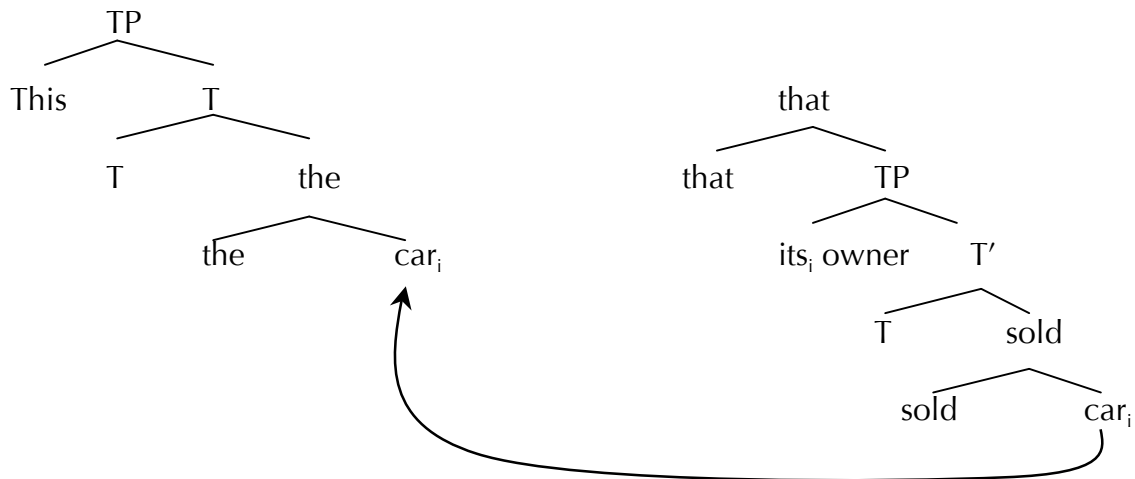
The CP would then adjoin to *car*, yielding (4.45):

(4.45)

This is [<sub>DP</sub> the [<sub>NP</sub> *car*<sub>*i*</sub> [<sub>CP</sub> (*car*)<sub>*i*</sub> [<sub>C</sub> that [<sub>IP</sub> [<sub>DP</sub> *its*<sub>*i*</sub> owner] sold (*car*)<sub>*i*</sub>]]]]]

As can be seen, though, when *car* moves to Spec-CP it takes scope over *its* as a result of movement (therefore violating (4.35) above), incorrectly predicting a WCO violation. One possible modification of this analysis, which would evade WCO, is that *car* undergoes sideward movement directly from its base position, without moving to Spec-CP, as in (4.46):

(4.46)



If this modified analysis is on the right track, we predict no WCO effects in relative clauses, yielding the right result.

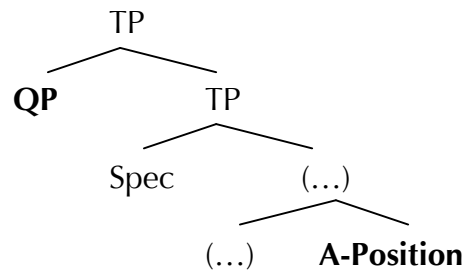
Even though lack of movement to Spec-CP might look problematic, it relates to Lasnik and Stowell's insight that in Weakest Crossover contexts, the null operator is not quantificational, hence their proposal that a null epithet (rather than a variable) is involved in these cases. Again, even though their approach does work, it involves proliferation of different kinds of empty categories that seem to be very restricted with respect to which syntactic environments they can appear in (recall section 4.1 above). Now there are two kinds of Null Operators, one quantificational, the other one non-quantificational, and each spawns a different type of trace in their base position. This last idea is also incompatible with Minimalism's incorporation of the Copy theory of movement.

I think that Lasnik and Stowell's idea can be translated into Minimalism in an interesting way if we assume the Sideward Movement approach. I am going to

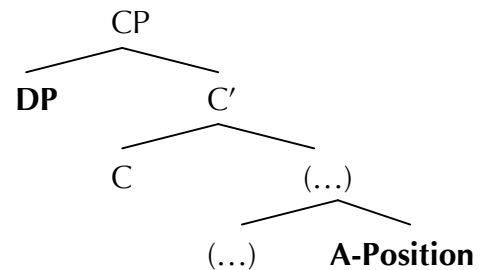
make the assumption that true quantification at LF is always and only achieved by adjunction to TP (in the case of a QP constituent) or movement to Spec,CP (in the case of a [+wh] DP constituent). This movement creates the familiar Operator-Variable configuration, with the Operator taking scope over the Variable (the copy left in the A-Position after A'-Movement applies). This is illustrated in (4.47):

(4.47)

a. Quantifier Raising:



b. [+wh] DP Movement:



If these assumptions are correct, WCO effects arise in the kind of Movement/configurations illustrated in (4.47), as shown in (4.48):

(4.48)

a. \*His<sub>i</sub> mother loves everyone<sub>i</sub> (at LF: [[**everyone**]<sub>i</sub> [his mother loves e<sub>i</sub>]])

b. \*Who<sub>i</sub> does his<sub>i</sub> mother love e<sub>i</sub>

Conversely, examples where no true quantification takes place (and so no WCO effects arise) are the result of Sideward Movement to another syntactic object

without movement to Spec-CP or adjunction to TP, rather than being the result of *movement of a special kind of empty category* as proposed by Lasnik and Stowell (1991). All Weakest Crossover contexts then, should be analyzable in terms of Sideward Movement.

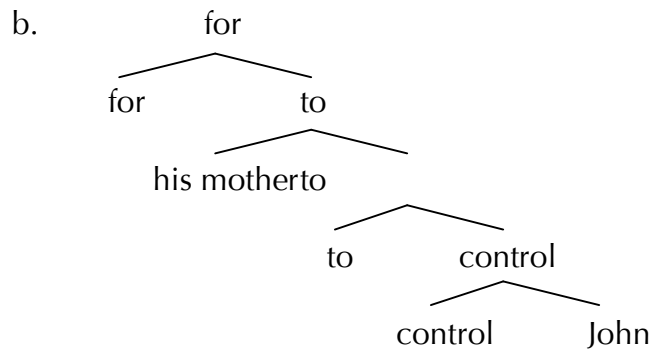
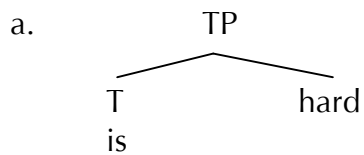
In this respect, Lasnik and Stowell also report that Tough-Movement (see e.g. Chomsky 1964, 1982) is a Weakest Crossover context since it does not give rise to WCO effects. For this case, I would like to propose the following Sideward Movement analysis (see also Hornstein 2001 and Taylor 2003). Let's consider a case of Tough-Movement like (4.49):

(4.49)

John<sub>i</sub> is hard for his<sub>i</sub> mother to control

First, the two separate syntactic objects in (4.50) are built.

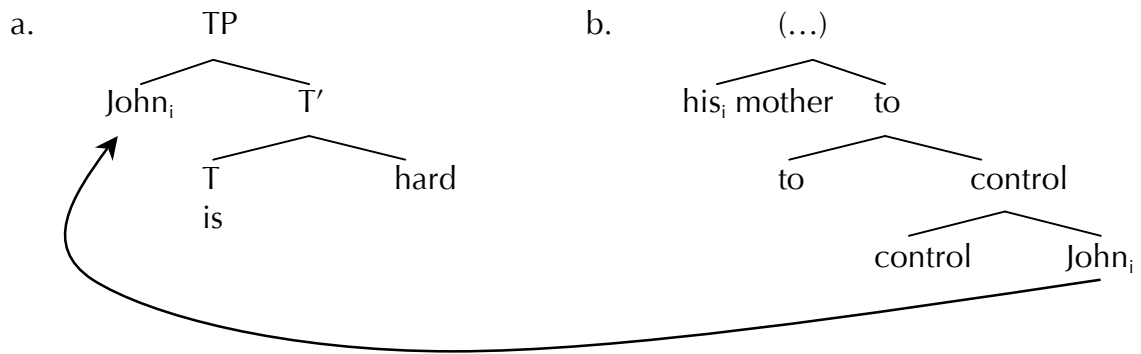
(4.50)





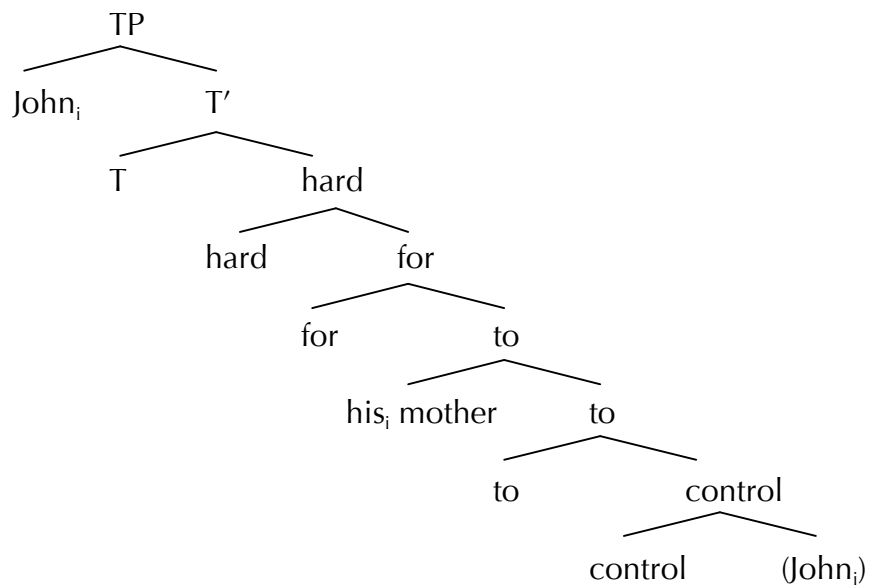
The EPP of the matrix T in (4.50a) needs to be satisfied. This can be done through Merge of an expletive (ultimately yielding *it is hard for his mother to control John*). It can also be done through Sideward Movement of *John*, as in (4.51):

(4.51)



Finally, the CP in (4.50b) (acyclically) adjoins to the Adjective *hard*, yielding (4.52).

(4.52)



This analysis solves one of the classic problems that arise in Tough-Movement cases, namely, the fact that *John* seems to be moving from an A-position to another A-position crossing an intermediate CP. Under this analysis, *John* never actually crosses an intermediate CP. Moreover, this analysis accounts for the fact that *John* does not get a  $\theta$ -role from the predicate *hard*, since it is never actually merged with it. It also accounts for the fact that *John is hard for his mother to control* is synonymous with *it is hard for his mother to control John*, since the difference between these two sentences reduces to whether the EPP of the matrix T has been satisfied by Merge of an Expletive or Sideward Movement of *John*. None of these operations contribute to the semantics of the sentence, since they are purely syntactic.

Notice that this analysis still obeys the requirement that a chain be formed involving both copies of *John*, since the two copies are in a c-command relation that ensures that a chain can be formed and eventually reduced.

In conclusion, the Sideward Movement approach provides an account of the main properties of ATB and PG constructions and it also offers a principled unified explanation for why ATB, PG, relative clause, and Tough-Movement constructions all display Weakest Crossover effects. In the spirit of the Minimalist Program, these results are obtained without relying on the stipulated inventory of empty categories that was proposed in the Government and Binding Theory.

## CHAPTER 5

### **General summary and implications for sentence processing**

In this dissertation I have investigated the syntax of TP-coordination, with respect to some of its structure and properties, and also the properties of related phenomena including ATB extraction and PG constructions.

#### **5.1. Summary**

In chapter 2 I discussed the syntactic analysis of simple cases of coordination. I briefly reviewed three of the main analyses that have been proposed for coordinate structures in the literature: the Ternary branching analysis, the X' analysis, and the adjunction analysis. I argued that even though both the X' and the adjunction analyses conform to the principles of X'-Theory, the adjunction analysis of Munn (1993) can account for asymmetries in coordinate structures regarding selectional requirements in a more straightforward way. Then I briefly discussed issues regarding word order in coordinate structures. Once the adjunction analysis is assumed, we need to account (as in any case involving adjunction) for why the lexical items in the adjunct follow the structure that the adjunct attaches to, even

though the adjunct c-commands the structure that it adjoins to. In section 2.3 I argued that the adjunction site of non-first conjuncts is not TP, as standardly assumed, but rather vP (VP in derivations where no v is present, i.e., unaccusatives and passives). This “lowered adjunction” analysis accounts for why in TP coordination the subject of the first conjunct c-commands into the second conjunct, as evidenced by previously unnoted condition C effects. This proposal lead to discussion of the validity of other c-command tests when more than one clause is involved. More specifically, I argued that bound variable readings for pronouns and NPI licensing can fail even when c-command obtains, as it can be independently blocked by intrinsic lexical properties of different conjunctions, which do not correspond to the traditional coordinating/subordinating distinction. In sections 2.4 and 2.5 I argued precisely that the coordination/subordination distinction cannot be expressed in syntactic terms, given that my version of the adjunction analysis (at least in the case of TP-coordination) implies that coordination and subordination display identical syntactic configurations when the conjunctions have the same (i.e., TP) selectional requirements.

Chapters 3 and 4 explore two main consequences of my version of the adjunction analysis. Chapter 3 builds on ideas put forward by Munn (1993) regarding the nature of the CSC. I first argued that maintaining the CSC as a syntactic constraint would yield a redundancy between the CSC and the CED, and that the CED is better motivated by very general properties of derivations, namely, the need for additional derivational workspaces to be created when phrasal

specifiers and adjuncts are built by the syntax. The CED however does not rule out extraction from a first conjunct, which I argue is a welcome result, since there is evidence that the syntax in fact allows extraction from a first conjunct. The ban on extraction only from a first conjunct is representational, and is evaluated at LF, which requires conjuncts to obey a parallelism requirement. In the remainder of the chapter I explore the parallelism requirement on the PF side of the grammar, and present novel data from Spanish that shows that conjuncts also need to be parallel at PF. Even more interesting is the fact that once extraction applies, the non-extracted elements need to display the same relative order in all conjuncts. Surprisingly, Spanish, unlike English, shows that the parallelism condition at LF does not require the gaps to be parallel with respect to grammatical function, since Spanish allows ATB extraction even when the gaps correspond to different grammatical functions. I briefly discuss the implication of these phenomena for the theory of linearization and the interaction between the syntax and the PF component of the grammar, in terms of the syntactic information that is visible once Spell-Out applies.

In chapter 4 I discuss the analysis of ATB and PG constructions. Given my version of the adjunction analysis of coordination, the only thing that differentiates these two constructions is the choice of lexical items. Their syntactic configurations are identical. This formalizes the long-standing intuition that ATB and PG should be analyzed along the same lines. I first review one of the approaches that try to unify the two constructions, the Null Operator analysis. I note that the general

assumptions that one must make for this approach to be tenable are incompatible with key ideas of the Minimalist Program regarding the notion of inclusiveness and the program's rejection of the Government and Binding inventory of empty categories and their special properties. Moreover, this analysis cannot account for the fact that only overt movement can license ATB and PG constructions, as evidenced by so-called wh-in situ languages like Chinese.

I then present the main idea behind the theory of Sideward Movement, and argue how it can straightforwardly account for the main properties of ATB and PG constructions, without appeal to a stipulated inventory of empty categories. I also discuss the fact that Sideward Movement is not considered part of current standard Minimalist theory, since it is based on slightly different assumptions about the nature of Merge and movement. I argue that while these differences could be reconciled, the idea that derivations proceed Phase by Phase seems to me to be really incompatible with Sideward Movement. Before I turn to discussion of the crossover properties of ATB and PG constructions, I discuss some asymmetries that are present in these constructions with respect to reconstruction and CED islandhood. I discuss data presented by Nissenbaum (2000) and others that show that only first conjuncts/main clauses are possible reconstruction sites. I formalize this difference in terms of (lack of) isomorphism between movement and chain formation: isomorphic chains allow for reconstruction in their different links, while non-isomorphic chains do not allow for reconstruction and their links are not visible at LF for computation of Binding Theory principles.

I then move on to argue that WCO phenomena can be better accounted for under the Sideward Movement analysis, while it is problematic for other approaches, like the Null Operator approach, the Parallel Merge approach and the PF deletion analysis. I formulate a derivational version of WCO, and argue that once co-reference is assigned to two syntactic objects, they cannot undergo a movement operation that will change their relative scope. I then show that this accounts for what Lasnik and Stowell dubbed “Weakest Crossover”, which under my approach obtains when late adjunction of the syntactic element that contained the syntactic object that underwent Sideward Movement takes place. I then claim that this approach can also account for other Weakest crossover contexts, like relative clauses and Tough-Movement.

In conclusion, this dissertation has been an attempt to examine coordination and related phenomena (ATB and PG constructions) and understand how these phenomena relate to one another. This has been done under the assumption that non-first conjuncts are concatenated by adjunction, just as adverbial clauses are, and the assumption that the syntax allows for constrained instances of Sideward Movement. By assuming and modifying these two main approaches put forward by Munn and Nunes respectively, I have tried to account for the properties of these phenomena, noting some problems and providing possible solutions. Taking the assumptions behind these two approaches seriously has proved to have rich theoretical and empirical implications, and I hope that the discussion of these

implications will shape future research in the field and thus contribute to our understanding of the phenomenon of coordination.

## 5.2. Possible implications for sentence processing

A final question that I would like to explore briefly in this dissertation concerns the implications of the theoretical discussion above for sentence processing. Since ATB and PG constructions display the interesting property that a single element is interpreted in two different thematic positions, they raise interesting parsing-related questions that to the best of my knowledge have not been systematically addressed (though see Phillips 2006). This final section is intended as a brief discussion of these issues.

In chapter 4 I presented two competing analyses of these constructions: the sideward movement approach (Nunes 2004) and the null operator movement approach (Munn 1993). The basic distinction between these two approaches, more informally, is that the number of dependencies posited in the sideward movement approach is smaller than the number of dependencies assumed by the null operator approach. These dependencies are sketched in (5.1) and (5.2):

(5.1) *Sideward Movement analysis*


**[Which book]** has Sue read **e** recently and Peter studied **e** for a while?





(5.2) *Null Operator analysis*

**[Which book]** has Sue read **e** recently **Op** and Peter studied **e** for a while?



The diagram consists of three curved arrows. The top arrow starts at the gap 'e' after 'read' and points to the overt element '[Which book]'. The middle arrow starts at the gap 'e' after 'studied' and points to the overt element '[Which book]'. The bottom arrow starts at the null operator 'Op' and points to the gap 'e' after 'studied'.

Informally, the Sideward Movement analysis creates a dependency between an overt element (*which book*) and each of the two gaps. Conversely, the Null Operator analysis assumes that there is one distinct dependency established for each gap, besides the necessary dependency between the two chains (since all the gaps refer to the same entity).

I am assuming here that the dependencies that have to be computed correspond with the chains that are formed in the syntax, and not necessarily correspond with the movement operation that spawned the creation of those chains. This is specially important in the Sideward Movement approach, since the overt element, though related to the second gap by a chain, was never related to that position by a movement operation (see section 4.2).

I believe that these claims about the structure of these constructions have implications for sentence processing mechanisms that are worth addressing. Assuming that the number of dependencies is an index of processing difficulty, we in principle have a metric that can help us determine which one of these approaches is more compatible with experimental findings. One possible way to approach this question is to construct a paradigm with sentences like (5.1-2) above, repeated here as (5.3i), sentences like (5.3ii), which involve two clearly distinct

dependencies, and sentences like (5.3iii), which involve no dependencies. We could then devise a self-paced reading task, for example, and compare relative reading times at the relevant points in the sentence.

(5.3)

- (i) **[Which book]** has Sue been reading **e** recently and **[which paper]** has Peter been studying **e** for a while?
- (ii) **[Which book]** has Sue been reading **e** recently and Peter been studying **e** for a while?
- (iii) Sue has been reading recently and Peter has been studying for a while.

(5.3iii) provides a reading time base-line, since it involves no dependencies. (5.3i) would help us measure the reading time cost of having one wh-dependency (**[which paper], e**), by comparing average reading times at *for a while* with reading times for the same region in (5.3iii). If ATB constructions like (5.3ii) indeed involve a mental representation like (5.2), reading times at *for a while* should be significantly longer than for (5.3i), since more dependencies are involved (the **(Op, e)** dependency *plus* that chain's dependency with the wh-chain in the first clause). Conversely, if they involve a representation like (5.1), reading times at *for a while* should be similar to those found for (5.3i).

If the experiment proposed here yields reliable differences in reading times, it will allow us to make interesting connections between syntactic theories and on-line processing mechanisms, which hopefully will be the basis for more future research in this domain.

## REFERENCES

## REFERENCES

- Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. "Parametrizing AGR: Word order, V-movement, and EPP-checking." *Natural language and linguistic theory* 16:491-540.
- An, Duk-Ho. 2007. "Asymmetric T-to-C Movement in ATB Constructions: A PF Deletion Analysis." *Proceedings of the 14th Annual Conference of the Student Organization of Linguistics in Europe*.
- Bobaljik, Jonathan, and Sam Brown. 1997. "Interarboreal operations: Head movement and the extension requirement." *Linguistic Inquiry* 28:345–356.
- Borsley, Robert. 2005. "Against ConjP." *Lingua* 115:461-482.
- Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton de Gruyter.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origins, and Use*. New York: Praeger.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. "Derivation by phase." In Michael J. Kenstowicz (ed.), *Ken Hale: A life in Language*, 1–52. Cambridge, MA: MIT Press.
- Chung, Sandra, William Ladusaw, and James McCloskey. 1995. "Sluicing and Logical Form." *Natural Language Semantics* 3:239-282.
- Citko, Barbara. 2000. *Parallel merge and the syntax of free relatives*. Doctoral dissertation. SUNY at Stony Brook.
- Citko, Barbara. 2005. "On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge." *Linguistic Inquiry* 36:475-496.
- Collins, Christopher. 1997. *Local Economy*. Cambridge, MA: MIT Press.
- Culicover, Peter W. and Paul M. Postal, eds. 2001 *Parasitic Gaps*. Cambridge, MA: MIT Press.

- Epstein, Samuel, Erich Groat, Hisatsugu Kitahara, and Ruriko Kawashima. 1998. *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.
- Epstein, Samuel, and T. Daniel Seely, eds. 2002. *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell.
- Epstein, Samuel, and T. Daniel Seely. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.
- Fernández-Salgueiro, Gerardo. 2007. "Reducing computation at the interface with the Sensory-Motor Systems: a derivational approach to (Chain) linearization." *Proceedings of the 14th Annual Conference of the Student Organization of Linguistics in Europe*.
- Fernández-Salgueiro, Gerardo. 2008. "Deriving the CSC and Unifying ATB and PG Constructions through Sideward Movement." *Proceedings of the 26<sup>th</sup> West Coast Conference on Formal Linguistics*, 156-162.
- Fernández-Salgueiro, Gerardo. To appear. "Romance Null Subjects at the Sensory-Motor Interface." To appear in *Syntax*.
- Fernández-Salgueiro, Gerardo, and Michael Marlo. 2006. "He himself and Binding Domains in a Minimalist Framework." In Cedric Boeckx (ed.), *Minimalist Essays*, 221–231. Philadelphia: John Benjamins.
- Fiengo, Robert. 1977. "On Trace Theory." *Linguistic Inquiry* 8:35-61.
- Fortin, Catherine. 2007. "Some (not all) nonsententials are only a phase." *Lingua* 117(1):67-94.
- Fox, Danny, and Howard Lasnik. 2003. "Successive cyclic movement and island repair: The difference between Sluicing and VP Ellipsis." *Linguistic Inquiry* 34:143-154.
- Fox, Danny, and David Pesetsky. 2004. "Cyclic Linearization of Syntactic Structure." *Theoretical Linguistics* 31:1-46.
- Frampton, John, and Sam Gutmann. 2002. "Crash-Proof Syntax." In Epstein and Seely, eds.
- Goodall, Grant. 1987. *Parallel structures in syntax: Coordination, causatives and restructuring*. Cambridge: Cambridge University Press.

- Henderson, Brent. 2007. "Matching and Raising Unified." *Lingua* 117(1):202-220
- Higginbotham, John. 1980. "Pronoun and bound variable." *Linguistic Inquiry* 11: 679-708.
- Hornstein, N. 2001. *Move! A Minimalist theory of construal*. Oxford: Blackwell.
- Hornstein, Norbert, and Jairo Nunes. 2002. "On Asymmetries between Parasitic Gap and Across-the-Board Constructions." *Syntax* 5,1:26-54.
- Jackendoff, Ray. 1977. *X'-syntax: a study of phrase structure*. Cambridge, MA: MIT Press.
- Johannessen, Janne. 1998. *Coordination*. Oxford: Oxford University Press.
- Koopman, Hilda, and Dominique Sportiche. 1982. "Variables and the Bijection Principle." *The Linguistic Review* 2:139-160.
- Lasnik, Howard, and Tim Stowell. 1991. "Weakest crossover." *Linguistic Inquiry* 22:687-720.
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason. 2008. "An asymmetry in voice mismatches in VP-ellipsis and pseudogapping." *Linguistic Inquiry* 39:169-179.
- Munn, Alan. 1993. *Topics in the Syntax and Semantics of Coordinate Structures*. Doctoral dissertation. University of Maryland, College Park.
- Munn, Alan. 2001. "Explaining Parasitic gap restrictions." In Culicover and Postal, eds.
- Nissenbaum, Jonathan. 2000. *Investigations of Covert Phrase Movement*. Doctoral Dissertation, MIT.
- Nunes, Jairo. 1995. *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*. Doctoral dissertation, University of Maryland, College Park.
- Nunes, Jairo. 2001. "Sideward Movement." *Linguistic Inquiry* 32:303-344.
- Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

- Phillips, Colin. 2006. "The Real-Time Status of Island Phenomena." *Language* 82(4):795-823.
- Postal, Paul. 1971. *Crossover Phenomena*. New York, NY: Holt, Rinehart and Winston.
- Ritter, Elizabeth, and Sara Rosen. 1993. "The Independence of External Arguments." In *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL)*. CSLI Publications.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Doctoral dissertation, MIT.
- Sells, Peter. 1984. *Syntax and Semantics of Resumptive Pronouns*. Doctoral Dissertation, UMass, Amherst.
- Stowell, Tim. 1981. *Origins of Phrase Structure*. Doctoral Dissertation, MIT.
- Taylor, Heather. 2003. *On Tough-constructions and Construal-as-Movement: Be Tough! It's too easy to give up*. MA Thesis, Eastern Michigan University.
- Thompson, Ellen. 1995. "The Structure of Tense and the Syntax of Temporal Adverbs." *Proceedings of the West Coast Conference on Formal Linguistics*, 499-514.
- Williams, Edwin. 1978. "Across-the-Board Rule Application." *Linguistic Inquiry* 9:31-43.