

**Economy and Semantic Interpretation - A study of scope  
and variable binding**

by

Daniel, J. Fox

M.A., Linguistics  
Tel-Aviv University  
(1993)

Submitted to the Department of Linguistics and Philosophy  
in Partial Fulfillment of the Requirements  
for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology  
July 1998

© Daniel Fox  
All rights reserved.

The author hereby grants MIT permission to reproduce and to distribute publicly paper  
and electronic copies of this thesis document in whole or in part.

Signature of Author: \_\_\_\_\_  
Department of Linguistics and Philosophy

July 23, 1998

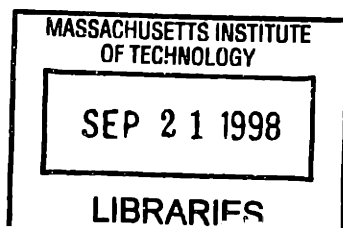
Certified by: \_\_\_\_\_  
Professor Noam Chomsky  
Thesis Supervisor

Certified by: \_\_\_\_\_  
Professor Kai von Fintel  
Thesis Supervisor

Certified by: \_\_\_\_\_  
Professor Irene Heim  
Thesis Supervisor

Certified by: \_\_\_\_\_  
Professor David Pesetsky  
Thesis Supervisor

Accepted by: \_\_\_\_\_  
Professor Michael Kenstowicz  
Chairman, Department of Linguistics



ARCHIVES

# **Economy and Semantic Interpretation - A study of scope and variable binding**

by  
Danny Fox

Submitted to the Department of Linguistics and Philosophy  
on July 23, 1998 in partial fulfillment of the requirements for the Degree of Doctor of  
Philosophy in Linguistics

## **ABSTRACT**

This thesis argues that principles of economy play an important role in determining the form of linguistic structures at the interface with the semantic component. The argument comes in the form of a variety of proposals regarding specific mechanisms that are active at this interface. Specifically, I suggest that Economy plays a role in determining relative scope, variable binding, and the form of Operator Variable Constructions.

**Economy and Scope** (chapters 2-3): In these chapters I investigate the scopal properties of sentences that are semantically identical under Surface- and Inverse-Scope (Scopally Uninformative sentences). Based on a variety of empirical tests, I argue for the following generalization:

- (1) Scopally Uninformative sentences are restricted to Surface-Scope. (i.e. Inverse-Scope is possible only when it is semantically distinct from Surface-Scope.)

This generalization argues for an economy condition (Scope-Economy) which rules out semantically vacuous applications of operations (such as QR and QL) which reverse the relative scope of two expressions.

**Economy and Variable Binding** (chapter 4): In this chapter I investigate the scopal properties of sentences that are semantically identical under Local- and Non-local Variable Binding (Binding Uninformative sentences). Using one of the tests from the previous chapters, I argue for the following generalization:

- (2) Binding Uninformative sentences are restricted to Local Variable binding (i.e. Non-local Binding is possible only when it is semantically distinct from Local-Binding.)

This generalization argues for an economy condition (VB-Economy) which rules out Non-local Binding when it is semantically vacuous (a condition which was suggested on independent grounds in Heim (1993)).

**Economy and Operator Variable Constructions** (Chapters 5-6): In these chapters, I use Binding Theory (BT) to argue that A-bar chains are converted to operator variable constructions under an economy condition which prefers to maintain as much of the material at the tail of the chain as possible for interpretability (OV-Economy). OV-Economy accounts for an observation of Fiengo and May that QR affects BT(C) only when

ACD is involved. The discussion yields new evidence that BT applies at LF (and only at LF) as suggested in Chomsky (1993).

Thesis Supervisors: Profs. Noam Chomsky, Kai von Stechow, Irene Heim and David Pesetsky

## Acknowledgments

This thesis is the result of five years of work. A lot has happened to me during this time, and there are many people who have made an impact on the way things turned out. I am especially indebted to the members of my committee: Noam Chomsky, Kai von Fintel, Irene Heim and David Pesetsky. They have all been extremely generous with their time, and were a source of encouragement and support. It is hard for me to imagine how my work would have turned out had I not been fortunate enough to be around these people. Other teachers from whom I have learned a lot include Yosi Grodzinsky, Ken Hale, Alec Marantz and Tanya Reinhart.

The discussions at the LF reading group were a great source of happiness. Thanks to Michel DeGraf, Diana Cresti, Kai von Fintel, Martin Hackl, Paul Hagstrom, Irene Heim, Jim Higginbotham, Sabine Iatridou, Lisa Matthewson, Renate Musan, Jon Nissenbaum, Philippe Schlenker, David Pesetsky, and especially to Orin Percus and Uli Sauerland.

This work would be impossible without the generous help of my fellow students. I would especially like to mention Jonathan Bobaljik, Marie Claude Boivin, Martin Hackl, Paul Hagstrom, Ingvar Lofstedt, Renate Musan, Jon Nissenbaum, Orin Percus, Uli Sauerland, Philippe Schlenker, Hubert Truckenbrodt and Susi Wurmbrand.

This work has its roots in various papers that I have written in the past five years. The people mentioned above provided me with a lot of help while I was writing these papers. I would also like to thank Ellena Anagnostopoulou, Sergey Avrutin, Gennaro Chierchia, Kyle Johnson, Michael Kenstowicz, Howard Lasnik, Martha McGinness, Shigeru Miyagawa, Maire Noonan, Colin Phillips, Roger Schwazchild, Ori Simchen, Norvin Richards, Chris Tancredi, Ken Wexler and Yoad Winter.

As David Pesetsky never failed to mention, the members of my committee were paid to read my thesis. Nevertheless, they did a good job and (as mentioned) I am very grateful. However, some people actually read (portions of) this stuff for free. For the comments I received while I was writing my thesis I would like to thank Marie Claude Boivin, Yosi Grodzinsky, Martin Hackl, Jon Nissenbaum and Hubert Truckenbrodt.

Finally, and most importantly, I would like to thank some dear friends and family members whose presence made life worthwhile. However, the debt that I have to these people goes way beyond what I can afford to mention publicly. I believe and hope that they know.



## **Preface:**

The work reported here began in September 1993 when I was a first year student at MIT. Chapter 2 is a revised version of material that was presented at MIT in November 1993 (Ling-lunch) and has since appeared in *Natural Language Semantics* (Fox 1995a). However, the manner of presentation has changed quite extensively (hopefully for the better), and there are some substantial revisions (most notably in sections 2.3-2.6). Chapter 3 is completely new. It involves a rejection of a hypothesis that I made in Fox (1995a) and allows for what I believe is a real simplification of the proposal. Chapter 4 is an almost unchanged version of a paper I presented in the optimality workshop at MIT (May 1995) which appeared in the proceedings of this conference (Fox 1998a). Chapters 5-6 combine material that appeared in Fox (1995b) with material that is going to appear in Fox (in press).

## Contents

<b>Chapter 1: An Overview</b>	<b>8</b>
1.1. Economy and Scope	9
1.2. Economy and Variable Binding	12
1.3. Binding Theory and the Representation of Scope	14
1.4. BT(C) and Scope Reconstruction	14
1.4.1. The nature of the argument	15
1.4.2. Consequences	16
1.5. BT(C) and QR	18
1.5.1. The Proposal	18
1.5.2. Consequences	20
1.6. BT(A) and Scope-Economy	21
<b>Part I: Interpretation-Sensitive Economy</b>	<b>22</b>
<b>Chapter 2: Economy and Scope</b>	<b>22</b>
2.1. Scope-Economy, Quantifier Raising and Quantifier Lowering	24
2.2. Scope-Economy and Parallelism	31
2.2.1. A Puzzle	32
2.2.2. Parallelism as a detector	33
2.2.3. The ESG	37
2.2.3.1. The ESG and scopal interaction among DPs	37
2.2.3.1. The ESG and QR	45
2.2.3.2. The ESG and QL	46
2.3. Scope-Economy and Coordination	49
2.3.1. The multi-dimensional analysis of coordination	50
2.3.2. Scope-Economy and Coordination - the case of QR	51
2.3.2.1. Setting the stage	52
2.3.2.2. Testing the generalization	54
2.3.2.3. The problem of Telescoping	56
2.3.2.4. Downward entailing operators	57
2.3.2. Scope-Economy and Coordination - the case of QL	59
2.4. The Locality of QR	63
2.5. Modularity - an argument in favor of a deductive system	67
2.6. What is Special about SSOs?	74
<b>Chapter 3: Asymmetries in Ellipsis and the Nature of Accommodation</b>	<b>77</b>
3.1. Problems with the account of the asymmetry in Fox (1995a)	80
3.1.1. The SCC	80
3.1.2. The presupposed generalization	82
3.2. An Independent Asymmetry in Parallelism	83
3.3. The Asymmetry of the ESG follows from the Asymmetry in Parallelism	91
3.4. A Constraint on Accommodation	97
3.4.1. Evidence for the constraint	98
3.4.2. Consequences for the ESG	102
3.5. Conclusion	106

<b>Chapter 4: Economy and Variable Binding</b>	<b>108</b>
4.1. The Basic Proposal	112
4.1.1. Rule H	112
4.1.2. Parallelism	114
4.1.3. Deriving Dahl's observation	116
4.1.4. An extension	117
4.1.5. New predictions	118
4.2. More on Local and Non-local Binding	119
4.3. When Local and Non-Local Binding Yield Different Interpretations	124
4.4. When Local Binding Is Impossible	129
4.4.1. C-command	129
4.4.2. The theta criterion	130
4.5. Rule H and Scope-Economy	132
4.5.1. The Ellipsis Binding Generalization	132
4.5.2. The Ellipsis Scope Generalization	133
4.5.3. An outline of a unification	133
4.6. Conclusion	135
<b>Part II: Binding Theory and the Representation of Scope</b>	<b>136</b>
<b>Chapter 5: Condition C and Scope Reconstruction</b>	<b>137</b>
5.1. Semantic vs. Syntactic Accounts of Scope Reconstruction	138
5.1.1. Syntactic accounts of Scope Reconstruction	141
5.1.2. Semantic accounts of Scope Reconstruction	142
5.1.3. Distinguishing the two accounts via BT(C)	143
5.2. A-bar Reconstruction	145
5.2.1. <i>How many</i> questions - an elaboration on Heycock (1995)	145
5.2.2. Variable Binding - an elaboration on Lebeaux (1990)	154
5.2.2.1. The correlation with BT(C)	155
5.2.2.2. The multitude of intermediate landing sites	157
5.2.3. Unselective Binding	158
5.3. A Reconstruction	162
5.4. Ramifications for the Interpretation of Chains	163
5.5. Conclusion	164
<b>Chapter 6: Economy and Operator Variable Formation</b>	<b>166</b>
6.1. The First Step (Chomsky 1993)	167
6.1.1. Chomsky's proposal	167
6.1.2. The interpretation of A-bar chains (a slight modification)	169
6.2. The Second Step (Fox 1995b)	171
6.3. F&M's Account	173
6.4. A Predicted Generalization	177
6.4.1. ACD constructions that depend on more than one instance of QR	178
6.4.2. Testing the predictions	180
6.5. Scope Reconstruction in A-bar Chains	181
6.5.1. The cases discussed in chapter 5	183
6.5.2. A case of partial reconstruction that doesn't need to include the adjunct	184
6.6. A Note on the A/A-bar Distinction	186
6.7. Condition A and Scope-Economy	188
6.8. Conclusion	191

## Chapter 1: An Overview

Much recent work in linguistic theory shares the idea that among the principles of grammar there are some principles of optimization. Such principles (usually known as principles of Economy or Optimality) become operative under (certain) circumstances in which grammar, as determined by other principles, is non-deterministic. Under such circumstances, the optimality principles choose, from a set of options, those objects (e.g. structures or steps in a derivation) which provide the most optimal solution to a given problem. This idea is sometimes further articulated when coupled with the notion that optimality is to be identified in some manner with “least effort”.<sup>1</sup> Obviously, the idea itself is far from an actual proposal. Indeed proposals vary extensively both in how they measure effort (the economy metric) and in how they specify the problems for which optimal solutions “are being designed” (the problem specification).

This work attempts to provide evidence in favor of the general idea. It argues that principles of “least effort” (henceforth, principles of Economy) play an important role in determining the form of linguistic structures at the interface with the semantic component. The argument comes in the form of a variety of proposals regarding specific mechanisms that are active at this interface. Specifically, I suggest that Economy plays a role in determining relative scope, accommodation, variable binding and the form of Operator Variable Constructions.

The proposals I make depart notably from a prevailing intuition regarding the problem specification.<sup>2</sup> The prevailing intuition is that Economy is designed to ensure that grammar yields an interpretable output in an optimal manner. Economy is oblivious to the nature of the output. For a linguistic object to be considered “a possible solution” (i.e., a member of the set from which “an optimal solution” is selected) all that matters is that it is interpretable at the relevant interface. In other words, the “specified problem” is the problem of yielding an interpretable output. The nature of the interpretation is irrelevant.

---

<sup>1</sup> For some discussion, see Chomsky (1986b, 1989, 1995, 1998a) and references therein. Chomsky sometimes relates the hypothesis that there exists principles of optimality to another hypothesis, namely that the language faculty is in some sense an optimal solution to design specifications which are provided by the interface systems.

<sup>2</sup> With respect to the economy metric, the assumptions I make are very much in line with prevailing intuitions.

My proposals, on the other hand, are based on the intuition that for some principles of economy the nature of the interpretation is not irrelevant. These principles are principles of Interpretation-Sensitive Economy. They are designed to ensure that *a particular* interpretation (“desired” by the outside system) is achieved in an optimal manner. For such principles, a linguistic object is considered “a possible solution” only if it is interpretable at the relevant interface, and *only if the interpretation is identical to some designated specification*. In other words, the “specified problem” is the problem of yielding an interpretable output *with a designated interpretation*.<sup>3</sup> (C.f. also some related work: Golan (1993), Reinhart (1994, 1995, 1997) and Adger (1996).)

In order to spell-out this intuition, some suggestion must be articulated regarding the nature of an output that is utilized by the relevant outside systems. When the relevant outside system is the semantic component, it is natural to think of a “desired outcome” as a designated semantic interpretation, which, in turn, can be thought of (to a first approximation) as a set of truth conditions (though see section 2.5.). The Economy principles for which I argue are thus designed to ensure that a given set of truth conditions is achieved with “no more effort than what is necessary”.<sup>4</sup>

### 1.1. Economy and Scope (chapters 2-3)

The first Economy principle I propose disallows semantically-vacuous applications of covert operations that reverse the relative scope of two expressions (henceforth, Scope-Economy).<sup>5</sup> In other words, Scope Shifting Operations (like Quantifier Raising, QR, and Quantifier Lowering, QL) are allowed to apply only when they are necessary to achieve a designated semantic interpretation.<sup>6</sup>

---

<sup>3</sup> The notion of a *designated interpretation* plays no role in the technical implementation that I outline for the economy principles (in contrast to the implementation I proposed in Fox 1995a). However, I think that the notion has a *motivational* import. (See Chomsky’s (1998b) discussion of the motivation for strong features from the perspective of optimal design. See also note 1.)

<sup>4</sup> For the third economy principle I propose (OV-Economy), it is not clear (at the moment) whether truth conditions are relevant. The reason is that in the cases under which the principle is operative (at least those I’ve looked at) independent factors determine that the set of competitors are interpretively identical.

<sup>5</sup> This principle was proposed independently in Fox (1995) (which is revised in chapters 2-3) and in Reinhart (1994, 1997).

<sup>6</sup> This general idea will be qualified. See 2.5.

- (1) **Scope-Economy:** Scope Shifting Operations (SSOs) can't be semantically vacuous.

To understand the nature of Scope-Economy, consider QR and its relevance for sentences such as those in (2) and (3). These sentences differ in that in (2) the interpretations are different when the subject takes scope over the object and when the scopal relation is reversed, while in (3) the interpretations under the two scopal relations are identical. The Economy principle predicts that sentences such as those in (3) will differ from their counterparts in (2) in not allowing QR to reverse the relative scope of the subject and the object.

- (2) a. A boy loves every girl.  
b. Many boys love every girl.
- (3) a. John loves every girl.  
b. Every boy loves every girl.

QR is a covert operation which has no obvious effects on phonology. Scope-Economy states that it does not apply when it has no effects on semantics. In order to test this statement, it is necessary to distinguish between linguistic representations which are interpreted in an identical manner by both phonology and semantics. For example, it is necessary to know whether the sentences in (3) can have the LF representations in  $(3_{SSO})$ , in which the subject and the object scope is reversed, or whether (as predicted by Scope-Economy) they are limited to the representations in  $(3_{-SSO})$ , in which scope has not been reversed.

- $(3_{-SSO})$  a.  $[_{IP} \text{John}_1 \dots [_{VP} \text{every girl}_2 [_{VP} t_1 \text{ loves } t_2]]]$   
b.  $[_{IP} \text{every boy}_1 \dots [_{VP} \text{every girl}_2 [_{VP} t_1 \text{ loves } t_2]]]$
- $(3_{SSO})$  a.  $*[_{IP} \text{every girl}_2 [_{IP} \text{John}_1 \dots [_{VP} t_2 [_{VP} t_1 \text{ loves } t_2]]]$   
b.  $*[_{IP} \text{every girl}_2 [_{IP} \text{every boy}_1 \dots [_{VP} t_2 [_{VP} t_1 \text{ loves } t_2]]]$

One way to distinguish between such “observationally identical” representations is to identify constraints that would treat the two representations in a different way. Such constraints could indicate the LF position of a quantifier even when semantics fails to distinguish among the different possibilities. In sections 2.2., I will utilize a constraint on phonological reduction and deletion (Parallelism) that has the desired property.

Scope-Economy allows an SSO to apply in sentences such as those in (2) where it has semantic effects. Such sentences, therefore, have both the LFs in  $(2_{SSO})$  and in  $(2_{\neg SSO})$ :

- (2 <sub>$\neg SSO$</sub> ) a. [IP A boy<sub>1</sub>..... [VP every girl<sub>2</sub> [VP t<sub>1</sub> loves t<sub>2</sub>]]]  
 b. [IP many boys<sub>1</sub>... [VP every girl<sub>2</sub> [VP t<sub>1</sub> love t<sub>2</sub>]]]
- (2<sub>SSO</sub>) a. [IP every girl<sub>2</sub> [IP a boy<sub>1</sub>..... [VP t<sub>2</sub> [VP t<sub>1</sub> loves t<sub>2</sub>]]]  
 b. [IP every girl<sub>2</sub> [IP many boys<sub>1</sub>... [VP t<sub>2</sub> [VP t<sub>1</sub> love t<sub>2</sub>]]]

Parallelism ensures that in ellipsis or phonological reduction environments two sentences will have isomorphic syntactic representations at LF. It is thus possible to create environments in which the sentences in (3) will have the sentences in (2) as their syntactic images. In such environments, the sentences in (2) can serve as a litmus test indicating the LF structure of the sentences in (3). The prediction of Scope-Economy is that in these environments the sentences in (2) will only have the semantic interpretation that results from the LF structures in  $(2_{SSO})$ , thus indicating that the sentences in (3) can only have the structures in  $(3_{SSO})$ . In other words, Economy predicts the Ellipsis Scope Generalization (ESG) in (4), which is exemplified in the ellipsis construction in (5-6).

- (4) **Ellipsis Scope Generalization (ESG):** A sentence that is semantically equivalent under its two scopal relations will disambiguate its image (in favor of  $\neg SSO$ ), even when this image does show a scopal ambiguity when it appears in isolation.
- (5) a. Some boy admires every teacher. Mary does, too. ( $\exists > \forall$ )  $*(\forall > \exists)$   
 b. Some boy admires every teacher. Every girl does, too. ( $\exists > \forall$ )  $*(\forall > \exists)$
- (6) a. Some boy admires every teacher. Some girl does, too. ( $\exists > \forall$ ) ( $\forall > \exists$ )  
 b. Some boy admires every teacher. Many girls do, too. ( $\exists > \forall$ ) ( $\forall > \exists$ )

Assume that Scope-Economy is operative in grammar (i.e., that it plays a role in accounting for the ESG). There are still open questions regarding the way it is implemented, some of which may have empirical consequences. One such question concerns the “locality” of the computations that are involved in the implementation. For example, does Scope-Economy consult the outside systems to figure out the exact semantic interpretation of sentences that involve multiple quantification? I will make what I believe is

a more natural assumption. I will assume that Scope-Economy applies a local algorithm in which the only relevant aspect of semantic interpretation is the semantic commutativity of two quantifiers whose relative scope is being reversed. The basic idea is that an SSO (QR or QL) (which is not forced for independent reasons) can apply only if it reverses the scope of two quantifiers that are semantically non-commutative.

If this implementation is correct, there should be some empirical consequences. Specifically, under this implementation it is no longer the case that Scope-Economy allows an SSO to apply whenever it has semantic effects. Rather, Scope-Economy allows an SSO to apply only when it has semantic effects of a very local type. Suppose that QR or QL moves an XP in a way that does not affect the relative scope of two non-commutative quantifiers but nevertheless has semantic consequences. Suppose, for example, that it allows an XP to bind a variable in a constituent that is not present at the point at which the SSO applies or that it is a necessary step for a later application of an SSO. Scope-Economy predicts that the movement would be impossible. I will try to demonstrate that this prediction is borne out based on an investigation of coordination (section 2.3.) and the locality of QR (section 2.4.)

## 1.2. Economy and Variable binding (chapter 4)

Scope-Economy compares representations that are interpreted in an identical manner by the phonological and the semantic component. For this reason, the effects of Scope-Economy are not very transparent and their detection requires sophisticated tools. A natural question to ask is whether these tools can be utilized to uncover other economy conditions of a similar nature.

In chapter 4 I suggest that the answer is positive. Specifically, I argue for a generalization about the effects of ellipsis on variable binding which is parallel to the ESG. This generalization, which I call the Ellipsis Binding Generalization (EBG), is stated in (7).

- (7) **Ellipsis Binding Generalization (EBG):** A structure that is semantically equivalent under two variable binding relations will disambiguate its image (in favor of local variable binding) even when this image does show a variable binding ambiguity when it appears in isolation.

The EBG suggests that there is an Economy condition which prefers representations in which a variable is bound from the closest position available for a given interpretation.



- (8) **VB-Economy:** Non local variable binding can't be semantically vacuous; i.e. if a variable is bound non-locally, the representation must be semantically distinct from local variable binding.

The effects of this condition, which I will call VB-Economy, cannot be detected in normal circumstances. The reason for this is plain. Whenever two representation (with local and non-local variable binding) are compared by VB-Economy, they are interpretively identical and we cannot tell which is available. In other words, VB-Economy, just like Scope-Economy, compares “observationally identical” representations. However, we can utilize Parallelism to distinguish between the two representations.

To see the way the EBG is derived, consider the sentences in (9). In (9a) the sentence that contains ellipsis will receive an identical interpretation with local and non-local variable binding. VB-Economy, thus, rules-out non-local variable binding. Consequently, Parallelism rules-out non-local variable binding in the sentence that contains the antecedent VP. In (9b,c) non-local variable binding is semantically distinct from local variable binding both in the sentence that contains the antecedent VP and in the sentence that contains the elided VP. Non-local variable binding is licensed in both sentences and Parallelism is satisfied.

- (9) a. Every boy said that Mary likes his mother.  
       But Mary didn't \*<say that she likes her mother> . (ruled out by OV-Economy)
- b. Every person said that Mary likes his mother.  
    Only Mary (herself) didn't <say that she likes her mother>.
- c. Every boy said that Mary likes his mother.  
    Every girl didn't <say that Mary likes her mother>.

VB-Economy was suggested on independent grounds by Heim (1993). Modifying work by Reinhart (1983), Heim demonstrates that certain obviation of Binding Theory (BT) can be accounted for by the postulation of VB-Economy. Given the logic outlined above, we can say that Binding Theory, just like Parallelism, enables us to distinguish between (interpretively identical) representation, and in doing so it vindicates the prediction of VB-Economy.

Given sufficient evidence for VB-Economy, we might want to ask the question of implementation that we asked about Scope-Economy. Specifically, we might want to

figure out a (perhaps local) algorithm which the system might use in order to determine whether two representations are interpretively identical. At the moment I can't think of a local statement of the problem for variable binding which would be comparable to the statement that I offer for scope. I will tentatively suggest a way in which we might proceed to give a unified statement of the two economy conditions. However, at the moment I don't find the suggestion extremely compelling. What we have at the moment, I believe, are two economy conditions which have a similar conceptual flavor. However, they are not identical. I believe that Scope-Economy is a better understood condition. The definition of semantic identity which it utilizes is straightforward and lends itself to (fairly) straightforward implementations.

### **1.3. Binding Theory and the Representation of Scope (chapters 5-6)**

In chapter 2, Parallelism is utilized to test the predictions of Scope-Economy and in chapter 4, the same is done for VB-Economy only this time with Parallelism and BT. The obvious questions to ask is whether BT can be used to test Scope-Economy. In order to address this question, some preliminary issues need to be settled. First we need to find out whether BT is sensitive to the LF position of quantifier phrases. The conclusions of chapters 5-6 is that condition C of BT (BT(C)) applies no sooner than LF, and is therefore a potential detector for the position of a QP. However, it will also be argued that an A-bar chain is converted to an Operator Variable Construction under an Economy condition (OV-Economy) which (modulo interpretability) doesn't delete material from the copy at the tail of the chain. This, as we will see, makes it impossible to test the predictions of Scope-Economy based on BT(C). The case of BT(A), which I address in the end of chapter 6, is different. It seems that although normal cases of A-bar movement don't affect BT(C), they can affect BT(A). Although this fact is ill-understood, it can be utilized to test the predictions of Scope-Economy and once again the predictions are borne out.

### **1.4. BT(C) and Scope Reconstruction (chapter 5)**

In this chapter I argue that BT(C) is affected by Scope Reconstruction. Specifically, building on work by Heycock (1995) and Lebeaux (1989), I demonstrate that whenever

overt movement manages to obviate BT(C) it does so only if there is no constituent that dominates the relevant r-expression which undergoes Scope Reconstruction.<sup>7</sup>

**1.4.1. The nature of the argument:** Consider the various options for interpreting the output of movement:

(10)  $[x_p \dots r\text{-exp}_1 \dots]_2 \dots t_2$

As is well known, XP can take scope in the position of  $t_2$  (Scope Reconstruction) or in the landing-site (Non Scope-Reconstruction). Consider next what happens when movement bleeds BT(C):

(11)  $[x_p \dots r\text{-exp}_1 \dots]_2 \dots \text{pronoun}_1 \dots t_2$

If BT(C) is insensitive to the LF position of quantifiers we would expect that (11), just like (10), would have two scopal interpretations. I will demonstrate (through a variety of constructions) that that this is not the case and that in the configuration in (11) the availability of Scope Reconstruction is determined by ET(C):

(12) **The BT(C) and Scope Reconstruction correlation:** In the structural configuration in (11), an element that contains the r-expression cannot undergo Scope Reconstruction.

**This correlation argues that Scope Reconstruction feeds BT(C).**

For an illustration of the argument, consider the contrast between (13a) and (13b); A-movement Scope Reconstruction (QL) is possible only in the former. This contrast follows from BT(C) if we assume that this condition applies to the output of QL.

(13)    a. [A student of his<sub>i</sub>] seems to David<sub>j</sub> [t to be in the other room].  
                ( $\exists > seem$ ) (*seem*  $> \exists$ )  
       b. [A student of David's<sub>j</sub>] seems to him<sub>i</sub> [t to be in the other room].  
                ( $\exists > seem$ ) ??(*seem*  $> \exists$ )

The contrast is more evident in (14) where the semantics of the construction requires Scope Reconstruction:

<sup>7</sup> Romero (1997) has independently come up with many of the arguments in chapter 5.

- (14) a. For these issues to be clarified,  
[Many more/new papers about his<sub>i</sub> philosophy] seem to Quine<sub>i</sub> [t to be needed].  
a. #For these issues to be clarified,  
[Many more/new papers about Quine's<sub>i</sub> philosophy] seem to him<sub>i</sub> [t to be needed].

Another illustration of the argument can be seen by the following contrast, due to Lebeaux (1989), from the domain of A-bar Scope Reconstruction:

- (15) a. [The papers that he<sub>i</sub> gave to Ms. Brown<sub>j</sub>]  
every student<sub>i</sub> hoped [<sub>CP</sub> t' that she<sub>j</sub> will read t].  
b. \*[The papers that he<sub>i</sub> gave to Ms. Brown<sub>j</sub>]  
she<sub>j</sub> hoped [<sub>CP</sub> t' that every student<sub>i</sub> will revise t].

Because a variable must be in the scope of its binder, both (15a) and (15b) require some form of Scope Reconstruction. The contrast between the two cases follows from BT(C). In (15b), Scope Reconstruction yields a BT(C) effect. In (15a), by contrast, the moved constituent can “reconstruct” to the intermediate trace, t', where the variable is bound and BT(C) is not violated. The examples in (15), thus, illustrates that A-bar Scope Reconstruction is possible only if it doesn't yield a BT(C) effect.

The feeding relationship which account for (13-15), and for many other cases which I discuss, argues strongly that BT(C) is sensitive to the LF position of quantifiers.<sup>8</sup>

**1.4.2. Consequences:** The correlation between BT(C) and Scope Reconstruction argues that BT(C) applies at the interface with the semantic component (and perhaps beyond). This consequence, in turn, suggests that BT(C) can serve as a powerful tool for distinguishing among various claims regarding the nature of LF and the inventory of semantic mechanisms. In Chapter 6, I will utilize this tool to investigate the way in which the output of QR is converted to an operator variable construction. However, already the discussion of the correlation between Scope Reconstruction and BT(C) allows us to draw a few conclusions.

---

<sup>8</sup> One of the more compelling cases I discuss comes up in chapter 6 (6.5.2). After introducing certain assumptions relating to the interpretation of A-bar chains, it turns out that such constructions allow for “partial reconstruction”. When there is partial scope reconstruction, the predicted correlation is that BT(C) will be affected only if partial reconstruction targets a constituents that includes the relevant r-expression.

The fact that Scope Reconstruction affects BT(C) argues strongly that Scope Reconstruction is the result of “literal” reconstruction. In other words, I believe that there is a strong argument against the assumption that the semantic type of traces is sufficiently flexible to yield “Scope Reconstruction” and “Non Scope-Reconstruction” from the same syntactic structure.<sup>9</sup> This, I believe, is an important result. What we learn from the correlation between BT(C) and Scope Reconstruction is that BT(C) has the potential of distinguishing what would otherwise appear to be notational variants. Specifically what we learn is that (in normal cases) traces, like pronouns, must be interpreted as variables that range over individuals (type *e*).<sup>10,11</sup>

As it turns out, the interaction between BT(C) and Scope Reconstruction could also be used to investigate certain properties of LF representations. For example, the well formedness of sentences such as (15a) indicates that the position of an intermediate trace is available for Scope Reconstruction. This, if true, can be utilized to study the nature of successive cyclic movement. For example, the fact the contrast in (15) resurfaces in (16-17) indicates that Spec CP is not the only landing site available for successive cyclic A-bar movement.

- (16) a. [The papers that he<sub>i</sub> wrote for Ms. Brown<sub>j</sub>]  
every student<sub>i</sub> [<sub>VP</sub> t' asked her<sub>j</sub> to grade t].  
b. \*[The papers that he<sub>i</sub> wrote for Ms. Brown<sub>j</sub>]  
she<sub>j</sub> [<sub>VP</sub> t' asked every student<sub>i</sub> to revise t].
- (17) a. [At the time he<sub>i</sub> saw Ms. Brown<sub>j</sub>]  
I introduced every student<sub>i</sub> [<sub>XP</sub> t' to her<sub>j</sub> t]?  
b. \*[At the time he<sub>i</sub> saw Ms. Brown<sub>j</sub>]  
I introduced her<sub>j</sub> [<sub>XP</sub> t' to every student<sub>i</sub> t]

(16) indicates that VP adjunction is a possible landing site and (17) indicates the same for another maximal projection. A plausible conclusion, it seems to me, is that intermediate

---

<sup>9</sup> This assumption was made (among others) in Chierchia (1995) and Sternefeld (1997).

<sup>10</sup> As we will see shortly, I assume that A-bar traces are not interpreted as variables but rather as constituents that include a variable (of type *e*) and a restrictor.

<sup>11</sup> Why this should be the case is a very interesting question. Why, for example, can't traces receive the semantic type of generalized quantifiers and yield the semantic effects of Scope Reconstruction with no syntactic reflex that could affect BT(C)? I will not provide an answer to this question but I will raise various possibilities and refer the reader to an interesting proposal made (in a different context) by Beck (1996).

adjunction is in principle possible to every maximal projection (For some discussion see Fox (1998b), Nissenbaum (1998), Richards (1997) and references therein.)

### 1.5. BT(C) and QR (chapter 6)

Consider the structural configurations in (18) and (19), where linear precedence represents structural prominence.

- (18) a. [<sub>DP<sub>i</sub></sub> A boy] ....pronoun<sub>1</sub>.... [<sub>DP<sub>j</sub></sub> every... r-exp<sub>1</sub>...]  
       b. [<sub>DP<sub>i</sub></sub> Many boys] ....pronoun<sub>1</sub>.... [<sub>DP<sub>j</sub></sub> every... r-exp<sub>1</sub>...]
- (19) a. [<sub>DP<sub>i</sub></sub> John]....pronoun<sub>1</sub>.... [<sub>DP<sub>j</sub></sub> every... r-exp<sub>1</sub>...]  
       b. [<sub>DP<sub>i</sub></sub> Every boy]....pronoun<sub>1</sub>.... [<sub>DP<sub>j</sub></sub> every... r-exp<sub>1</sub>...]

In (18), QR of DP<sub>j</sub> over DP<sub>i</sub> is possible whereas in (19) it should be ruled-out by Scope-Economy. If this is the case, there is an obvious potential prediction. It is a potential prediction that in (18) BT(C) would be obviated via inverse scope whereas in (19) there would be a condition C violation. Unfortunately, this prediction is not borne out. Sentences such as (18) are bad independently of the relative scope of DP<sub>j</sub> and DP<sub>i</sub>. This, of course, doesn't cast doubt on the validity of Scope-Economy (since Scope-Economy does not restrict the sentences in (18)). Rather, it shows that BT(C) cannot serve as a diagnostic for QR. The obvious question is why.

**1.5.1. The proposal:** One possibility is to assume with Chomsky (1981) that BT(C) applies before QR. However, such an assumption is suspicious for architectural reasons (Chomsky 1993) and is very dubious given the correlation between BT(C) and Scope Reconstruction presented in chapter 5. The discussion in Chomsky (1993) provides us with a way out. Specifically, Chomsky suggests that, contrary to appearances, BT(C) applies after QR. The unacceptability of the configuration in (18) (even when an SSO has applied), he suggests, follows from a more general observation, namely that A-bar movement is, under normal circumstances, unable to bleed BT(C). The explanation that Chomsky provides for this general observation is based on the copy theory of movement and an economy condition on operator variable formation (OV-Economy).

The aim of chapter 6 is to suggest a slight modification of Chomsky's proposal (which would make a general theory of the semantics of A-bar chains more feasible) and to provide new evidence for it based on Antecedent Contained Deletion (ACD). The new

evidence is based on the observation that a natural formulation of OV-Economy predicts that in ACD constructions QR would bleed BT(C). I demonstrate, building on certain observations made in Fiengo and May (1994), that this prediction is correct. Furthermore, I extend Fiengo and May's observations and establish that the length of QR which is needed for ACD resolution coincides with the domain in which BT(C) is obviated. This result argues that Chomsky's proposal is not only architecturally parsimonious but actually needed on empirical grounds.

To appreciate the nature of the argument consider the contrast in (20).

- (20) a. ??I reported him<sub>i</sub> to [every cop that John<sub>i</sub> was afraid of t]  
 b. I reported him<sub>i</sub> to [every cop that John<sub>i</sub> was afraid I would]

Assume (for concreteness) that every QP must undergo short QR (adjunction to VP) for type reasons.<sup>12</sup> The output of QR (assuming the copy theory of movement) is given in (20').

- (20') a. ??I<sub>i</sub> [every cop that John<sub>i</sub> was afraid of t]  
           [t<sub>i</sub> reported him<sub>i</sub> to [every cop that John<sub>i</sub> was afraid of t]]  
 b. I<sub>i</sub> [every cop that John<sub>i</sub> was afraid I would]  
           [t<sub>i</sub> reported him<sub>i</sub> to [every cop that John<sub>i</sub> was afraid I would]]

The structure must now be converted to an Operator Variable Construction. I assume that this conversion is governed by an Economy condition along the lines of Chomsky (1993). This Economy condition (OV-Economy) states (informally) that the resulting Operator Variable Construction is as close as possible (up to interpretability) to the output of QR:

- (21) **OV-Economy:** When an Operator Variable Construction, *OV*, is formed from a chain,  $\alpha$ , the structure of *OV* is the closest structure to  $\alpha$  which is interpretable.

I will assume in the spirit of Chomsky (1993) that in principle a chain can be converted to two interpretable Operator Variable Constructions. One contains a copy of the restrictor of the quantifier at the tail of the chain and one doesn't. Under normal circumstances OV-Economy prefers the former, and this preference accounts for the fact that (under normal

---

<sup>12</sup>This is an assumption that I make following Heim and Kratzer (1998). However, it is possible to reformulate almost everything I say without this assumption. Nevertheless, I believe that the assumption is motivated on independent grounds (see chapter 2 note 5).







## Part I: Interpretation-sensitive Economy

### Chapter 2: Economy and Scope

It is well known that for certain sentences which involve more than one quantificational expression, there is semantic evidence for scope ambiguity. In the sentences in (1), for example, we can use semantic judgments (judgments about truth, entailment, etc.) to detect an ambiguity which we then characterized in terms of scope. Under such a characterization, one interpretation results from a scopal configuration in which the subject DP takes scope over the quantificational expression that it c-commands in Surface Structure ( $\exists > \forall$ ; *many*  $> \forall$ ;  $\exists > \textit{likely}$ ). In the other interpretation, the relative scope of the two quantificational expressions is reversed ( $\forall > \exists$ ; *many*  $> \forall$ ; *likely*  $> \exists$ ). I will call the former interpretation Surface Scope and the latter Inverse Scope.

- |     |  |                               |                               |
|-----|--|-------------------------------|-------------------------------|
| (1) | a. A student admires every teacher.                  | $(\exists > \forall)$         | $(\forall > \exists)$         |
|     | b. Many students admire every teacher.               | $(\textit{many} > \forall)$   | $(\forall > \textit{many})$   |
|     | c. Someone from NY is likely [t to win the lottery]. | $(\exists > \textit{likely})$ | $(\textit{likely} > \exists)$ |

There are also sentences for which there is semantic evidence for lack of scope ambiguity. The contrast between sentences of this sort and sentences such as those in (1) has figured prominently in various attempts to discover the mechanisms responsible for scope determination. (See, for example, the papers in Szabolsci 1997.) However, there is also a similarity between the two types of sentences. Both types of sentences have semantic properties that *reveal* the scopal relationships between the quantificational expressions that they contain. In this sense, the sentences of both types are Scopally Informative. This chapter brings to the fore a third set of sentences. In these sentences (which I will call Scopally Uninformative) there is no semantic evidence that might reveal the relative scope of two quantificational expressions:

- |     |   |                                 |
|-----|---|---------------------------------|
| (2) | a. Every student admires every teacher.   | <i>(Scopally Uninformative)</i> |
|     | b. John admires every teacher.            | <i>(Scopally Uninformative)</i> |
|     | c. John is likely [t to win the lottery]. | <i>(Scopally Uninformative)</i> |

Scopally Uninformative sentences are semantically equivalent under their different scopal relationships. For such sentences, it is impossible to use intuitions about meaning as

evidence for determining scopal relationships (though see note 62). This, of course, does not mean that scopal relationships fail to exist in such sentences. Nevertheless, uncovering the scopal relationships can not rely on simple judgments of truth, entailment and the like. The aim of this chapter is to suggest alternative methods that will allow us to investigate these relationships. Specifically, I will identify various mechanisms that are sensitive to the distinction between Surface and Inverse Scope and use them to determine the scopal properties of Scopally Uninformative sentences. This investigation will reveal the generalization in (3).

- (3) Scopally Uninformative sentences are restricted to Surface Scope.

If this generalization is correct, we would like to understand its source. Assume that Inverse Scope is the result of some operation (or set of operations) which we can call *OP*, to keep at this point to a theory neutral term. What (3) tells us is that *OP* cannot be semantically vacuous. This, if correct, lends support to an idea that has been very influential in recent work: the idea that *in general* operations in the linguistic system cannot be vacuous (in terms which need to be made precise). Furthermore, (3) suggests how the general idea should be articulated in the case of *OP*, namely that in this particular case vacuity is to be measured in semantic terms:<sup>1</sup>

- (4) **Economy condition on Scope Shifting (Scope-Economy):**<sup>2</sup> *OP* can apply only if it affects semantic interpretation (i.e., only if Inverse Scope and Surface Scope are semantically distinct).<sup>3</sup>

This much can be extracted from what is to follow independent of specific assumptions about the nature of *OP*. In other words, the argument that *OP* cannot be semantically vacuous will go through irrespective of whether or not *OP* is a movement

<sup>1</sup> For other operations, vacuity seems to be measured in other terms. For NP movement for example vacuity might be measured in terms of the checking of certain uninterpretable features. For extraposition and scrambling vacuity might be measured in phonological terms. See section 2.6.

<sup>2</sup> This Economy condition was suggested independently in Fox (1995a) and Reinhart (1994; 1997). Reinhart's argument for the Economy condition, however, were not based on Scopally Uninformative sentences. The arguments were based on certain assumptions about WCO. Unfortunately, these assumptions conflict with some of the arguments for the generalization in (3). See note 23.

<sup>3</sup> This idea will be qualified shortly. Specifically, the relevant notion of "semantic interpretation" will be limited to properties of "logical-syntax".

operation, and (assuming that it is a movement operation) the argument is independent of the precise nature of the operation. However, there is a general picture regarding *OP* which I believe is very well-motivated. Specifically, I believe that there are good reasons to believe that Inverse Scope is the result of an *optional* operations of Quantifier Raising or Quantifier Lowering/Reconstruction. I will, therefore, state what follows assuming this general picture with the hope that certain interaction between Scope-Economy and other well-motivated assumptions will yield interesting results.<sup>4</sup>

In the next section, I will re-state Scope-Economy in the context of certain background assumptions about the nature of *OP*. Once this is done, I will characterize the nature of the predictions that Scope-Economy makes, and will outline the structure of the chapter.

## 2.1. Scope-Economy, Quantifier Raising and Quantifier Lowering

There is widespread agreement that Inverse Scope is the result of an operation that reverses the relative scope of two quantificational expression. I will assume, following Chomsky (1976) and May (1977), that the relevant operation is a form of covert movement. Furthermore, I will assume, following May (1977), that covert movement (which is scope related) comes in two varieties. One variety involves Quantifier Raising (QR) and the other involves Quantifier Lowering (QL). (I will sometimes use the term Scope Shifting Operation (SSO) to refer to both types of movement.)

However, unlike May (1977;1985), I will not assume that a quantificational phrase must always be affected by QR. Rather, I will adopt the standard assumption that a QP is a second order predicate (type  $\langle et, t \rangle$ ). Therefore, I will assume with Heim and Kratzer (1998) that when a QP is the sister of a one place predicate (either by base generation [e.g.

---

<sup>4</sup> The predicted interactions between Scope-Economy and Parallelism (2.2.) are independent of the assumptions that I make regarding the nature of *OP* and can easily be translated to the terminology that is used when alternative theoretical assumptions are adopted. (See Jacobson (1997).) The predicted interactions between Scope-Economy and constraints such as the CSC (2.3.), Shortest Move (2.4.) and BT (section 6.7.) depend on the underlying assumption that *OP* is a movement operation. To the extent that the predictions are borne out, they lend support to the underlying assumption.

Furthermore, I believe that there are very strong independent arguments for an account of Inverse Scope that involves movement, most notably from Antecedent Contained Deletion (ACD) and from considerations of locality. These arguments will be reinforced in chapters 6 where I consider the possible affects of covert movement on Binding Theory.

as subject of a transitive verb or as object of an unaccusative verb] or due to a movement operation other than QR) it need not be affected by QR. When a QP is not a sister of a one place predicate, I assume (again, with Heim and Kratzer) that it must move by QR to a clause denoting expression (type  $t$ ), and that movement forms a one place predicate (through  $\lambda$ -abstraction over the position of the trace).<sup>5</sup>

In order to predict the generalization in (3), I would like to add two economy conditions to the system described above. One condition (which I call Scope-Economy) blocks optional instances of SSOs from applying when they are semantically vacuous. This condition ensures that an SSO which moves a quantifier out of a position in which it can be interpreted, yields an interpretation which would be unavailable otherwise:

- (5) **Scope-Economy:** SSOs which are not forced for type-considerations must have a semantic effect.<sup>6</sup>

In order to derive the generalization in (3), it is necessary to ensure (among other things) that a QP in object position would not raise over the subject in Scopally Uninformative sentences. (5) is insufficient; an Object QP must move for type-considerations, and (5) does not ensure that this movement will not cross the subject position. I therefore propose to add a locality condition on QR, which I call Shortest Move, with the hope that it follows

---

<sup>5</sup> The last assumption is not crucial for anything I say in this chapter. However, I believe that it becomes crucial once ACD is taken into account. Some of the arguments for Scope-Economy which are presented in this chapter are based on the assumption that QR is not licensed in order to satisfy Parallelism (in contrast to the operation that converts the output of QR to an interpretable structure, see chapter 6). If this assumption is correct, the fact that ACD is possible in sentences such as *John stood near every boy that Mary did* lends empirical support to Heim and Kratzer's claim

. The availability of ACD where there is no scope ambiguity virtually forces the assumption that short QR is always motivated on independent grounds. (Heim and Kratzer provide us with these grounds.) This assumption is reinforced by the observation that ACD is impossible when the quantificational phrase is replaced by a referring expression which can be interpreted in situ (*\*John stood near Bill, who Mary didn't* Lasnik (1995). If Lasnik is correct, the acceptability of *John likes Bill who Mary doesn't* is due to object shift).

<sup>6</sup> Note that obligatory instances of QR always have semantic effects. (With QR the structures are interpretable; without QR they aren't.) It is therefore conceivable that Scope-Economy applies to obligatory QR (as well as to optional QR). In order to implement this possibility, we would like to see how the local algorithm for determining semantic non-vacuity would be stated (in a uniform way) for the obligatory and the optional instances of QR.

from general considerations of locality:

- (6) **Shortest Move:** QR must move a QP to the closest position in which it is interpretable. I.e., a QP must always move to the closest clause denoting element that dominates it.<sup>7,8</sup>

Scope-Economy, together with Shortest Move, yields the restriction on Scope Inversion stated in (3). To see this, let's begin with a discussion of sentences such as those in (7) and (8).

- (7) a. A boy loves every girl.  
b. Many boys love every girl.
- (8) a. John loves every girl.  
b. Every boy loves every girl.

In each of these sentences, the object QP is uninterpretable in its base position and must move by QR.<sup>9</sup> The first instance of QR is thus determined by type considerations and is not restricted by Scope-Economy. It is licensed in all of the sentences, but restricted by Shortest Move. Shortest Move determines that it must target the closest clause-denoting expression dominating the object, namely the VP:

- (7') Obligatory instance of QR:<sup>10</sup>  
a. [IP A boy<sub>1</sub>.....[VP every girl<sub>2</sub> [VP t<sub>1</sub> loves t<sub>2</sub>]]]  
b. [IP Many boys<sub>1</sub>.....[VP every girl<sub>2</sub> [VP t<sub>1</sub> love t<sub>2</sub>]]]
- (8') Obligatory instance of QR:

<sup>7</sup> At this point, I have no evidence that could indicate whether Shortest Move applies to QL. If it does, the definition should be modified accordingly (where "c-commands" is replaced with "c-commanded by").

<sup>8</sup> Shortest Move predicts that QR is extremely local, perhaps more local than it really is. If QR turns out to be less local than what Shortest Move predicts, this condition will have to be modified. What is crucial for the purposes of this chapter is that obligatory QR satisfies Shortest Move (as it is defined here), and that optional QR moves QP to the closest position in which it crosses the relevant scope bearing element (see section 2.4.).

<sup>9</sup> In these particular cases one might suggest to replace short QR with covert object shift. However, as pointed out by Kennedy (1997) and Johnson and Tomioka (in press), this cannot be a general solution.

<sup>10</sup> In chapter 6 I will argue (following Chomsky 1993) that both the obligatory and the optional instance of QR leave a copy at the tail of the chain which is partially interpreted in that position. This copy is irrelevant for the considerations in this chapter, and will thus be ignored.

- a.  $[_{IP} \text{John}_1 \dots [_{VP} \text{every girl}_2 [_{VP} t_1 \text{ loves } t_2]]]$
- b.  $[_{IP} \text{Every boy}_1 \dots [_{VP} \text{every girl}_2 [_{VP} t_1 \text{ loves } t_2]]]$

Scope-Economy is responsible for determining whether a second (optional) instance of an SSO is licensed. Both QR and QL need to be considered. Let's begin with QR. In the constructions in (7'), the optional instance of QR has semantic effects while in (8') it doesn't. Scope-Economy thus licenses only the former:

(7'') Optional instance of QR:<sup>11</sup>



- a.  $[_{IP} \text{every girl}_2 [_{IP} \text{A boy}_1 \dots [_{VP} t'_2 [_{VP} t_1 \text{ loves } t_2]]]]$
- b.  $[_{IP} \text{every girl}_2 [_{IP} \text{Many boys}_1 \dots [_{VP} t'_2 [_{VP} t_1 \text{ love } t_2]]]]$

(8'') Optional instance of QR:

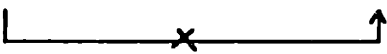
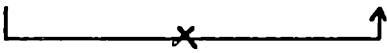
- a.  $*[_{IP} \text{every girl}_2 [_{IP} \text{John}_1 \dots [_{VP} t'_2 [_{VP} t_1 \text{ loves } t_2]]]]$  (*violates Scope-Economy*)
- b.  $*[_{IP} \text{every girl}_2 [_{IP} \text{Every boy}_1 \dots [_{VP} t'_2 [_{VP} t_1 \text{ loves } t_2]]]]$  (*violates Scope-Economy*)

Consider now the effects of Scope-Economy on QL. In contrast to QR, QL is never forced for type considerations. (The subject can always be interpreted in its surface position.) QL is totally optional and (by Scope-Economy) is allowed to apply only when it has semantic effects. For this reason QL, like QR, can yield Inverse Scope from (7'), but not from (8'):

(7''') Optional instance of QL:

- a.  $[_{IP} \dots [_{VP} \text{every girl}_2 [_{VP} \text{A boy}_1 \text{ loves } t_2]]]$   

- b.  $[_{IP} \dots [_{VP} \text{every girl}_2 [_{VP} \text{Many boys}_1 \text{ love } t_2]]]$   


(8''') Optional instance of QL:

- a.  $*[_{IP} \dots [_{VP} \text{every girl}_2 [_{VP} \text{John}_1 \text{ loves } t_2]]]$  (*violates Scope-Economy*)  

- b.  $*[_{IP} \dots [_{VP} \text{every girl}_2 [_{VP} \text{Every boy}_1 \text{ loves } t_2]]]$  (*violates Scope-Economy*)  


QL and QR provide two derivations for Inverse Scope in (7) and given Scope-Economy,

<sup>11</sup> In chapter 5-6, I will provide independent evidence that  $t'$  exists in QR and in other forms of A-bar movement.





for every model, and for every  $\phi \in D_{\langle e, \langle e, t \rangle \rangle}$ ,  
 $[[XP_1]](\lambda x [[XP_2]](\lambda y \phi(y)(x))) = [[XP_2]](\lambda y [[XP_1]](\lambda x \phi(y)(x)))$ .<sup>14</sup>

It might be plausible (for computational reasons) to assume that Scope-Economy applies a local algorithm in determining whether or not an SSO has semantic effects (though see Chomsky (1998) for cautionary remarks). If it does, the definition in (11) is a reasonable first step in specifying this algorithm.

Before we see the evidence for Scope-Economy, I would like to say a few words about the nature of this evidence. The evidence consists in part of arguments in favor of the generalization in (3). Establishing this generalization requires a method for distinguishing a representation with Inverse Scope from a representation with Surface Scope in cases where they are semantically identical (and of course phonologically identical as well). Distinguishing between such “observationally identical” representations is not a trivial matter. What I will attempt to do is to identify grammatical constraints which are sensitive to whether or not Inverse Scope is possible. In this chapter, the constraint that I will propose for this identification will be the widely discussed constraint on parallelism in constructions that involve phonological deletion or reduction (Parallelism). (In section 6.7., I will suggest that Condition A of Binding Theory can play a similar role.) This constraint will indicate that Inverse Scope is impossible in Scopally Uninformative sentences. As such, it will provide strong evidence for Scope-Economy.

The second type of evidence is based on certain assumptions about the structure of coordination. Under these assumptions, the relative scope of coordination and a quantificational phrase is invisible at the point at which an SSO applies. Therefore, scope relative to coordination cannot motivate an SSO.<sup>15</sup> However, once an SSO applies, it has

---

<sup>14</sup> For the relative scope of generalized quantifiers and other scope bearing elements like intensional verbs, modals and negation, we will need a more general definition of commutativity. For the purpose of the definition, we can take all of these scope bearing elements to be of semantic type  $\langle \tau, \tau \rangle$  (where  $\tau$  is the type of a proposition  $\langle s, t \rangle$ ), and for simplicity we will lift generalized quantifiers to type  $\langle e\tau, \tau \rangle$ .

$\langle O, XP \rangle$  is scopally commutative ( $[[O]] \in D_{\langle \tau, \tau \rangle}$  and  $[[XP]] \in D_{\langle e\tau, \tau \rangle}$ ) if for every model, and for every  $\phi \in D_{\langle e, \tau \rangle}$ ,  
 $[[O]]([[XP]](\lambda x \phi(x))) = [[XP]](\lambda x [[O]](\phi(x)))$ .

<sup>15</sup> Here (and elsewhere) I intend the word ‘motivation’ to be understood as a shorthand. The

semantic consequences. These consequences can be utilized to test the predictions of Scope-Economy.

The third type of evidence that I will provide is best understood under the implementation of Scope-Economy that I proposed in (11). Under this implementation it is no longer the case that an SSO can apply whenever it has semantic effects. Rather, Scope-Economy allows an SSO to apply only if it has semantic effects of a very specific (and local) type. Suppose that QR or QL have semantic consequences that do not depend on reversing the relative scope of two scopally non-commutative expressions. Suppose, for example, that the SSO is a necessary step for a later application of an operation  $\alpha$  that has its own semantic consequences. In such a case, Scope-Economy predicts that the consequences of  $\alpha$  will be visible only if the SSO also reverses the relative scope of two scopally non-commutative expressions.

I will focus on one case of this sort. Specifically, I will argue that a first instance of QR can bring an XP to the domain in which it can be affected by a second instance of QR (given considerations of locality, e.g., Shortest Move) only if the first instance (on its own) reverses the relative scope of two scopally non-commutative expressions.<sup>16</sup>

Assuming Sufficient evidence for Scope-Economy, there would still be open questions regarding implementation. Among others, we would like to know how the

---

claim that an SSO *is motivated* is to be understood as a shorthand for the claim that it satisfies the relevant economy conditions, i.e. that it crosses a scopally non-commutative expression satisfying Shortest Move.

<sup>16</sup> In Fox (1995a), I suggest that there is another case of this sort. Specifically I suggest that QL can bring an XP into the domain of existential closure (Heim (1982), Kamp (1981)) only if (independently) QL reverses the relative scope of two scopally non-commutative expression:

- (i)    a. Jewish woman are related to Chomsky.                      *False*
- b. Jewish woman are related to every Jewish man.    *True* but only under Inverse Scope

However, there was a problem in that suggestion. If existential closure is relevant for the interpretation of indefinites, then indefinites are not quantificational, and shouldn't be affected by scope-shifting operations. (Thanks to Gennaro Chierchia for discussion of related issues.)

What I would like to suggest is that bare plurals are existential QPs which (because of an ill-understood constraint, C) cannot be interpreted existentially in SPEC IP (see von Stechow 1995). In (ib), in contrast to (ia), Scope-Economy licenses QL and thus allows satisfaction of C.

In a sentence involving a Stage level predicate, the subject can undergo QL in order to have its time argument bound by a quantifier (over events) which is absent with individual level predicates (a possibility suggested to me by Irene Heim, cf. Musan 1995).

cognitive system in which an SSO applies (Syntax) determines whether or not an application of an SSO is semantically vacuous. The answer to this question bears on issues that have broad theoretical implications. One of these issues has to do with modularity. Is it the case that syntax can “see” all of what we might (pre-theoretically) call the interpretive properties of linguistic expressions?

This doesn’t seem very plausible. Instead, I will argue that there is a very narrow class of formal logical properties that certain words have and that these properties alone are accessible to syntax. These properties, which we could call properties of “logical-syntax”, are the properties which determine whether or not an SSO can apply. What I will suggest (tentatively) is that the linguistic system contains a deductive system with various formal rules of inference which can “prove” logical equivalence in various cases. When logical equivalence is proven, an SSO is blocked.

The remainder of this chapter has the following structure. In section 2.2., I discuss certain well-known puzzles that arise from the scopal interaction of quantifiers in VP ellipsis contexts. I go over some previous unsatisfactory attempts at an explanation and present data (mostly from Fox 1995a, but also some new data) that will motivate an explanation in terms of Scope-Economy. The explanation, as will be seen shortly, will be based on the idea that Parallelism can detect scopal relationships in Scopally Uninformative sentences. In section 2.3., I will present data that have to do with the scope of quantifiers in coordinated structures. These data will follow from an interaction of Scope-Economy and certain assumptions about the structure of coordination. Section 2.4. will be devoted to a proposed account for the locality of QR which is a necessary consequence of Scope-Economy and Shortest Move. In section 2.5., I will discuss broader potential ramifications of Scope-Economy. Specifically, I will discuss ramifications for modularity, and introduce the suggestion that grammar contains a deductive system. Finally in section 2.6., I will speculate on the reasons why SSOs, in contrast to other operations, are subject to Scope-Economy.

## **2.2. Scope-Economy and Parallelism**

In this section, I will use Parallelism as a detector of scopal relationships in Scopally Uninformative sentences. Parallelism ensures that in certain environments (environments that involve phonological reduction or deletion (ellipsis)) two sentences will receive isomorphic syntactic representations at LF. In such environments, one sentence,  $S_1$ , can

indicate the LF structure of another,  $S_2$ . Specifically (for scope), if  $S_1$  is Scopally Informative it can inform us of the scopal relationships in  $S_2$ , even if the latter is Scopally Uninformative. More specifically, if  $S_2$  is Scopally Uninformative and if (as predicted by Scope-Economy and Shortest Move) it is restricted to Surface Scope, we can detect this restriction by looking at  $S_1$ . The result of a variety of experiments based on this logic will indicate that, indeed, Scopally Uninformative sentences are restricted to Surface Scope.

I begin this section with a discussion of puzzling facts that emerge from work by Williams (1977), Sag (1976) and Hirschbühler (1982). I then show that if we assume Scope-Economy (and consequently assume that Scopally Uninformative sentences are restricted to Surface Scope) the facts are no longer puzzling. They are interpreted as involving a simple case in which a Scopally Informative sentence indicates that Scopally Uninformative sentences are confined to Surface Scope. I then move on to show that the logic described above makes very intricate predictions, and finally, I attempt to demonstrate that the predictions are in fact borne out.

**2.2.1. A Puzzle:** It is widely known that ellipsis sometimes disambiguates constructions involving multiple quantification. More specifically, a construction involving multiple quantification, which would normally show scopal ambiguity, sometimes loses this ambiguity when it serves as an antecedent for ellipsis.

Consider the fact exemplified in (12-13). The scopal ambiguity in a sentence such as (12) disappears when the sentence is an antecedent for VP ellipsis as in (13). For the first sentence in (13) to be true, there must be a single boy that admires all of the teachers, whereas in (12) the boys can vary with the teachers.

- |      |  |                       |                        |
|------|--|-----------------------|------------------------|
| (12) | A boy admires every teacher.                 | $(\exists > \forall)$ | $(\forall > \exists)$  |
| (13) | A boy admires every teacher. Mary does, too. | $(\exists > \forall)$ | $*(\forall > \exists)$ |

This phenomenon was already noted in Sag (1976) and Williams (1977). However, the accounts that Sag and Williams proposed, although very elegant, were both based on a mistaken factual assumption. Specifically, both Sag and Williams assumed that ellipsis will always bar Inverse Scope in a sentence of the type in (12). However, Hirschbühler (1982) showed that this assumption is incorrect. Despite ellipsis, the sentences in (14) allow the subject to receive narrow scope relative to a QP that it c-commands at surface structure.

- (14) a. A Canadian flag is in front of every building, and an American flag is too.  
 b. A Canadian flag is in front of many buildings and an American flag is too.  
 c. A Canadian flag is in front of most buildings and an American flag is too.<sup>17</sup>  
 d. One guard is standing in front of every building and one policeman is too.  
 e. A boy admires every teacher. A girl does, too.

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

We therefore confront a puzzle: what is the difference between (13) and (14)? Why is Inverse Scope possible in (14) but not in (13)? Previous attempts to deal with this puzzle (e.g. Cormack (1984) and Diesing (1992)) assumed that the relevant difference between (13) and (14) has to do with the grammatical properties of the subject in the sentence that contains ellipsis (henceforth, the ellipsis-sentence). In (13) the subject of the ellipsis sentence is a referring expression and in (14), the subject is quantificational.

However, this assumption seems to be incorrect. Although the relevant difference does concern properties of the ellipsis sentence, it is not limited to properties of the subject alone. Rather, as we shall see, it relates to a semantic property of the ellipsis sentence as a whole. I will argue that the crucial difference concerns the fact that the ellipsis sentence in (13), but not in (14), is Scopally Uninformative. *Mary admires every teacher* is semantically identical under Surface and Inverse Scope, whereas its counterpart in (14), *A girl admires every teacher* isn't.

**2.2.2. Parallelism as a detector:** It is a well-known fact that constituents that involve phonological reduction or deletion require an antecedent with which they share various interpretive properties.<sup>18</sup> Among many other properties, the scope bearing elements in the antecedent sentence ( $\beta_A$ ) must receive parallel scope to the corresponding elements in the ellipsis/reduction sentence ( $\beta_E$ ). This is demonstrated by the sentences in (15). (*Small Italics* are used to indicate Phonological Reduction.)

- (15) a1. [ $\beta_A$  I introduced one of the boys to every teacher] and  
           [ $\beta_E$  Bill did, too].

<sup>17</sup> Examples (14b,c) are designed to rule out an analysis in which the object quantifier in the antecedent sentence has scope over both conjuncts (ATB scope). See Hirschbühler (1982), Fiengo and May (1994), and Fox (1995a).

<sup>18</sup> There are certain cases in which VP ellipsis doesn't have an antecedent and an antecedent therefore needs to be accommodated (See Johnson 1997 and references therein). Such cases do not affect the arguments made here. (For some relevant discussion see chapter 3.)

- a2. [ $\beta_A$  I introduced one of the boys to every teacher] and  
 [ $\beta_E$  Bill *introduced one of the boys to every teacher*, too].
- b1 [ $\beta_A$  One of the boys was introduced to every teacher], and  
 [ $\beta_E$  one of the girls was, too].
- c2 [ $\beta_A$  One of the boys was introduced to every teacher], and  
 [ $\beta_E$  one of the girls *was introduced to every teacher*, too].
- c1. [ $\beta_A$  An American runner seems to Bill to have won a gold medal], and  
 [ $\beta_E$  a Russian athlete does too].
- c2. [ $\beta_A$  An American runner seems to Bill to have won a gold medal], and  
 [ $\beta_E$  a Russian athlete *seems to Bill to have won a gold medal*, too].

All of the sentences in (15) are scopally ambiguous. However, the scopal relationships in  $\beta_A$  cannot be different from the scopal relationships in  $\beta_E$ . If one is interpreted with Surface Scope, so must the other; ambiguities do not multiply in ellipsis/reduction contexts.

Facts such as those in (15) reveal the existence of a principle which I have called Parallelism. In chapter 3, I will further investigate this principle. For the time being, I will state what I think comes close to a necessary consequence of Parallelism. In chapter 3, we will see that this statement is not totally accurate. However, it is good enough for the time being:

- (16) **Parallelism** (a consequence of): In an ellipsis/phonological-reduction construction the scopal relationship between the elements in  $\beta_A$  must be identical to the scopal relationship between the parallel elements in  $\beta_E$ .

It should also be noted that Parallelism cannot follow from operations such as LF copying (as suggested, e.g., in Williams 1977). There are two reasons for this. First, as noted in Lasnik (1972), Chomsky and Lasnik (1993) and Tancredi (1992), Parallelism holds in constructions that involve phonological reduction (for example, the (b) sentences in (15)). In phonological reduction constructions there is no LF copying and Parallelism must be independently postulated. Second, Inverse Scope in sentences such as those in (15) involves operations (QR and QL) which relate positions internal and external to the putatively copied material. It is far from clear how the Parallelism effects in these cases would follow from a copying algorithm. I will therefore assume that VP ellipsis involves phonological deletion which is subject to Parallelism at a semantic level. In chapter 3, I will discuss accounts of Parallelism in terms of an independently needed theory of Focus

(as in Tancredi (1992) and Rooth (1992)) and show that such accounts when combined with Scope-Economy yield additional predictions that are corroborated.

But now let us return to the puzzle that was presented in the previous section. The problem was to explain the contrast between the sentences in (13) and (14). Why does ellipsis restrict the sentences in (13) but not in (14) to Surface Scope? Given that Parallelism is independently needed, Scope-Economy provides the answer. To see this, compare (13) and (14e).

- (13) A boy admires every teacher. Mary does, too <admire every teacher>.  
 $(\exists > \forall)$   $*(\forall > \exists)$
- (14e) A boy admires every teacher. A girl does, too <admire every teacher>.  
 $(\exists > \forall)$   $(\forall > \exists)$

The relevant difference between the two constructions, I propose, is that in (13) the ellipsis sentence is Scopally Uninformative. Therefore, Scope-Economy restricts the ellipsis sentence in (13) to Surface Scope. Consequently, Parallelism blocks Inverse-Scope in the antecedent sentence. In (14e), the ellipsis sentence is Scopally Informative and is therefore unrestricted by Scope-Economy. Both the ellipsis and the antecedent sentence can receive Inverse Scope as long as Parallelism is maintained. The availability of Inverse Scope in the other constructions in (14) follows in an identical manner (since, as mentioned, in all cases the ellipsis and the antecedent sentence are Scopally Informative).

Given that Parallelism applies to phonological reduction as well as to phonological deletion, Scope-Economy predicts that the facts in (13) and (14) will not change when the deleted VPs are replaced by phonologically reduced VPs. This is demonstrated below.

- (17) A boy admires every teacher. Mary *admires every teacher*, too.  
 $(\exists > \forall)$   $*(\forall > \exists)$
- (18e) A boy admires every teacher. A girl *admires every teacher*, too.  
 $(\exists > \forall)$   $(\forall > \exists)$

The discussion that follows will be restricted to ellipsis, but this is done only for the sake of brevity.

By now it should be obvious that Parallelism can be used to investigate the scopal properties of Scopally Uninformative sentences. Suppose that S is such a sentence. Semantics cannot help us figure out the assignment of scope. However, Parallelism provides us with a sentence, S' (the syntactic image of S) which must be isomorphic to S.

Whenever S' is Scopally Informative we can use its scopal properties in order to figure out the scopal relationship in S. Scope-Economy predicts the following generalization.

- (19) **Ellipsis Scope Generalization (ESG):** A sentence, S, will disambiguate its syntactic image, S', (in favor of Surface Scope), whenever S is semantically equivalent under Surface and Inverse Scope (i.e. whenever S is Scopally Uninformative).<sup>19</sup>

Stated somewhat differently, the ESG is a predicted interaction of Scope-Economy and Parallelism. From Scope-Economy it follows that Inverse Scope is impossible in S if S is semantically equivalent under Surface and Inverse Scope. (Since Inverse Scope would involve a semantically vacuous SSO.) From Parallelism it follows that the LF of S' involves Inverse Scope iff the LF of S does. The ESG is derived from these two principles in the following way. If S is semantically distinct under Surface and Inverse Scope, Inverse Scope is licensed in S. In such a case, S does not block Inverse Scope in S'. If, on the other hand, S is semantically equivalent under Surface and Inverse Scope, then

- (a) Scope-Economy doesn't allow Inverse Scope in S. And consequently,
- (b) parallelism doesn't allow Inverse Scope in S'.<sup>20</sup>

The contrast between the sentences in (13) and (14) provides initial support for the ESG and (consequently) for Scope-Economy. However the ESG has a wider extent. If we can test and corroborate it to its full extent, we will have very strong support for our initial account of the contrast between (13) and (14), and (consequently) for Scope-Economy. The aim of the next sub-section is to provide this support.<sup>21</sup>

---

<sup>19</sup> In Fox (1995a), I assumed that the ESG is slightly different. Specifically, I assumed that S can disambiguate S' only if S' is the ellipsis sentence. In this chapter, as in Fox (1995a), I concentrate on cases in which S' is the ellipsis sentence. However, in contrast to Fox (1995a), I don't restrict myself to such cases. (See example (36).) In Chapter 3, I argue that despite certain examples that suggest otherwise, disambiguation goes in both directions.

<sup>20</sup> This prediction depends on the assumption that an SSO is never licensed in S in order to satisfy Parallelism (given that it has "already" applied in S'). For further discussion see Chapter 3.

<sup>21</sup> In Fox (1995a) I discuss apparent counter-examples to the ESG and argue that upon closer scrutiny they end up supporting the generalization. I will not go over these arguments here. Johnson and Lappin (1997) provide two counter-examples, which I cannot account for. I think that many ill-understood factors enter into the analysis of the two sentences they discuss (involving conjuncts like *one divides every prime number* which



**2.2.3. The ESG:** The ESG states that in ellipsis constructions the ellipsis sentence will disambiguate the antecedent sentence if and only if the ellipsis sentence is Scopally Uninformative (and vice versa, see note 19). I will start by showing that the generalizations holds for scopal interactions between two DPs (2.2.3.1). Then I will argue that it holds for scopal interactions between DPs and sentential operators which are accounted for by QR (2.2.3.2.) and QL (2.2.3.3.).

**2.2.3.1. The ESG and scopal interactions among DPs:** To see that the ESG holds for scopal interactions between two DPs, we will go over the paradigm in (20-35). This paradigm is constructed by minimally varying the semantic properties of the subject in the ellipsis sentence. Just as in (13) and (14), Inverse Scope is possible in the antecedent sentence only when the ellipsis sentence is Scopally Informative.

In the construction in (20), the antecedent sentence is restricted to Surface Scope. This provides corroboration for the ESG. A definite description and a universal quantifier are (in this case) scopally commutative. In other words, the ellipsis sentence in the constructions in (20) have the same meaning under Surface and Inverse Scope. For this reason the antecedent sentence is expected to be restricted to Surface Scope.

- (20) a. Someone in the audience knows the capital of every country. The lecturer does, too.  $(\exists > \forall) *(\forall > \exists)$   
 b. One of the film reviewers admires every movie. The organizer of the film festival does, too.  $(\exists > \forall) *(\forall > \exists)$

The next question to ask is whether there are environments in which a definite description and a universal quantifier are scopally non-commutative. If such environments exist, we should be able to form minimal pairs with the sentences in (20). Such minimal pairs would provide a pretty decisive test for the validity of the ESG. One possibility that comes to mind is to embed a variable inside the restrictor of the definite description. Such a variable could be bound by the universal quantifier only under Inverse Scope, and thus could make

---

is a “law-like” sentence about arithmetic, and *the prime-minister bears responsibility for every parliamentary office* which very likely involves incorporation of ar. indefinite quantifier *some responsibility* which might reverse the predictions of Scope-Economy). In any event, I don’t agree with their conclusion that given the counter-examples the proposal “should be abandoned”. I believe it would be more reasonable to try and investigate the properties of the counter-examples.

Surface and Inverse Scope semantically distinct.<sup>22</sup>

An obvious confound for the experiment is Weak Crossover (WCO). Sentences such as those in (21) are marginal to some speakers when the variable inside the subject is bound under Inverse Scope.

- (21) a. ??[Its<sub>i</sub> prime-minister] knows the capital of [every country]<sub>i</sub>.  
 b. ??[Its<sub>i</sub> producer] admires [every movie]<sub>i</sub>.

There are, however, environments in which WCO is very weak, perhaps non-existent. Although we don't understand what characterizes these environments, we can use them for our experimental purposes.<sup>23,24</sup>

- (22) a. [The person who produced it<sub>i</sub>] admires [every movie]<sub>i</sub>.  
 b. [The expert who was invited to talk about it<sub>i</sub>] knows the capital of [every

---

<sup>22</sup> The definition of Scopal-commutativity in (11) did not make reference to assignment functions and thus is not sensitive to variable binding. This, however, could be corrected with very little effort.

(11') **Scope-Economy:** An SSO can move  $XP_1$  from a position in which it is interpretable only if the movement crosses  $XP_2$  and  $\langle XP_1, XP_2 \rangle$  is not scopally commutative.

$\langle XP_1, XP_2 \rangle$  is scopally commutative (when both denote generalized quantifiers) if for every model,  $M$ , assignment function,  $g$ , and  $\phi \in D_{\langle e,et \rangle}$

$$[[XP_1]]^{M,g}(\lambda x [[XP_2]]^{M,g}(\lambda y \phi(y)(x))) = [[XP_2]]^{M,g}(\lambda y [[XP_1]]^{M,g}(\lambda x \phi(y)(x))).$$

<sup>23</sup> Reinhart's argument for Scope-Economy (Reinhart (1994, 1997)) are based on the idea that this condition can account for WCO under the assumption that variables are invisible to Scope-Economy. Unfortunately, this argument is inconsistent with the arguments for Scope-Economy based on (23-25), (28-29) and (57-58).

<sup>24</sup> In Fox (1995a) I didn't give enough thought to the possible effects of WCO. I used constructions such as (i) in which there is no WCO. However, the lack of WCO (as suggested to me by Kai von Steiger) makes the paradigm somewhat suspicious. As we can see from (ii), in which WCO re-emerges, the obviation of WCO depends on having the ellipsis sentence and the antecedent sentence within the same construction. This might suggest an ATB analysis (as suggested to me by Daniel Hardt; see note 17). An ATB analysis is very unlikely for the (b) cases in (23-24).

- (i) One of the film reviewers admires every movie<sub>i</sub>, and its<sub>i</sub> director/the director does, too.  
 (ii) One of the film reviewers admires every movie<sub>i</sub>. ??Its<sub>i</sub> director/The director does, too.

- country]<sub>1</sub>.  
 c. [His<sub>1</sub> tutor] is admired by [every boy]<sub>1</sub>.<sup>25</sup>

Consider now the pairs in (23-25). These pairs show that when the definite description inside the subject of the ellipsis sentence contains a variable which can be bound by a universal quantifier under Inverse Scope, the ellipsis sentence is not disambiguated. This is exactly what is expected under the ESG.

- (23) a. Someone in the audience knows the capital of every country. [The person who was invited to talk about it] does, too.  $(\exists > \forall) (\forall > \exists)$   
 b. Someone in the audience knows the capital of every country. [The person who was invited to talk about these countries] does, too.  $(\exists > \forall) *(\forall > \exists)$
- (24) a. One of the film reviewers admires every movie. [The person who produced it] does, too.  $(\exists > \forall) (\forall > \exists)$   
 b. One of the film reviewers admires every movie. [The person who produced the film festival] does, too.  $(\exists > \forall) *(\forall > \exists)$
- (25) a. At least one teacher is admired by every boy. His tutor is, too.  $(\exists > \forall) (\forall > \exists)$   
 b. At least one teacher is admired by every boy. The head master/My tutor is, too.  $(\exists > \forall) *(\forall > \exists)$

It is important to point out that there is a syntactic difference between the subjects of the ellipsis sentence in the (a) and the (b) sentences in (23-25). The difference is that in the (a) sentences the subjects of the ellipsis sentence contain a bound-variable. One might suggest that this difference could be utilized for an account that cares about the formal properties of the subject and is agnostic to scopal commutativity. However, this doesn't seem to be the case. To see this, consider (26).<sup>26</sup> In this construction, there is a variable within the subject of the ellipsis sentence. Nevertheless, the antecedent sentence is still disambiguated in favor of Surface Scope. The reason is that the variable in the subject of the ellipsis sentence cannot be bound by the object quantifier under Inverse Scope. For this reason, Inverse Scope is semantically identical to Surface Scope, and consequently the ellipsis sentence disambiguates the antecedent sentence.

- (26) Every producer<sub>1</sub> thinks that one film reviewer admires every movie and that his<sub>1</sub> wife does, too.  $(\exists > \forall) *(\forall > \exists)$

<sup>25</sup> The obviation of WCO in (a) and (b) seems to be related to the embedding of the pronoun. In (c) it is probably related to the existence of a trace in object position, perhaps below the position of *every boy*. See chapter 5, section 1.

<sup>26</sup> Thanks to an *NALS* reviewer for providing me with a sentence of this type.

Consider now (27). In this construction, the subject of the ellipsis sentence is a bona fide quantifier. Nonetheless, contrary to Diesing and to Cormack, the antecedent sentence is restricted to Surface Scope. The reason for this restriction is that the ellipsis sentence *every girl admires every teacher* is Scopally Uninformative.<sup>27</sup> (Two universal quantifiers are scopally commutative.)

- (27) A boy admires every teacher. Every girl does, too.  $(\exists > \forall) *(\forall > \exists)$

Now consider (28-29).

- (28) a. Someone in the audience knows the capital of every country. [Every person who was invited to talk about it] does, too.  $(\exists > \forall) (\forall > \exists)$   
 b. Someone in the audience knows the capital of every country. [Every person who was invited to talk about these countries] does, too.  $(\exists > \forall) *(\forall > \exists)$
- (29) a. One of the film reviewers admires every movie. [Every person who was involved in its production] does, too.  $(\exists > \forall) (\forall > \exists)$   
 b. One of the film reviewers admires every movie. [Every person was involved in this production] does, too.  $(\exists > \forall) *(\forall > \exists)$
- (30) a. At least one teacher is admired by every boy. [Every one of his tutors] is, too.  $(\exists > \forall) (\forall > \exists)$   
 b. At least one teacher is admired by every boy. [Everyone of my tutors] is, too.  $(\exists > \forall) *(\forall > \exists)$

The (b) constructions in (28-30) are similar to (27) in that the ellipsis sentence involves two universal quantifiers, and is thus Scopally Uninformative. The ellipsis sentence in the (a) constructions involves two universal quantifiers as well. However, in this case Inverse Scope (given that it enables a variable within the subject to be bound) allows the domain of quantification of the subject universal quantifier to be determined by the elements that the QRed expression quantifies over. Specifically, in (28a) Inverse Scope allows the invited people to vary with respect to the countries. Similarly, in (29a), it allows the production staff to vary with respect to the movies, and in (30a), it allows the tutors to vary with the students. In these constructions, there are semantic consequences to Inverse Scope in the ellipsis sentence, and, thus, Inverse Scope is possible in the antecedent sentence as well.

---

<sup>27</sup> I thank Kyle Johnson for suggesting this line of research.

The constructions in (31-32) differ minimally from (27) and from the (b) constructions in (28-29). The difference is that the semantic properties of the subject quantifier in the ellipsis sentence in (31-32) make the truth conditions different under the two different scopal relations.

- (31) a. One of the film critics admires every movie. Almost everyone in the film festival does, too.  $(\exists > \forall) \quad (\forall > \exists)$   
 b. One of the film critics admires every movie. Most visitors to the film festival do, too.  $(\exists > \forall) \quad (\forall > \exists)$   
 c. One of the film critics admires every movie. Many visitors do, too.  $(\exists > \forall) \quad (\forall > \exists)$
- (32) a. One student (in the school) knows the capital of every country. Almost every teacher (in the school) does too.  $(\exists > \forall) \quad (\forall > \exists)$   
 b. One student knows the capital of every country. Most teachers, do too.  $(\exists > \forall) \quad (\forall > \exists)$   
 c. One student knows the capital of every country. Many teachers, do too.  $(\exists > \forall) \quad (\forall > \exists)$

In the (a) constructions, the modifier *almost* makes the ellipsis sentence Scopally Informative. When the subject in the ellipsis sentence of (31a), for example, has wide scope, the sentence can be true only if there is a set of people in the film festival that is sufficient in cardinality to be considered “almost everyone”, and only if each member of this set admires all of the movies. However, when the object of (31a) has wide scope, the sentence can be true also when the set of admiring people vary with respect to the movies. In other words, the sentence can be true even when there is no single sufficiently large set of people that admires all of the movies. All that is required is that for each movie there would be a sufficiently large set of admirers. The ellipsis sentence is thus Scopally Informative and therefore does not block Inverse Scope in the Antecedent Sentence. The other sentences in (31-32) are parallel to (31a) and again the antecedent sentence is not restricted to Surface Scope.

Compare now the sentences in (33). (33a) is a Scopally Uninformative sentence. (33b), which is syntactically very close to (33a), is Scopally Informative (see Keenan and Faltz (1985)). (For the sentence to be true under Surface Scope ( $or > \forall$ ), there must be a way of choosing a person from the set containing John, Bill and Fred such that the chosen person likes every teacher; under Inverse Scope ( $\forall > or$ ) the choice can vary with the teachers.)

- (33) a. John, Bill and Fred like every teacher. (*Scopally Uninformative*)  
 b. (Either) John, Bill or Fred likes every teacher. (*or* >  $\forall$ ) ?( $\forall$  > *or*)

The speakers I consulted find Inverse Scope somewhat difficult in (33b). Nevertheless, the reading seems to be available. Scope-Economy predicts that (to the extent that Inverse Scope is available in (33b)) this sentence will contrast with (33a), in that only (33a) will disambiguate its image under ellipsis. This prediction seems to be correct:

- (34) a. A girl admires every teacher. John, Bill and Fred do, too. ( $\exists$  >  $\forall$ ) \*( $\forall$  >  $\exists$ )  
 b. A girl admires every teacher. (Either) John, Bill or Fred does, too. ( $\exists$  >  $\forall$ ) ?( $\forall$  >  $\exists$ )
- (35) a. A girl admires every teacher. Both John and Bill do, too. ( $\exists$  >  $\forall$ ) \*( $\forall$  >  $\exists$ )  
 b. A girl admires every teacher. Either John or Bill does, too. ( $\exists$  >  $\forall$ ) ?( $\forall$  >  $\exists$ )

To summarize, (20-35) demonstrate that the antecedent sentence in a VP ellipsis construction can receive Inverse Scope only if the ellipsis sentence is semantically distinct under Surface and Inverse Scope. To make this point clear, I list the sentences that disambiguate the antecedent sentence in (I) and those that do not in (II). The difference is that in (I) a Scope Shifting Operation does not affect truth conditions, while in (II) it does.

- (I) a. Mary likes every teacher.  
 b. The person who produced the film festival admires every movie.  
 c. Every person who was involved in this production admires every movie.  
 d. John and Bill like every teacher.  
*(Scopally Uninformative)*
- (II) a. Some girl likes every teacher.  
 b. The expert who was invited to talk about it knows the capital of every country.  
 c. Every person who was involved in its production admires every movie.  
 d. Almost everyone in the film festival admires every movie.  
 e. Many visitors to the film festival admire every movie  
 f. An American flag is in front of many buildings.  
 g. John or Bill like every teacher.  
*(Scopally Informative)*

What we've learned in this sub-section is that in ellipsis constructions the sentences in (I) restrict an otherwise ambiguous sentence,  $S_{\text{Scope-Indicator}}$ , to Surface Scope. This corroborates the predictions of Scope-Economy. Reversing the scope of the DPs in (I) would be semantically vacuous and is predicted by Scope-Economy to be impossible. In a

Parallelism environment,  $S_{\text{Scope-Indicator}}$  can be used to test this prediction. The results of the experiment confirm the prediction.

More specifically, because Inverse Scope in a sentence such as  $S_{\text{Scope-Indicator}}$  can result from two possible derivations (compare (8'') and (8''')), we conclude that Scope-Economy blocks both of these derivations for the Scopally Uninformative sentences in (I). This suggests that both optional QR and optional QL are subject to Scope-Economy. In the next two sub-sections, I will provide additional evidence that this is the case.

First, however, I would like to discuss additional support for the ESG which comes from pseudo-gapping.<sup>28</sup> Pseudo-gapping is a construction similar to VP ellipsis in its compliance to Parallelism. In this respect pseudo-gapping can, like ellipsis, be used to test Scope-Economy. However, pseudo-gapping is also different from VP ellipsis (and superficially similar to gapping) in that part of the VP is not elided.<sup>29</sup> This property of pseudo-gapping is important for our purposes because it allows us to vary the semantic properties of a sentence (which must comply to Parallelism) while keeping its subject constant.<sup>30</sup> Such an experiment can give the final blow to an account of the facts we've discussed based on properties of the subject in the ellipsis sentence.

Consider the instances of pseudo-gapping in (36a). In this sentence both the first and the second conjunct are Scopally Informative and neither is restricted to Surface Scope. The sentence thus conforms to the ESG. The sentence also exemplifies the fact that pseudo-gapping obeys parallelism. The relative scope of the two quantifier must be identical in the two conjuncts.<sup>31</sup> If the sentence allows the boys to vary with respect to the professors, it must also allow the girls to vary with respect to the parents. Similarly, if variance is disallowed for the boys it is also disallowed for the girls.

<sup>28</sup> I thank Martha McGinness for suggesting pseudo-gapping as a further test for the validity of the ESG.

<sup>29</sup> It seems reasonable to claim that pseudo-gapping involves movement of material out of the VP, followed by VP ellipsis. For analyses along these lines, see Jayaseelan (1990) and Lasnik (1995).

<sup>30</sup> Phonological down-stressing has the potential of serving the same role that pseudo-gapping does. If conjunctions involving phonological down-stressing in one of the conjuncts must obey Parallelism (see Lasnik (1972), Chomsky and Lasnik (1992), Tancredi (1992), Rooth (1992) as well as Chapter 3), then the absence of full ellipsis could allow us to vary lexical choices within the VP. However, I found that judgments with these constructions were too difficult. This is not surprising since the phonological realization of VP with a focused object and a non-focused verb is not necessarily distinct from the phonological realization of a focused VP.

<sup>31</sup> In Chapter 3, we will understand the relevant properties of Pseudo-Gapping which make it comply with Parallelism.









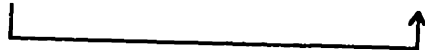
Consider the sentence in (40). This sentence is ambiguous with respect to the relative scope of the subject quantifier and of the attitude verb *seems*. If the subject quantifier has wide scope, the sentence can be true only if Bill has some American runner in mind, and only if it seems to Bill that that particular American runner won a gold medal. If the attitude verb has wide scope, the sentence merely requires that Bill have the belief that some American runner or other won a gold medal. Bill need not have any particular American runner in mind. The sentence would be true, for example, in a situation in which Bill sits in the Olympic cafeteria, hears the American anthem and concludes that an American runner won the medal.

- (40) An American runner seems to Bill to have won a gold medal.

$(\exists > \textit{seems})$  ( $\textit{seems} > \exists$ )

The two readings of the sentence in (40) are the result of the two positions in which the subject can be interpreted. If the subject is interpreted in its SS position, it has wide scope relative to the attitude verb. However, if it is lowered to the embedded IP at LF, the attitude verb has wide scope. (See, among others, May (1985) and Diesing (1992).) The two LFs for (40) are thus represented in (40').

- (40') a. An American runner<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal].  
 b.        seems to Bill [[An American runner] to have won a gold medal].



The obvious question to ask now is what happens when we embed (40) in a VP ellipsis construction. Let's start with (41), which shows the same ambiguity which we have seen in (40).

- (41) An American runner seems to Bill to have won a gold medal, and a Russian athlete

---

have to stipulate that all PPs move to a case position. Further, it is not clear how such approaches would account for cases of ambiguity other than those involving two arguments of a verb. For example, it is not clear how such approaches would deal with the scopal ambiguity of object quantifiers and heads such as modals, negation and attitude verbs (as in (38) above and in (64) below), or with cases of inverse linking. Yet another problem is that they cannot deal with certain cases in which scope is not clause-bound. (see Appendix, see, also Kennedy (1997)). Finally, as pointed out to me by Noam Chomsky, the data in section 2.3. indicate that Inverse Scope can result from two different derivations, as outlined in 2.1. (c.f. (7'', 7'''))

does, too.

$(\exists > \textit{seems}) (\textit{seems} > \exists)$

However, since Parallelism must be maintained, the relative scope of the subject and the attitude verb must be the same in the two conjuncts. Bill could know the identity of the Russian and the American runner, or alternatively he could be sitting in the cafeteria hearing the consecutive playing of the two anthems. What is important, however, is that if (for the sentences to be true) Bill is required to know the identity of the American runner, he must also know the identity of his Russian colleague and vice versa. (41'a,b) are possible LFs for (41) while (41'c,d) are impossible.

- (41') a. An American runner<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal] and a Russian athlete<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal].
- b. \_\_seems to Bill [[An American runner] to have won a gold medal] and \_\_seems to Bill [[a Russian athlete] to have won a gold medal].
- c. \* An American runner<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal] and \_\_seems to Bill [[a Russian athlete] to have won a gold medal].  
(violates Parallelism)
- d. \* \_\_seems to Bill [[An American runner] to have won a gold medal] and a Russian athlete<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal].  
(violates Parallelism)

Consider now (42). This sentence, contrary to (41), is restricted to Surface Scope. For the sentence to be true, Bill must know the identity of the American runner.

- (42) An American runner seems to Bill to have won a gold medal and Sergey does, too.  
 $(\exists > \textit{seems}) *(\textit{seems} > \exists)$

This follows naturally from the assumption that QL is restricted by Scope-Economy. If QL, just like QR, is subject to Scope-Economy, it can apply in a structure only if it has semantic consequences. Since in a sentence such as (43) the meaning remains the same whether or not QL takes place, QL is impossible. For the same reason QL cannot apply to the second conjunct in (42), and because of Parallelism it cannot apply to the first conjunct. (42'c,d) are ruled out by Parallelism and (42'b) is ruled out by Scope-Economy.

- (43) Sergey seems to Bill to have won a gold medal.

- (42') a. An American runner<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal] and  
Sergey<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal].
- b. \*\_\_seems to Bill [[An American runner] to have won a gold medal] and  
\_\_seems to Bill [[Sergey] to have won a gold medal].
- c. \*An American runner<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal] and  
\_\_seems to Bill [[a Russian athlete] to have won a gold medal].
- d. \*\_\_seems to Bill [[An American runner] to have won a gold medal] and  
Sergey<sub>1</sub> seems to Bill [t<sub>1</sub> to have won a gold medal].

### 2.3. Scope-Economy and coordination

Consider a situation in which movement of QP over  $\alpha$  (either by QR to a position above  $\alpha$  or by QL to a position below  $\alpha$ ; henceforth crossing  $\alpha$ ) has semantic effects. Suppose, now, that (under certain circumstances), the relevant semantic effects are “invisible” to Scope-Economy. What this means is that although crossing  $\alpha$  has semantic effects, it is not licensed by Scope-Economy. If such circumstances exist, Scope-Economy would make very intricate predictions. Specifically, Scope-Economy would predict that QP could cross  $\alpha$  only if the structure contains another scope bearing element,  $\beta$ , such that (a) QP is allowed by Scope-Economy to cross  $\beta$  (QP and  $\beta$  are scopally non-commutative) and (b) Crossing  $\beta$  necessarily involves crossing  $\alpha$ .

In this section, I will suggest that such a circumstance exists, and that the predictions of Scope-Economy are borne out. Specifically, I will suggest that semantic effects of movement across coordination are invisible to Scope-Economy, and that we therefore predict the following generalization.

- (44) **The Coordination Scope Generalization:**<sup>38</sup> An SSO can move a QP across coordination only if the structure contains a scope bearing element,  $\beta$ , such that
- (a) QP and  $\beta$  are scopally non-commutative.
  - (b) The SSO that moves QP across  $\beta$  also moves QP across coordination ( $\beta$  and QP are on different sides of the coordination.)

This generalization has two parts depending on whether the relevant SSO is QR or QL. In

---

<sup>38</sup> As will become clearer in the next sub-section, ‘movement across coordination’ refers to cases in which an XP moves (sometimes Across The board) from within a conjunct to a position that c-commands coordination (or vice versa in the case of lowering).

this section, I will provide evidence that both parts of the generalization are correct. Furthermore, I will argue that the invisibility of coordination follows from a multi-dimensional analysis of coordination which has been proposed to account for the Coordinate Structure Constraint (CSC). In order for the discussion to be more concrete, I will begin with a discussion of the multi-dimensional analysis.<sup>39</sup>

**2.3.1. The Multi-dimensional analysis of coordination:** In this sub-section, I will present assumptions that are shared by a family of accounts of the CSC. As we will see, these assumptions will yield the invisibility which was characterized somewhat abstractly above.

The effects of the CSC are exemplified in (45). The relevant generalization is that extraction out of a coordinate structure is possible only if it occurs across the board (ATB).<sup>40</sup>

- (45) a. \* Who do you think Mary likes *t* and Bill hates Sue?  
b. Who do you think Mary likes *t* and Bill hates *t*.

Certain accounts of this generalization are based on the two assumptions in (46).<sup>41</sup>

- (46) a. Extraction out of a coordinate structure is possible only when the structure consists of two independent sub-structures, each composed of one of the coordinates together with material above it up to the landing site (henceforth, Component Structures).  
b. Grammatical constraints are checked independently in each of the Component Structures.

---

<sup>39</sup> The discussion of the multi-dimensional analysis is provided for concreteness. Specifically, it is provided in order for us to be clear about the set of operators that are relevant in coordination for the licensing of an SSO. If one feels strong distaste for the multi-dimensional analyses, one can read this section as providing evidence for Scope-Economy based on an ill-understood invisibility assumption.

Another possibility that might be considered is to capitalize on the fact that a universal quantifier and a coordinator *and* are scopally non-commutative. However, this will raise problems for cases that involve downward entailing quantifiers, such as (62), or for the case of ATB lowering discussed in 2.3.2.

<sup>40</sup> There are certain counter-examples to the CSC due to Lakoff (1986). My hope is that the correct account of these examples will not affect the arguments presented in this section (see previous note).

<sup>41</sup> These assumptions are made most explicit (and, as far as I know, are most natural) under theories which analyze coordinate structures with the use of multi-dimensional phrase structures (cf. Goodall (1987), Muadz (1991) and Moltman (1992)).

With these assumptions, the contrast in (45) follows as a direct consequence of the constraint against vacuous quantification. When movement takes place out of one of the coordinates (and not the other), as in (45a), the Component Structures are those in (47). As we see, the second Component Structure involves vacuous quantification. When movement is ATB, both Component Structures are well-formed, as we can see in (48).<sup>42</sup>

- (47) Component Structures of (45a)  
 1. Who do you think Mary likes t  
 2. Who do you think Bill hates Sue
- (48) Component Structures of (45b)  
 1. Who do you think Mary likes t  
 2. Who do you think Bill hates t

Given (46b), Scope-Economy, like any other grammatical constraint, should be satisfied independently in each Component Structure. The prediction is that an SSO which moves a QP across coordination must be motivated internal to the Component Structures in which it occurs. This yields the invisibility which was necessary for deriving the Coordination Scope Generalization. More specifically, if a QP crosses coordination, it must cross some quantificational expression  $\beta$  (such that  $\beta$  and QP are scopally non-commutative) within the Component Structure(s) in which it applies. Global semantic effects that result from the complete structure (i.e. from the composition of the derivative components and the semantics of coordination) are invisible to Scope-Economy.

**2.3.2. Scope-Economy and coordination - the case of QR:** In this section, I will argue that the Coordination Scope Generalization holds of QR. The generalization as it applies to QR can be stated more explicitly in the following manner:

---

<sup>42</sup>Note that this account doesn't predict that ATB movement must take place in parallel. That is, it doesn't predict that ATB extraction from different structural positions would be ill-formed. I don't think that this is necessarily a bad result. Although ATB extraction tends to be better when it takes place from parallel positions, it is known that this is only a tendency. Non-parallel ATB extraction such as that in (i) are perfectly acceptable. (see Williams (1978))

- (i) John, who Mary invited t to the party and Bill said t would come late.


- (49) **The Coordination QR Generalization (CQRG):** In a structure such as (i), an optional instance of QR can move QP outside of the coordination only if there is some scope bearing element  $\beta$  c-commanding the coordination such that (a)  $\beta$  and QP are scopally non-commutative and (b) QR moves QP over  $\beta$ , as in (ii).

- (i)  $[_{YP} \dots [_{\alpha 1} \dots QP \dots]]$  and  $[_{\alpha 2} \dots]$   
 (ii)  $QP_x [_{YP} \beta \dots [_{\alpha 1} \dots X \dots]]$  and  $[_{\alpha 2} \dots]$

The CQRG follows straightforwardly from Scope-Economy under a multi-dimensional analysis. Given Scope-Economy, an optional instance of QR must move QP above a scopally non-commutative expression,  $\beta$ . Given the multi-dimensional analysis of coordination,  $\beta$  must be an element in the Component Structure in which QP applies ( $[_{YP} \dots [_{\alpha 1} \dots QP \dots]]$ ). Furthermore, given Shortest Move, QR will target the maximal projection that immediately dominates  $\beta$ . Therefore, the only way for QR to move QP over *and* is for it to move over  $\beta$  and for  $\beta$  to c-command  $\alpha 1$  (and hence c-command coordination as well).

However, before we test the CQRG, we should go over some background regarding the affects of the CSC on QR.

**2.3.2.1. Setting the stage:** May (1985) and Ruys (1993) have argued that QR obeys the CSC (see also Lakoff (1970) and Rodman (1976)). One of the arguments is based on the observation that a QP within a coordinated VP cannot out-scope the subject. Thus, the option for Inverse Scope in (50a) disappears in (50b).

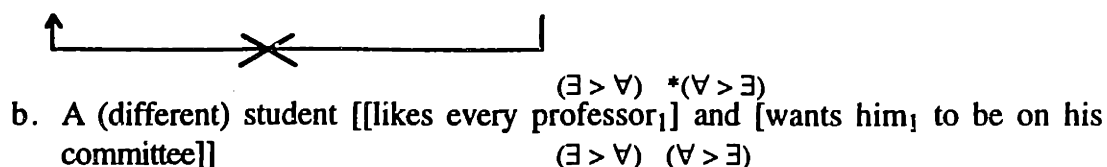
- (50) a. A (different) student likes every professor.  $(\exists > \forall)$   $(\forall > \exists)$   
 b. A (#different) student [[likes every professor] and [hates the dean]]  
  $(\exists > \forall)$   $*(\forall > \exists)$

This observation supports the claim that QR obeys the CSC, and, thereby supports the claim that QR involves syntactic movement.

Ruys (1993), however, has observed that a pronoun in one of two conjuncts can license QR out of the other in apparent violation of the CSC. This is demonstrated by the contrast in (51).

- (51) a. A (#different) student [[likes every professor] and [hates the dean]]





(51b), in contrast to (51a), allows the object to QR over the subject in apparent violation of the CSC. Ruys suggestion (stated in the terms of a multi-dimensional analysis) is that the pronoun in the second conjunct allows the quantifier to move out of the first conjunct without violating the constraint on vacuous quantification in any of the Component Structures. More specifically, he suggests that the pronoun is analyzed in LF as a resumptive pronoun. The two Component Structures are provided in (52), and neither involves vacuous quantification.

- (52)
1. Every professor<sub>1</sub>, a student likes t<sub>1</sub>.
  2. Every professor<sub>1</sub>, a student wants him<sub>1</sub> to be on his committee.

Although Ruys doesn't point this out explicitly, I think that his observation lends support to an explanation of the CSC along the lines discussed above. (51) shows that there is no requirement of ATB movement. The appearance of a requirement is an artifact of the two Component Structures which are checked independently for vacuous quantification. In (51b), where vacuous quantification is avoided by a resumptive pronoun, there is no need for ATB movement.

In (51b), movement takes place only out of one of the conjuncts. However, the principles from which the CSC follows do not require ATB movement. What they require is that each of the Component Structures obey all grammatical constraints, in particular the ban on vacuous quantification. Movement takes place only out of one of the Component Structures. However, the other Component Structure doesn't suffer from vacuous quantification due to the presence of the pronoun.

Ruys' suggestion is further supported by his observation that quantifiers behave like in situ Wh-phrases. (53a), just as (51a), involves vacuous quantification in one of the Component Structures and is thus ruled out. In (53b), just as in (51b), a resumptive pronoun obviates the CSC.<sup>43</sup>

<sup>43</sup> Note that this point does not depend on the claim that in situ WH-operators undergo LF

- (53) a. \*Which student likes which professor and hates the dean?  
 b. Which student likes which professor and wants him to be on his committee?

**2.3.2.2. Testing the generalization:** What we've learned in the previous section is that there is a restriction on QR which follows from vacuous quantification (the CSC) and that this restriction is independent of Scope-Economy. We can thus restate our expectation for coordination in (49') where a (resumptive) pronoun  $x$  is inserted into the second conjunct  $\alpha_2$  so as to nullify the effects of vacuous quantification:

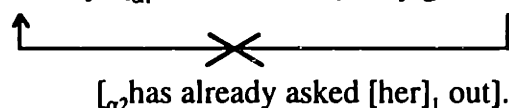
- (49') **The Coordination QR Generalization (CQRG):** In a structure such as (i), an optional instance of QR can move QP outside of the coordination only if there is some scope bearing element  $\beta$  c-commanding the coordination such that (a)  $\beta$  and QP are scopally non-commutative and (b) QR moves QP over  $\beta$ , as in (ii).

- (i) [ $_{YP} \dots$  [ $_{\alpha_1} \dots$  QP .. ] and [ $_{\alpha_2} \dots x \dots$ ]]  
 (ii) QP <sub>$x$</sub>  [ $_{YP} \beta \dots$  [ $_{\alpha_1} \dots x \dots$  ] and [ $_{\alpha_2} \dots x \dots$ ]]

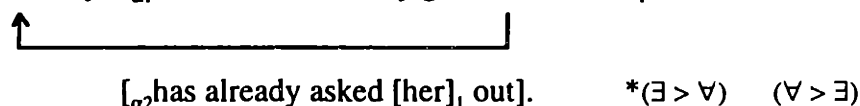
How would we know if QR has moved QP outside of the coordination? The most obvious way is to see whether the variable  $x$  can be bound by QP.

Consider the contrast in (54).

- (54) a. \*Billy [ $_{\alpha_1}$  wants to date [every girl in this class] <sub>$i$</sub> ] and



- b. A boy [ $_{\alpha_1}$  wants to date [every girl in this class] <sub>$i$</sub> ] and




---

movement. Analyses that do not assume LF movement (e.g., Baker (1970) and more recently Reinhart (1994) and Tsai (1994)) assume that the in situ WH-operator provides a variable that is bound by a matrix (null) operator. This matrix operator will bring about vacuous quantification in (53a) but not in (53b). Note that if there are arguments against LF WH-movement (c.f., Reinhart 1994), then the ill-formedness of (53a) provides a strong argument for an analysis that divorces the CSC from movement (such as the multi-dimensional analysis that I assume).

This contrast follows from Scope-Economy. In (54b), QR over the subject reverses the relative scope of two scopally non-commutative expressions (*every girl in this class* and *a boy*), while in (54a) it doesn't. Therefore, in (54a) there is nothing that allows QR to move the object QP outside of coordination. Consequently, binding the variable in  $\alpha_2$  is impossible. It is also important to notice that (54b) is restricted to Inverse Scope.<sup>44</sup> This is exactly what we expect under Scope-Economy; optional QR to any position above coordination will have no motivation.

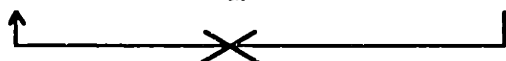
However, the predictions are more intricate. Specifically, we predict that sentences such as those in (54) will be acceptable if and only if the subject and the object in the first conjunct are scopally non-commutative:

- (55) **The CQRG (a consequence of):** In a structure such as (i),  $QP_2$  can bind  $x$  only if  $QP_1$  and  $QP_2$  are scopally non-commutative.

(i)  $[_{VP}QP_1 [_{VP1}\dots QP_2\dots]]$  and  $[_{VP2}\dots x\dots]$

This prediction, as far as I can tell, is correct:

- (56) a. \*Every boy  $[_{\alpha_1}$ wants to date [every girl in this class],] and








<sup>44</sup> Demonstrating the restriction to Inverse Scope in (54b) is not trivial given the entailment relations. (There is no situation in which the sentence is false under Inverse Scope and true under Surface Scope. See Abusch (1994) and references therein.) The restriction is apparent when we consider operators that reverse entailment relations. Consider (i).

- (i) I doubt that a boy  $[_{\alpha_1}$ wants to date [every girl in this class],] and  
 $[_{\alpha_2}$ has already asked [her], out].  $*(\exists > \forall) \quad (\forall > \exists)$

For this sentence to be true, the speaker must doubt that every girl in the class has an admirer (with the appropriate properties). The speaker cannot utter the sentence to express a doubt that all of the girls have the same admirer. This is clear when we compare (i) with (ii).

- (ii) I doubt that there is a boy who  $[_{\alpha_1}$ wants to date every girl in this class] and  
 $[_{\alpha_2}$ has already asked every girl out].

- [ $\alpha_2$ has already asked [her]<sub>i</sub> out].
- b. Many boys [ $\alpha_1$ want to date [every girl in this class]<sub>i</sub>] and  
  
 [ $\alpha_2$ have already asked [her]<sub>i</sub> out].      *\*(Many >  $\forall$ )* ( $\forall$  > *Many*)
- (57) a. \*The boy sitting next to me [ $\alpha_1$ wants to date [every girl in the class]<sub>i</sub>] and  
  
 [ $\alpha_2$ has already asked [her]<sub>i</sub> out].
- b. The boy sitting next to her<sub>1/\*2</sub> [ $\alpha_1$ wants to date [every girl in the class]<sub>i</sub>] and  
  
 [ $\alpha_2$ has already asked [her]<sub>i</sub> out].
- (58) a. \*Every boy who loves Mary [ $\alpha_1$ wants to date [every girl in the class]<sub>i</sub>] and  
  
 [ $\alpha_2$ has already asked [her]<sub>i</sub> out].
- b. Every boy who loves her<sub>1/\*2</sub> [ $\alpha_1$ wants to date [every girl in the class]<sub>i</sub>] and  
  
 [ $\alpha_2$ has already asked [her]<sub>i</sub> out].<sup>45</sup>

What is common to all of the (b) sentences in (54-58) is that there is no motivation for QR of *every girl in the class* to a position above  $\alpha_1$  in the Component Structure that contains this DP. We thus expect that in all these cases movement of *every girl in the class* to a position above  $\alpha_1$  is not licensed, and, subsequently, binding of the pronoun in  $\alpha_2$  is impossible.

**2.3.2.3. The problem of Telescoping:** Unfortunately, the paradigm in (54-58) doesn't always hold. As noted by Reinhart (1987) and by Ruyé (1993) constructions such as (59) are well-formed.<sup>46</sup>

<sup>45</sup> Some speakers find (58b) as well as (57b) slightly marginal. However, these speakers do find a contrast between (a) and (b). I think it is reasonable to assume that the marginality of (58b) and (57b) is due to WCO. See the discussion of WCO in 2.2.3.1.

<sup>46</sup> A few speakers, myself not included, find (51b) (under Inverse Scope) better than (59). Although these judgments support my general claim, I ignore them, since they are very

(59) Mary likes every professor<sub>1</sub> and wants him<sub>1</sub> to be on her committee.

(59) seems to disconfirm the CQRG. However, there is a confound in the experiment. The confound is that constructions that contain universal quantifiers sometimes exemplify a phenomena which we might call *illusory variable binding*. Sometimes a universal quantifier can appear to bind a pronoun even though it is obvious that it doesn't c-command the pronoun at any level of representation. Such, illusory binding, called *telescoping* by Roberts (1987), is illustrated in (60)

- (60) a. Every graduating student walked up to the stage. He shook the dean's hand and returned to his seat. (Roberts 1987)  
 b. Every story pleases these children. If it is about animals, they are excited. If it is about witches, they are enchanted, and if it is about humans they never want it to stop. (Belvadi 1989)

The precise restrictions on telescoping are not clear (cf. Poesio and Zucchi (1992)). It is therefore unclear whether it can apply in (59). However, I don't think that this should cast serious doubt on the validity of the evidence I provided in (54-58). What we saw in (54-58) is that whenever we can tell that variable binding is impossible in a configuration of the form  $[_{VP_1} \dots QP_2 \dots]$  and  $[_{VP_2} \dots x_2 \dots]$  (whenever telescoping is impossible) we observe that variable binding becomes possible if the subject and  $QP_2$  are scopally non-commutative. This is exactly what we expect given Scope-Economy.

**2.3.2.4. Downward entailing operators:** One way to control for telescoping is by the use of downward entailing operators, such as *only one*. With such operators (as noted by Evans (1980)) there is an interpretive difference between real and illusory variable binding. This is illustrated in (61).

- (61) a. Only one student walked up to the stage. He shook the dean's hand and returned to his seat.  
 b. Only one student walked up to the stage, shook the dean's hand and returned to his seat.

In (61a) *only one student* doesn't c-command the pronoun. Still, the pronoun is

---

subtle (perhaps for the reasons I discuss below).

somehow related to the operator. This relation, however, is not that of syntactic binding, which is available in (61b). The two are semantically distinct. (61a) makes a stronger assertion. For this utterance to be truthful, there must be no more than one student that walked up to the stage. Further, the unique student that walked up to the stage is required to do two other things. (61b) makes a weaker assertion. For this sentence to be true, there could be 3 students who walked up to the stage, as long as only one of them did the two other things. This difference follows naturally from the assumption that in (61a) the pronoun is outside the scope of the operator and is related to it by some other strategy (e.g., an e-type strategy). This is illustrated in the pseudo-logical paraphrases in (61').

- (61') a. [Only one x: (student(x) & walked up to the stage (x))] & [the unique student that walked up to the stage shook the deans hand and returned to his seat]  
 b. Only one x: [student(x) & walked up to the stage (x) & shook the deans hand (x) & returned to one's seat (x)]

We could thus replace a downward entailing quantifier in the position of *every professor* in (59) and see whether variable binding is available. That only illusory binding is possible is indicated by the utterance in (62). The second sentence in this utterance contradicts the first one. The reason for this contradiction is the lack of real variable binding.

- (62) John loves 3 of the women he knows. However, he loves only one of them and expects her to love him back. (contradiction)

Had variable binding been possible, the meaning of the second conjunct in (62) would be parallel to that of (61b). In other words, the sentence would merely require that there be no more than one woman that has two properties (the property of being loved by John and the property of being expected by him to show some love in return). The existence of no more than one woman who has both properties doesn't contradict the existence of 3 who have one of them. Therefore, had syntactic binding been available in the second sentence of (62), we would expect there to be no contradiction with the first sentence. What (62) would mean, had syntactic binding been possible, can be illustrated by an utterance which involves syntactic binding but is parallel to (62) in all other respects. In (63) we see that such an utterance involves no contradiction.

- (63) John loves 3 of the women he knows. However, there is only one of them that he

loves and expects to love him back.

(no contradiction)

The lack of syntactic binding in the second sentence of (62) argues that *only one* cannot QR out of the coordinate structure. The reason, I suggest, is that the movement is not licensed within the Component Structure in which it would take place. Further evidence that this is the case is given by the contrast between (62) and (64).

- (64) John can love 3 of the women he knows. However he **can** [love only one of them] and [expect her to love him back].  
(not necessarily a contradiction)


In (64), optional QR of *only one* has motivation (within the relevant Component Structure). The motivation for QR is to have scope over the modal *can* (*can* and *only one* are scopally non-commutative). A by-product of QR is that syntactic binding is possible in the second Component Structure, and that consequently there is no contradiction with the first sentence.<sup>47</sup>

The utterance in (65) contrasts with (64) in the size of the constituent that is being coordinated in the second sentence. In (65), this constituent includes the modal *can* (as we can tell by the appearance of the agreement marker *s* in the second conjunct). Therefore, although QR is licensed, Shortest Move restricts its length; the QP *only one* cannot move outside the coordination and therefore cannot bind the pronoun in the second conjunct. Only e-type binding is possible, and the semantic result is a contradiction.

- (65) John can love 3 of the women he knows. However he [**can** love only one of them] and [expects her to love him back].  
(contradiction)

**2.3.3. Scope-Economy and coordination - the case of QL:** Let's consider again the contrast in (50), which argues that QR obeys the CSC.

<sup>47</sup>As pointed out to me by Irene Heim and Satoshi Tomioka, (64) has two non-contradictory interpretations. Under one interpretation, *only one* has scope over the modal *can*. Under the other interpretation, the sentence means something paraphrasable as *it is not the case that John can love more than one woman and expect her to love him back*. As Irene Heim pointed out, the second interpretation might require a decomposition of *only one* into *not more than one* (perhpas) followed by QR of *not*. Note that this doesn't bear on the argument I make. For the argument to go through it is enough to establish that (64) has one representation which involves QR out of the coordination. Such a representation is demonstrably unavailable for (62).

- (50) a. A student likes every professor.  $(\exists > \forall) (\forall > \exists)$   
 b. A student [[likes every professor] and [hates the dean]].  
  $(\exists > \forall) *(\forall > \exists)$

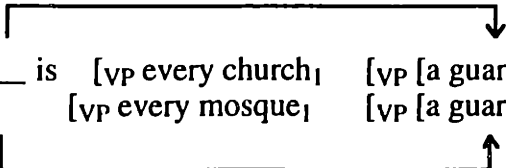
As mentioned, the contrast argues that QR obeys the CSC. However, when we consider the option for Quantifier Lowering and the existence of VP internal subjects, there is a gap in the explanation of the contrast. We understand why in (50b) the object can't raise by QR over the subject, but we don't understand why the subject can't lower by QL under the object. (See (7'''-8''').)

In order to deal with this, one might want to seek a principle that, in general, blocks the subject in coordinate structures from lowering into the coordinates. Example (66), however, shows that this is a wrong strategy. In (66), although the object quantifier is within a coordinated VP, it can have scope over the VP external subject. Assuming that QR obeys the CSC, the subject must be lowered under the object.

- (66) A guard is standing in front of every church and sitting at the side of every mosque.

Further, the interpretation of (66) indicates that it is the subject that lowers under the object rather than the object that QRs over the subject. To see this, we must first observe that in order for the subject to have narrow scope relative to the object of one of the VPs, it must also have narrow scope with respect to the other object. Putting aside the CSC, there are two principled ways for the subject to have narrow scope relative to both objects. One would be by moving both objects over the subject as in (66'a) and the other would be by combining obligatory VP-level QR in both conjuncts with ATB lowering of the subject to the two coordinated VPs, as in (66'b).

- (66') a. every mosque<sub>1</sub> every church<sub>2</sub> [A guard is {standing in front of t<sub>2</sub>} and {sitting at the side of t<sub>1</sub>}].

- b.  is [vp every church<sub>1</sub> [vp [a guard] standing in front of t<sub>1</sub>]] and [vp every mosque<sub>1</sub> [vp [a guard] sitting at the side of t<sub>1</sub>]].

There is an obvious difference in the interpretation of the LFs in (66'a) and in (66'b). For the LF in (66'a) to be true, each pair of a mosque and a church must have a single guard



standing both in front of the church and in front of the mosque. For the LF in (66'b) to be true, the guards can vary independently with respect to the churches and the mosques. (66'b) is an LF equivalent to the one achieved by independent wide scope for the object in the two conjuncts of (67).

- (67) A guard is standing in front of every church and  
a guard is sitting at the side of every mosque.

It seems clear that the interpretation of (66) is that of (66'b) and not that of (66'a). It is therefore reasonable to conclude that the LF of (66) is achieved by ATB lowering and not by QR out of both conjuncts. (66), therefore, doesn't raise any problem for the claim that QR obeys the CSC. However, we want an explanation for the difference between (50b) and (66). Why is ATB lowering allowed in (66) and not in (50b)?

Scope-Economy together with the multi-dimensional analysis of coordination provide the answer. First, notice that from the multi-dimensional analysis of coordination it follows that lowering is possible into a coordinate structure only if it is ATB. To see this, consider (66'b) and its two Component Structures in (68). Had lowering taken place only in one of the Component Structures, the trace in the other Component Structure would not have an antecedent (as we see in (69)).

- (68)
- |    |       |   |  |
|----|-------|---|--|
| 1. | __ is | [ <sub>VP</sub> every church <sub>1</sub> | [ <sub>VP</sub> [a guard] standing in front of t <sub>1</sub> ]]   |
| 2. | __ is | [ <sub>VP</sub> every mosque <sub>1</sub> | [ <sub>VP</sub> [a guard] sitting at the side of t <sub>1</sub> ]] |
- 

- (69)
- |    |       |   |  |
|----|-------|---|--|
| 1. | __ is | [ <sub>VP</sub> every church <sub>1</sub> | [ <sub>VP</sub> [a guard] standing in front of t <sub>1</sub> ]] |
| 2. | __ is | [ <sub>VP</sub> every mosque <sub>1</sub> | [ <sub>VP</sub> t sitting at the side of t <sub>1</sub> ]]       |
- 

Consider the point in the derivation of (66) at which the object has already undergone short QR (adjunction to VP), but the subject has not yet undergone ATB lowering. The structure at this point is represented in (66'').

- (66'') a guard<sub>2</sub> is [<sub>VP</sub> every church<sub>1</sub> [<sub>VP</sub> t<sub>2</sub> standing in front of t<sub>1</sub>]] and  
[<sub>VP</sub> every mosque<sub>1</sub> [<sub>VP</sub> t<sub>2</sub> sitting at the side of t<sub>1</sub>]]

The two Component Structures:

1. a guard<sub>2</sub> is [vp every church<sub>1</sub> [vp t<sub>2</sub> standing in front of t<sub>1</sub>]]
2. a guard<sub>2</sub> is [vp every mosque<sub>1</sub> [vp t<sub>2</sub> sitting at the side of t<sub>1</sub>]]

To determine whether Scope-Economy allows QL to move the subject into both conjuncts we must take an independent look at each of the Component Structures. Scope-Economy, like any other grammatical constraint, is checked independently in the two Component Structures. Therefore, at each Component Structure, QL must reverse the relative scope of two non-commutative quantifiers. As we see in (66''), in both Component Structures, QL alters the scope of an existential and a universal quantifier. Since these quantifiers are scopally non-commutative, QL is licensed in both Component Structures.

In (50b), by contrast, lowering the subject into the second conjunct will not reverse the relative scope of two non-commutative quantifiers within the second Component Structure and is thus ruled out.

More generally, we predict the following generalization:

- (70) **The Coordination QL Generalization (CQLG):** In a structure such as (i), ATB QL is possible only if there are two expressions  $\beta$  and  $\beta'$  such that (a) both  $\langle\beta, QP\rangle$  and  $\langle\beta', QP\rangle$  are scopally non-commutative and (b) QL (in one Component Structure) moves QP under  $\beta$ , and (in the other Component Structure) under  $\beta'$ , as in (ii).

- (i) [<sub>YP</sub>QP<sub>1</sub>..... [ <sub>$\alpha_1$</sub> .....t<sub>1</sub>.....] and [ <sub>$\alpha_2$</sub> .....t'<sub>1</sub>.....]]
- (ii) [<sub>YP</sub>..... [ <sub>$\alpha_1$</sub>  $\beta$ ...QP...] and [ <sub>$\alpha_2$</sub>  $\beta'$ ..QP..]]

Further evidence in favor of this generalization is provided in (71).

- (71) a. A guard is standing in front of every church and sitting at the side of every mosque.  
 b. # A guard is standing in front of every church and sitting at the side of this mosque.

In (71a(=66)), Scope-Economy licenses QL in both Component Structures in (66''). ATB QL is licensed and results in (66'b). In (71b), by contrast, QL into the second conjunct is not licensed. Consequently, the sentence is semantically anomalous (since we cannot

imagine a single guard standing in front of all churches).<sup>48</sup>

To sum up, in this section we have seen that an invisibility assumption (which follows from a multi-dimensional approach to coordination) provides a test for Scope-Economy. The basic idea is that the CSC follows (as a by-product) from the necessary formation of Component Structures. These structures must independently obey all grammatical constraints. In particular, movement within a Component Structure must be motivated (internal to the structure). However, once movement takes place, it has effects also outside of the Component Structure (binding of a pronoun and scope relative to coordination). Testing for these effects confirms Scope-Economy. When an SSO doesn't have the semantic effects (required by Scope-Economy) within a Component Structure (when it doesn't reverse two scopally non-commutative expressions), it can't have other semantic effects that would result from movement.

## 2.4. The Locality of QR

In the previous sections I have claimed that QR is restricted by Scope-Economy and Shortest Move. In this section, I would like to discuss a consequence that this claim has for the locality of QR. Specifically, I would like to point out that this claim could help us understand the clause boundedness of this operation.

The clause boundedness of QR is attested in sentences such as (72), which are

---

<sup>48</sup> If the CQLG is correct, we expect the following paradigm:

- (i) A guard is standing in front of every church and sitting at the side of every mosque.
- (ii) A guard is standing in front of every church and sitting at the side of almost every mosque.
- (iii) A guard is standing in front of every church and sitting at the side of most of the mosques.
- (iv) A guard is standing in front of every church and sitting at the side of two of the mosques.
- (v) #A guard is standing in front of every church and sitting at the side of one of the mosques.
- (vi) #A guard is standing in front of every church and sitting at the side of this mosque.

Checking (i-vi) I found that all speakers accept (i) and reject (vi). With respect to (ii-v) I found some variation. About half the speakers I checked displayed judgments in accord with the CQLG.

restricted to Surface Scope.<sup>49</sup>

- (72) a. #Someone said that every man is married to Sue.       $\#(\exists > \forall)$        $\ast(\forall > \exists)$   
 b. #Someone said that Sue is married to every man.       $\#(\exists > \forall)$        $\ast(\forall > \exists)$

This restriction is extremely puzzling under the assumption that QR is a form of A-bar movement. Interestingly, is an immediate consequence of Scope-Economy and Shortest Move.<sup>50</sup> To see this, let's start with the structures that would result from the sentences in (72), once obligatory QR has applied. In (72a), the QPs are interpretable in their surface positions, and there is therefore no obligatory QR. In (72b), Obligatory QR moves *every man* to a position in which this QP can be interpreted. Shortest move determines that the landing site of this movement will be the embedded VP, yielding the structure in (72'b):

- (72') a. Someone said [<sub>CP</sub>that [<sub>IP</sub>every man is married Sue]].  
 b. Someone said that [<sub>IP</sub>Sue<sub>i</sub> is [<sub>VP</sub> every man<sub>2</sub> [<sub>VP</sub>t<sub>1</sub> married to t<sub>2</sub>]]].

At this point, *every man* is 'stuck'. There is no way for an optional step of QR to satisfy both Shortest Move and Scope-Economy. Shortest Move determines that if a second step of QR is to apply, it would target the nearest clause-denoting maximal projection (the embedded CP in (72'a) and the embedded IP in (72'b)). Scope-Economy determines that such a step is unavailable. (It doesn't reverse the relative scope of two scopally non-commutative operators, and, as such is unmotivated.)

Shortest Move predicts that every step of QR is restricted to the closest XP that dominates QP (where XP ranges over clause-denoting maximal projections). Scope-Economy determines whether QR is licensed. If a QP is in an interpretable position which is separated from another expression,  $\alpha$ , by an XP (and even if  $\langle \alpha, QP \rangle$  are scopally non-commutative), Scope-Economy and Shortest Move predict that the scopal relationship between the quantifiers will be restricted to Surface Scope.

This suggestion makes a prediction, which is unexpected under any other approach

<sup>49</sup>The factual claim that quantifiers are clause-bound has not remained unchallenged. See May (1988), Reinhart (1991; 1995) and Wilder (1997).

<sup>50</sup> As pointed out by Hornstein (1995), the restriction is expected if QR is a form of A-movement. However, this has been contested by Kennedy (1997) and Johnson and Tomioka (1997). Furthermore, in chapter 6 we will see that treating QR as A-bar movement allows us to account for various interactions between QR and Binding Theory.

to locality. The prediction is that the locality of QR would be obviated in cases where movement to each intervening XP is motivated by Scope-Economy. Evidence that this prediction might be correct can be drawn from an observation made in Moltmann and Szabolcsi (1994):<sup>51</sup>

- (73) a. One girl knows that every boy bought a present for Mary.  $(\exists > \forall) *(\forall > \exists)$   
 b. One girl knows what every boy bought for Mary.  $(\exists > \forall) (\forall > \exists)$

In (73a), just as in (72), the universal quantifier in the embedded sentence cannot out-scope the existential quantifier in matrix subject position. The reason for this is that Scope-Economy rules out movement to the embedded CP, and that longer movements are ruled-out by Shortest Move. In (73b), by contrast, movement of the universal quantifier over the *Wh*-operator is semantically motivated (it provides a pair list reading, or, more accurately, it reverses the relative scope of a universal and an existential quantifier).<sup>52</sup> After the

<sup>51</sup> Moltmann and Szabolcsi (1994) analyze the contrast in (73) without abandoning the clause-boundedness of QR (though, see next note). In particular they claim that in (73b), the effects of  $(\forall > \exists)$  result from QR of the embedded interrogative accompanied by QR internal to the interrogative (over the *wh*-phrase). Moltmann and Szabolcsi provide one strong argument for this view. They show that when a pronoun in the embedded interrogative is bound by the matrix subject, wide scope for the universal quantifier is ruled-out:

- (i) A girl<sub>i</sub> know what every boy bought for her<sub>i</sub> mother.  $(\exists > \forall) *(\forall > \exists)$

Many speakers I consulted agree with Moltmann and Szabolcsi. However, it appears that the contrast between (73b) and (i) re-emerges in (ii), where an extension of Moltmann and Szabolcsi's analysis would be extremely implausible.

- (ii) a. A girl<sub>i</sub> expected every boy to come to the party.  $(\exists > \forall) (\forall > \exists)$   
 b. A girl<sub>i</sub> expected every boy to come to her<sub>i</sub> party.  $(\exists > \forall) *(\forall > \exists)$

(ii) suggests to me that Inverse Scope is very difficult for a surface structure configuration of the form  $[QP_1 \dots QP_n \dots \text{pronoun}_i]$  (but cf. Hornstein (1995:160, 180) and Pica and Snyder (1994: 345)). If this is in general true, Moltmann and Szabolcsi's observation doesn't pose a serious threat to the proposal I am making.

<sup>52</sup> It is well-known that pair-list reading are possible whenever the universal quantifier c-commands the base-position of the *wh*-phrase. (See, among other, Aoun and Li (1993).):

- (i) what did you say that every boy bought.  
 (ii) Someone knows what you said that every boy bought.  $(\exists > \forall) (\forall > \exists)$

I assume that the *wh*-phrase has a landing site in every maximal projection and that the *wh*-

universal quantifier moves over the *Wh*-operator, it can adjoin to the matrix VP (reversing the relative scope of an Intensional verb and a universal quantifier). Finally it can move over the subject position, yielding the attested  $(\forall > \exists)$  interpretation.<sup>53</sup>

Notice, however, that given the way Shortest Move is defined, the predicted restriction on QR might be too stringent. Consider, for example, the sentences in (74).

- (74) a. Someone expects Sue to marry every boy.  $(\exists > \forall) (\forall > \exists)$   
 b. He demanded that we read not a single book. (Kayne 1981).  
*(demand > not a single) (not a single > demand)*

In these sentences, QR appears to cross many clause-denoting expressions. Given the results of this chapter, there are two possible explanations for this fact. The first possibility is that the sentences in (74) contain null (modal) operators which provide motivation for the necessary intermediate instances of QR.<sup>54</sup> The second possibility is that Shortest Move is a

---

trace (copy) provides motivation for every step of QR.

<sup>53</sup> It is possible that this suggestion makes further predictions. One potential prediction is that replacing the subject of (72b) with an existential quantifier would allow matrix scope for the embedded universal quantifier. A few speakers I consulted found a contrast between the sentences in (i) and (ii). According to these speakers, the universal quantifiers in (i) but not in (ii) can have matrix scope. However, this contrast is not necessarily predicted. The contrast would be predicted only if adjunction to IP can serve as an escape-hatch for crossing an embedded CP.

- (i) a. \*A different person said that Sue is married to every man.  
 b. At least one person said that Sue is married to each of these men.  
 (ii) a. A different person said that someone will marry every boy  
 b. At least one person said that someone is married to each of these men.

<sup>54</sup> David Pesetsky (p.c.) has pointed out that such a proposal might shed light on the fact that a subjunctive clause (in contrast to an indicative clause) is not a barrier to QR.

Similarly consider the following contrast.

- (i) a. Someone expects Sue to marry every man.  
 b. Someone believes Sue to be married to every man.  
 (ii) a. Someone expects Sue to know every man.  
 b. Someone believes Sue to know every man.

Although the judgments are subtle, it seems that Inverse Scope is easier in the (a) sentences, which plausibly contain a modal operator in the embedded clause. (Consider the

weaker condition than what I've assumed.

In order to investigate the second possibility we must ask what aspects of Shortest Move were crucial in order to capture the generalizations reported in this chapter. It seems to me that the account depends on three claims; (a) that optional QR applies only when it reverses the relative scope of non-commutative expressions (Scope-Economy), (b) that obligatory QR (assuming that it exists) is necessarily short (it targets the closest VP) and (c) that optional QR targets the closest position that dominates the relevant scope bearing element which is being crossed. A very natural way to capture these generalizations is by Scope-Economy and Shortest Move (as I've defined it). However, there are other possibilities. Among other things, it is possible to assume that (a) and (c) are part of the definition of optional QR and that Shortest Move applies only to obligatory QR. Whatever turns out to be the right account of (a-c), it should yield certain locality restrictions on QR. (Any statement of the relevant economy conditions that would capture (a-c) will restrict intermediate steps of QR.) For this reason, it seems to me that Scope-Economy (under any formulation) yields the attested result that QR is more restricted than other cases of A-bar movement.

## 2.5. Modularity - an argument in favor of a deductive system<sup>55</sup>

In the previous sections we have seen a variety of arguments in favor of the claim that optional SSOs are allowed to apply only when they have certain effects on semantic interpretation (Scope-Economy). If the claim is correct, it raises various questions regarding the possible interactions between systems that construct linguistic objects and systems that interpret these objects.

However, these questions can be formulated only given certain assumptions about the nature of the relevant systems. Based on the results of research in generative grammar, the following assumptions seem very plausible:

---

paraphrase: *for every world consistent with x's desire, w, there is an extension of w, w', such that  $\phi$ .*) Thanks to Jon Nissenbaum and Rajesh Bhatt.

<sup>55</sup> Many of the arguments that appear here have appeared in a different context in Fox and Nissenbaum (1996). Additional considerations came up in discussions with Noam Chomsky and Philippe Schlenker.

- There is a separate cognitive system, *Syntax*, which consists of *Syntactic Operations* that combine primitive object (e.g., words or features) to form *Syntactic Structures*.
- There are interpretive systems (distinct from Syntax) that access Syntactic Structures and determine the way these structures can contribute to thought and communication (*Semantic Systems*).<sup>56,57</sup>

Given these assumptions, we might try to ask how exactly Syntax interacts with Semantic Systems. In a state of ignorance, the best strategy is to postulate a simple working hypothesis with the hope that it can be corrected as we gain further understanding. A very simple working hypothesis for the interaction between Syntax and Semantic Systems, which we can call the Modularity Hypothesis (MH), is that although the products of Syntax are interpreted, Syntactic Operations apply blindly, oblivious to the resulting interpretations:

- **Modularity Hypothesis (MH):** The application of Syntactic Operations is not conditioned by the interpretation of the Syntactic Structures that would result.

MH has been a very productive working hypothesis. However (assuming that optional QR and QL are Syntactic Operations, see note 63) it is inconsistent with Scope-Economy. Scope-Economy states that the relevant Syntactic Operations are conditioned by the nature of the semantic interpretations that they generate, and thus is inconsistent with the claim that Syntactic Operations are oblivious to their interpretive consequences.

If Scope-Economy is correct, as I have tried to argue, MH needs to be modified. What are the possible modifications? The most radical modification is that Syntactic Operation are (in principle) conditioned by every property that relates to the way a sentence can contribute to thought or communication. However, this is not a very plausible

---

<sup>56</sup> The term "Semantics" has many uses. Here the term "Semantic Systems" is defined so as to refer to the internal (cognitive) systems which take a Syntactic Structure and contribute to determining the way it can be used in thought and communication. Under this definition, a system that contains syntactic rules of inference (if it exists) is a Semantic System.

<sup>57</sup> Obviously, there are also Interpretive systems that access Syntactic Structures and determine the way in which they can contribute to phonetic realizations (*Morpho-Phonological Systems*). However, Scope-Economy has very little to say about the interactions between these systems and Syntax (though it might be relevant to discussion of Marantz's (1997) notion of late insertion (See Fox (1995a: note 9)).

See Bobaljik (1998) for a discussion of an economy condition which raises questions about the interactions between Syntax and Morpho-Phonology which are very similar to the considerations that I address here.



possibility. We know very well that the entirety of effects that a sentence can have on thought processes is totally chaotic (See, for example, Davidson's (1986) discussion of the considerations that enter into the construction of a "passing theory" that would determine the truth conditions of a sentences. See also Chomsky (1992:214-215).) The correct modification will have to be less radical.

We've seen that Scope-Economy *is* sensitive to semantic properties. However, the semantic properties we've looked at concern the meanings of logical terms, such as universal and existential quantifiers, conjunction, disjunction, negation, modal/intensional operators (which can be viewed as universal and existential quantifiers over possible worlds) and variable binding.<sup>58</sup> The properties of such logical terms, which we can call properties of "logical syntax", can be formalized by deductive systems. Suppose that among the faculties of the mind there is such a (language dedicated) deductive system (DS), and that DS consists of various rules of inference (familiar from proof theory). We can then say that optional QR and QL are ruled out by Scope-Economy whenever DS is capable of "proving" logical equivalence for two logical representations (with and without the SSO).

What I think Scope-Economy suggests is that DS, and DS alone, can condition the Syntactic Operations of QR and QL:

- **Revised Modularity Hypothesis (RMH):** Among the Semantic Systems there is a formal deductive system (DS) which conditions certain Syntactic Operations. Syntactic Operations are not conditioned by other Semantic Systems.

One argument in favor of RMH has already been provided. If Scope-Economy is correct, MH needs to be revised by an alternative working hypothesis. Radical revision is implausible, and RMH is a coherent alternative (For a somewhat related discussion see Hintikka (1977) and the response in Chomsky (1979).)

The next argument is based on the observation that semantic properties that are not plausibly attributable to DS are irrelevant for Scope-Economy. As mentioned, the properties that *are* relevant to Scope-Economy are the properties of logical terms. This can be seen in more detail by comparing the cases in which logical equivalence is provable (and an SSO is blocked) with the cases in which it isn't and an SSO is licensed:

---

<sup>58</sup> As mentioned (notes 2, 23), Reinhart's argument for Scope-Economy was based on the suggestion that this condition is not sensitive to variable binding.

- (75) a. **John** likes every teacher. (Scopally Uninformative)  
 b. **A boy** likes every teacher. (Scopally Informative)
- (76) a. **Every boy** likes every teacher. (Scopally Uninformative)  
 b. **Many boys** like every teacher. (Scopally Informative)
- (77) a. **Every person who was involved in the production of the film festival** admires every movie. (Scopally Uninformative)  
 b. **Every person who was involved in its production** admires every movie (Scopally Informative)
- (78) a. **The person who produced the film festival** admires every movie. (Scopally Uninformative)  
 b. **The person who produced it** admires every movie. (Scopally Informative)
- (79) a. **John and Bill** like every teacher. (Scopally Uninformative)  
 b. **John or Bill** like every teacher. (Scopally Informative)
- (80) a. **John does speak more than three languages** (Scopally Uninformative)  
 b. **John doesn't speak more than three languages.** (Scopally Informative)
- (81) a. **John** seems to have won a gold medal. (Scopally Uninformative)  
 d. **An American athlete** seems to have won a gold medal. (Scopally Informative)
- (82) a. **A guard** is standing in front of a **building.** (Scopally Uninformative)  
 b. **A guard** is standing in front of every **building.** (Scopally Informative)
- (83) a. **A guard** is standing in front of **this building.** (Scopally Uninformative)  
 b. **A guard** is standing in front of every **building.** (Scopally Informative)

In the (a) sentences, which I have called Scopally Uninformative, Scope-Economy doesn't allow Inverse Scope. The reason (under RMH) is that DS can prove that the boldfaced quantificational expressions in the (a) sentences are scopally commutative, whereas for the (b) cases proof of commutativity is impossible (since commutativity doesn't hold).

The question to ask is what happens in cases in which commutativity *does* hold but not due to properties that are plausibly attributable to DS. In such cases, RMH would lead us to expect that Inverse Scope would be possible despite commutativity. Evidence that this is the case can be extracted from the sentences in (84).

- (84) a. In our class that consists of 40 students, at least one girl is taller than every boy.  
 b. An odd number of guards is standing in front of an odd number of buildings.

c. Rob doesn't speak more than half of the 9 languages spoken in Sydney.<sup>59</sup>

If we take into account the contribution that arithmetic expressions make to truth conditions, the sentences in (84) are equivalent under Surface and Inverse Scope. For (84a) and (84c), observing truth conditional equivalence is trivial, for (84b) it involves a more intricate proof. (See Westerstahl 1994.) What is important, however, is that for all the sentences, a proof of truth conditional equivalence must access the meaning of arithmetic terms, and cannot be determined by properties of logical syntax which are plausibly attributable to a language dedicated deductive system. (In (84a), the proof depends on the fact that 40 is a finite number, in (84b) it depends on the meaning of the word *odd*, and in (84c) it depends on the meaning of *half* and on the fact that 9 is an odd number.<sup>60</sup>)

If we assume that Scope-Economy accesses DS and no other semantic system (i.e. that it is only sensitive to the meaning of logical terms) we would predict that the truth conditional equivalence of Surface and Inverse Scope in sentences such as those in (84) would not affect the licensing of SSOs. In other words, we would predict that sentences as those in (84) would not be limited to Surface Scope. This predictions seems to be correct.

The first indication that this is the case is the fact that speakers have a clear feeling that the sentences in (84) are scopally ambiguous.<sup>61</sup> The speakers I consulted report that the sentences have two meanings, and it takes some work to illustrate to them that (as far as truth conditions go) the meanings are identical.<sup>62</sup> The second (perhaps more important)

---

<sup>59</sup> I thank Philipp Schlenker and Ed Keenan for providing me with (84b), and for mentioning Westerstahl's paper. (84c) is from Fox and Nissenbaum (1996).

<sup>60</sup> If, as I suggest, Scope-Economy uses a local algorithm that considers properties of two elements whose relative scope is being reversed, (84a) is irrelevant. The meaning of the comparative will not be accessible to Scope-Economy. Therefore, even if the deductive system cared about the meaning of the numeral expression 40, Scope-Economy would not rule out Inverse Scope. (84a) provides an argument for a deductive system only if I am wrong about the local algorithm for Scope-Economy (which is indeed a possibility)

<sup>61</sup> For (84b), the intuitions are pretty difficult since the most prominent reading is the cumulative interpretation under which we *really* cannot distinguish Surface from Inverse Scope.

<sup>62</sup> How can we make sense of this fact? What I think it suggests is that speakers have an intuition about verification procedures (cf. Reinhart 1995). In (84a), for example, they feel that there is one interpretation (Surface Scope) in which the default procedure for verification will involve scanning the girls to see whether there is one who is taller than all the boys. The other interpretation (Inverse Scope) suggests a different verification procedure. Under the Inverse Scope procedure, one is guided to start with the boys and to see whether each one of them is shorter than some girl.

indication is that we can use the tests designed in the previous sections (as well as other tests) in order to demonstrate that the intuitions are real. I.e., we can demonstrate that the sentences in (84) differ from the Scopally Uninformative sentences in (75-83) in a way that indicates that in the former (in contrast to the latter) Inverse Scope is possible.

In order to see the availability of Inverse Scope in (84a), consider the contrast in (85). In (85a), Scope-Economy rules out QR of *every boy* to a position in which it can bind the pronoun *him*. As discussed in 2.3., the reason for this is that in the Component Structure in which QR applies, there are no semantic consequences to Inverse Scope. Under RMH, the unavailability of QR would be the result of a proof by DS that an interpretation of the relevant Component Structure with Surface Scope is equivalent to an interpretation with Inverse Scope.

- (85) In our class, which consists of 40 students
- a. \*Mary is [taller than every boy<sub>i</sub>] and [has threatened to beat him<sub>i</sub> up].
  - b. At least one girl is [taller than every boy<sub>i</sub>] and [has threatened to beat him<sub>i</sub> up].

In (85b), Inverse Scope will not have semantic consequences within the relevant Component Structure. Nevertheless, Inverse Scope is licensed. The conclusion is that Scope-Economy is not sensitive to the semantic properties that would yield truth

---

If this speculation is correct, it might suggest that (in general) sentences provide us with verification procedures. If we learn how to investigate these procedures, Scopally Uninformative sentences might turn out to be Scopally Informative. In such a case, we might be able to find direct tests for Scope-Economy. Consider, for example, the two sentences in (i):

- (i)
- a. Every boy loves every girl.
  - b. Every girl is loved by every boy.

Here intuitions are very unreliable, and, needless to mention, should be viewed with skepticism. Nevertheless, it does seem that the default verification procedure for (ia) involves scanning the boys to see whether each one of them loves every girl, whereas in (ib), default verification will start out with the girls. DS can prove “that the two verification procedures are equivalent”, hence Scope-Economy doesn’t allow Inverse Scope, and consequently each of the sentences in (i) is limited to one verification procedure.

It is very difficult to probe intuitions about verifications procedures because we are capable of using various sorts of deductions (way beyond those determined by DS), and these deductions obscure the default (putative) procedures (though sometimes the procedures are very apparent as in the difference between *All Ravens are black* and *All non black things are non ravens*). Nevertheless, if we had a way to probe intuitions about verification procedures, we could broaden the realm of linguistic evidence, and (among other things) subject Scope-Economy to further tests.

conditional equivalence for Surface and Inverse Scope in (84a).

In order to see the availability of Inverse Scope in (84b), Consider the contrast in (86). In (86a), as I have claimed in 2.2.3., Scope-Economy blocks Inverse Scope in the antecedent sentence, and Parallelism blocks Inverse Scope in the ellipsis sentence yielding an anomalous interpretation. Assuming that DS is responsible for proving logical equivalence, we would say that DS treats the subject quantifier in (86a) as an existential QP (ignoring the adjective *odd*), and proves that such a quantifier phrase is scopally commutative with respect to a definite description.

- (86) a. #An odd number of guards stood in front of this building and two policemen did, in front of every church.  
 b. ?An odd number of guards stood in front of an odd number of buildings and two policemen did, in front of every church.

In (86b), Inverse Scope is licensed in the antecedent sentence despite scopal commutativity. We can see this by the availability of Inverse Scope in the ellipsis sentence. The conclusion is that Scope-Economy is not sensitive to the semantic properties that would yield truth conditional equivalence for Surface and Inverse Scope in (84b).

To see the availability of Inverse Scope in (84c) requires a more complicated test which was designed in Fox and Nissenbaum (1996). To set the stage for this test, consider the contrast in (87).

- (87) a. Rob doesn't speak more than 2 of the 9 languages spoken in Sydney.  
 Perhaps he will never learn to speak them. (them = the languages spoken in Sydney that Rob doesn't speak)  
 b. There is no way that Rob speaks more than 2 of the 9 languages spoken in Sydney.  
 \*Perhaps he will never learn to speak them. (them = the languages spoken in Sydney that Rob doesn't speak)

In (87a), there is a pronoun in the second sentence which refers to the languages spoken in Sydney that Rob doesn't speak (probably an e-type pronoun). Such an interpretation for the pronoun depends on Inverse Scope. In (87b), there is an island intervening between the two scope bearing elements, thus blocking Inverse Scope. Consequently, the pronoun cannot get the intended interpretation. Given the contrast in (87), it is plausible to view the intended interpretation of the pronoun as an indicator of the availability of Inverse Scope. The contrast in (88) suggests that Inverse Scope is possible in the first sentence of (88b), which is identical to (84c).

- (88) a. Rob doesn't speak more than half of the 9 languages spoken in Sydney.  
 Perhaps he will never learn to speak them. (them = the languages spoken in Sydney that Rob doesn't speak)
- b. There is no way that Rob speaks more than half of the 9 languages spoken in Sydney.  
 \*Perhaps he will never learn to speak them. (them = the languages spoken in Sydney that Rob doesn't speak)

To sum up, we have seen that the sentences in (84) are different in an important way from the (a) sentences in (75-83). Scope-Economy is sensitive to the truth conditional equivalence of Surface and Inverse Scope in (75-83), but is oblivious to this equivalence in (84). This difference is exactly what we would expect under the existence of a deductive system together with RMH. We can therefore take the difference as evidence in favor of the postulation of a deductive system.<sup>63</sup>

The discussion in this section is admittedly very speculative, and might be extremely pre mature. However, if there is something to the speculations, the potential gain might be significant. Specifically, we can hope to distinguish between properties of meaning which are determined by the linguistic system (hence accessible to Scope-Economy) from properties of meaning that come about from other cognitive systems (hence, like our knowledge of arithmetic, are incapable of conditioning Syntactic Operations).

## 2.6. What is special about SSOs?

As far as I can tell, there are no operations other than optional SSOs which are subject to Scope-Economy. Why should this be the case? In this section, I would like to make a speculation inspired by a proposal made in Chomsky (1995: 377) and Krifka (1998). This speculation capitalizes on the fact that optional SSOs are the only optional operations that do not affect phonology.

---

<sup>63</sup> Noam Chomsky has pointed out to me that MH is threatened by Scope-Economy only under the assumption that optional QR and QL are Syntactic Operations. If the operations are part of the interpretive system itself, there would be no need to modify MH. However, there is a lot of evidence that SSOs affect structure. SSOs are structure building operations with many properties of (non-putative) syntactic movement (see chapter 6). Given this fact, Chomsky suggests that they should (perhaps) be viewed as structure building operations that apply in a post-syntactic system. Modularity questions could be avoided all together (he suggests) if the rules of inference needed for Scope-Economy apply in the same system that applies SSOs (i.e., if SSOs apply in DS).

Suppose that it is in general necessary that all optional operations have either a phonological or a semantic effect.<sup>64</sup> In the previous sections, I claimed that the relevant notion of a semantic effect has to do with scopal non-commutativity. Suppose that the relevant notion of a phonological effect has to do with word order (though see Krifka 1998). We can then view Scope-Economy as a special case of a more general condition, which we can call Output-Economy.

(89) **Output-Economy:** Optional operations must affect the output.

Output-Economy has as its consequence two different economy conditions:<sup>65</sup>

- (90) a. **Scope-Economy:** Covert optional operations (i.e., QR and QL) cannot be scopally vacuous (i.e., they must reverse the relative scope of two non-commutative quantificational expressions).  
 b. **Word-Order-Economy:** Overt optional operations cannot be string vacuous (i.e., they must reverse the relative order of two [perhaps phonologically overt] expressions)<sup>66</sup>.

If we could provide evidence for Word-Order-Economy, we will have an argument for this line of reasoning. In order to engage in such an investigation, we need to identify certain optional cases of overt movement. Among the operations that come to mind are optional scrambling and extraposition. I think that there is some indication that, in the case of extraposition, further investigation might yield an interesting confirmation. (For a related discussion of Scrambling [and a related speculation] see Krifka (1998).)

To set the stage for the test, consider the contrast in (91-92). The (a) sentences are ruled out by Condition C of the Binding Theory (BT(C)). The (b) sentences show that optional extraposition has an effect on BT(C). (See Taraldsen 1981.)

---

<sup>64</sup> Chomsky assumes that all movement is feature driven. Under this assumption, optional movement operations are those operations for which the triggering feature is not obligatorily present in the Numeration. The generalized economy condition states that these features can be inserted only if they have effects on the output.

<sup>65</sup> The way I state things here assumes that the kind of output effects that an operation must have depends on whether it is overt or covert. If this is correct, it would mean that an effect on Scope cannot obviate the requirement that an overt optional operation will have effects on phonology. This is more in line with Krifka's suggestion than with Chomsky's.

<sup>66</sup> As pointed out to me by Noam Chomsky, it is possible that Word-Order-Economy restricts certain instances of obligatory movement. See Chomsky (1986a: 48-54) and references therein.

- (91) a. ??I introduced him<sub>i</sub> to the woman that John<sub>i</sub> likes yesterday.  
 b. I introduced him<sub>i</sub> to the woman yesterday that John<sub>i</sub> likes.
- (92) a. ??/\* I want him<sub>i</sub> to explain the argument that John<sub>i</sub> was making with all my heart .  
 b. ? I want him<sub>i</sub> to explain the argument with all my heart that John<sub>i</sub> was making.

Given this contrast, the ill-formedness of the sentences in (93) is somewhat mysterious.

- (93) a. ??I Introduced him<sub>i</sub> to the woman that John<sub>i</sub> likes.  
 b. ??/\* I want him<sub>i</sub> to explain the argument that John<sub>i</sub> was making.

Why can't the derivation for these sentences include an instance of string-vacuous extraposition that would be structurally identical to the extraposition we've seen in (91b) and (92b)? Output-Economy provides the answer. Extraposition is an overt optional operation. Therefore, by Word-Order-Economy, it cannot be string-vacuous.

Another domain where we might hope to find interesting results is the domain of optional overt SSOs. If such operations exist, Output-Economy predicts that they will not have to affect semantic interpretation. Consider Hungarian. This language has been claimed to have both overt and covert QR. (See Szabolcsi (1998).) If this claim is correct, Output-Economy would predict that in Hungarian QR will need to affect semantic interpretations only when it is covert. This claim should be testable with the tools that were introduced in this chapter.



### Chapter 3: Asymmetries in Ellipsis and the Nature of Accommodation

In Chapter 2, I have argued that Scope Shifting Operations (SSOs) are subject to a Semantics-Sensitive Economy condition (Scope-Economy, henceforth, for brevity, just Economy). Under this condition, an SSO is licensed only when it affects Semantic Interpretation.<sup>1</sup> Because SSOs (generally) don't affect phonetic interpretation, the argument for Economy had to be based on various mechanisms that distinguish "observationally identical" representations. In particular, I focused on cases in which a certain constraint *C* can detect whether or not an SSO has applied. I have shown that in such cases, *C* indicates that the predictions of Economy are correct; the SSO applies only when it has semantic effects.

It is worth noting, however, that the predictions depend on the assumption that the SSO -- although licensed in order to get a distinct interpretation -- is not licensed in order to satisfy *C*. For illustration, let us look at the case of ATB QL, in which the relevant constraint is the CSC (see 2.3.3):

- (1) a. A guard is [t standing in front of every church] and  
          [t sitting at the side of every mosque].
- b. #A guard is [t standing in front of every church] and  
          [ $\bar{t}$  sitting at the side of this mosque].

The CSC (which, as I have mentioned, could follow from the ban on vacuous quantification and the requirement that every trace be bound) determines that QL is allowed to occur into one of the conjuncts in the sentences in (1), only if it occurs into the other conjunct as well. In (1a) both instances of QL are licensed, while in (1b) QL into the second conjunct doesn't yield a distinct semantic interpretation, and is therefore not licensed. The CSC thus indicates that QL cannot occur if it doesn't have a semantic effect. However, this prediction is based on the assumption that QL cannot occur into the second conjunct in order to satisfy the CSC (given that it has "already" occurred into the first conjunct). In other words, what we need to assume is that while Economy "sees" the semantic effects of the relative scope of two quantifiers, it does not "see" the effects of the CSC.

---

<sup>1</sup> The term Semantic Interpretation is to be understood with the qualifications stated in Chapter 2. See, in particular, 2.5.



In the (a) sentences of (3) and (4), Parallelism indicates that Economy rules out QR in the conjunct in which it would not yield a semantic effect. In the (b) sentences, however, Parallelism seems to make the opposite indication. This asymmetry was accounted for under a version of the Strict Cycle Condition (SCC) which ensures that operations apply to the complement of a head before they apply to the specifier. Given independent arguments that in coordination the second conjunct is the complement and the first is the specifier (as in (5)), it follows that SSOs apply in the second conjunct before they apply in the first conjunct.

(5) [XP [and[YP]]]

The disambiguation in the (a) sentences of (3) and (4) could now follow even if one assumes that Parallelism can motivate an SSO while the lack of disambiguation in the (b) sentences (henceforth, anti-disambiguation) required this assumption. The idea was that when the availability of the first SSO is considered, Parallelism is irrelevant; the SSO can apply only if the two scopal relations are semantically distinguishable. However, when the second SSO is considered, Parallelism is the condition that determines whether or not it applies. Because the SCC determines that the SSO is considered in the second conjunct first, the asymmetry is predicted.

In this chapter, I argue against the account of the asymmetry in (3) and (4) that I provided in Fox (1995a). Specifically, I aim to show that the asymmetry follows from an independently needed asymmetry in Parallelism. Building on work of Rooth (1992) and Tancredi (1992), I will argue that Parallelism imposes a weaker requirement than the one I have assumed in the previous chapter. Under certain circumstances, the sentence that contains ellipsis (henceforth  $\beta_E$ ) need not be directly isomorphic to the sentence that serves as its antecedent (henceforth  $\beta_A$ ). Under the relevant circumstances, a sentence  $\beta_{AC}$ , which follows (together with reasonable presuppositions) from  $\beta_A$  can be accommodated, and as long as  $\beta_{AC}$  is isomorphic to  $\beta_E$  Parallelism is satisfied. This asymmetry in the definition of Parallelism derives, as I will show, the asymmetry in (3) and (4). Furthermore, I will show that there are cases in which the asymmetry in (3) and (4) is not attested. These cases, which are incompatible with the assumption that Economy takes Parallelism into account, are expected under the alternative that I will propose.

The chapter has the following structure. In 3.1., I present various problems for the proposal made in Fox (1995a). In 3.2., I present the arguments from Rooth and Tancredi

for an asymmetry in Parallelism and provide an appropriate working-definition of the constraint. In 3.3., I demonstrate that the new definition of Parallelism yields the right results for (3) and (4) and avoids the problems presented in 3.1. This demonstration will rely crucially on the assumption that Parallelism cannot motivate QR and, as such, will provide an argument for the assumption. Because not all constraints on accommodation are well known, it will be difficult to characterize the set of predictions that the proposed account makes. Nevertheless, in 3.4. I will discuss a natural constraint on accommodation and show that this constraint makes certain clear predictions that are borne out.

### **3.1. Problems with the account of the asymmetry in Fox (1995a)**

There are two types of problems with the account I provided in Fox (1995a) (Henceforth the ordering account). The first type has to do with the assumptions that are made about ordering. As we will see, there is no way for these assumptions to follow from the SCC. The second type of problem is more substantial. The ordering account presupposes a generalization which is wrong. It presupposes that a first (Scopally Uninformative) conjunct will never disambiguate the second conjunct. However, it will turn out that in many cases disambiguation works in both directions. We will see that the asymmetry in (3) and (4) must be tied to the semantic properties of the quantifiers in these constructions.

**3.1.1. The SCC:** The assumption about ordering that I made in Fox (1995a) does not follow from standard accounts of the SCC. Standard assumptions about the SCC entail that no operations apply to XP after operations have applied to a constituent that dominates XP. However, nothing is said about the ordering of operations that apply internal to a specifier relative to operations that apply to a complement. For this reason, I had to assume a non-standard characterization of the SCC. Jacobson (1997), however, presents a good argument that even a non-standard characterization of the SCC would not yield the assumptions about ordering needed for the ordering account.

The first observation that Jacobson makes is that for the ordering account to work, the SCC must apply intersententially. This can be seen by the observation that the asymmetry seen in (3) and (4) persists when  $\beta_E$  and  $\beta_A$  are not contained in the same sentence. This is demonstrated by the contrast in (6). The (a) and the (b) sentences in (6) contrast in the same way that the (a) and the (b) sentence in (3) and (4) do. If the anti-



- (8) At least one critic from the *Times* admires every movie. Pauline Kael does, too, and at least one critic from the *Post* does, too. (Jacobson 1997)
- (9) Someone from NY is very likely to win the lottery. Jon is, too. However, it is not the case that someone from Boston is.

The first sentences of (8) and (9) are disambiguated in favor of the interpretation which does not result from the application of an SSO. (In (8) the subject *At least one critic from the Times* must take scope over the object *every movie*, and in (9) the subject *Someone from NY* must take scope in its surface position.) This is not predicted by the ordering account. Given that the ordering account is committed to an intra-sentential application of the SCC, it is predicted that the availability of an SSO in the third sentence will license (via Parallelism) an SSO in each of the preceding sentences.<sup>4</sup> When (6) and (7) are put together with (8) and (9), it is hard to imagine a way in which the assumption about ordering which are needed for the ordering account would follow from the SCC.

**3.1.2. The Presupposed Generalization:** The ordering account is based on the claim that in ellipsis constructions scopal disambiguation comes about only when Economy rules out an SSO in the *second* conjunct. If we assume that Economy doesn't see Parallelism, the predicted generalization is that disambiguation would come about whenever Economy rules out an SSO in *either* of the conjuncts. This difference can be seen in the two ways in which the Ellipsis Scope Generalization (ESG) could be stated:

- (10) **ESG** (from Fox 1995): In constructions that involve phonological reduction or deletion, Inverse Scope is possible only if it is semantically distinct from Surface Scope *in the sentence that includes the phonologically reduced/elided VP*.
- (11) **ESG** (from Ch. 1): In constructions that involve phonological reduction or deletion, Inverse Scope is possible only if it is semantically distinct from Surface Scope *both in the sentence that includes the phonologically reduced/elided VP and in the sentence that includes the antecedent VP*.

The asymmetry in (3) and (4) seemed to suggest that the generalization presupposed by the ordering account, (10), is correct. However, once we consider a broader set of examples, it

---

<sup>4</sup> This prediction was pointed out to me by David Pesetsky when I first proposed the ordering account. However, at the time we were unable to come up with constructions that would be simple enough to test the prediction (See Fox (1995: note 69).).



Consider the asymmetry attested in (14) and (15). In (14a) and (15a), phonological reduction is very natural. However, in (14b) and (15b), where the order of the sentences is reversed, reduction feels inappropriate.

- (14) a. First Bill called Mary an idiot. Then [John]<sub>F</sub> *insulted her*.  
 b. \*First John insulted Mary. Then [Bill]<sub>F</sub> *called her an idiot*.
- (15) a. John talked to every woman he saw. [Bill]<sub>F</sub> *talked to many women, too*.  
 b. \*Bill talked to many women. [John]<sub>F</sub> *talked to every woman he saw, too*.

Parallelism was the name we gave to the constraint that governs phonological reduction and deletion (ellipsis). In order to account for the asymmetry in (14) and (15) there must be some asymmetry in the definition of Parallelism.

In order to spell-out this asymmetry, I will have to get a little more technical. But first I would like to use the contrast in (14) to illustrate the intuitive source of the asymmetry (See Rooth (1992) and Tancredi (1992)). Let us start with (14a). The basic idea is that Parallelism is licensed because of a conclusion that we can plausibly draw from the proposition that Bill called Mary an idiot. Specifically, this proposition allows us to conclude the proposition that Bill insulted Mary and the latter proposition is clearly an appropriate antecedent for phonological reduction. In (14b), however, there is no inference that can license phonological reduction. The proposition that Bill insulted Mary does not allow us to infer the proposition that Bill called Mary an idiot, and therefore phonological reduction is not licensed. The basic intuition is that phonological reduction is licensed either by direct Parallelism with a linguistic object that is present in discourse or through indirect Parallelism, which involves accommodation (via inference) of a new linguistic object that satisfies direct Parallelism. This intuition is stated in (16) along the lines of Rooth (1992), where Parallelism is seen as a consequence of the theory of Focus.<sup>6</sup>

---

<sup>6</sup> The condition in (16) is stated as a condition on the *licensing* of phonological reduction. An alternative way of saying things is that phonological reduction is free. However, it yields certain presuppositions which need to be satisfied. When I say that Parallelism is not licensed, this could be read as a statement that focus presuppositions are not met.



- (16) **Parallelism:** Phonological reduction of a constituent  $\alpha$  is licensed only if at LF there is some constituent  $\beta$  that reflexively dominates  $\alpha$  and the discourse contains an antecedent LF,  $\beta_A$ , such that, either:<sup>7</sup>
- a. **Direct Parallelism:**  $\beta_A$  is an appropriate alternative to  $\beta$ . [That is to say  $[[\beta_A]]$  belongs to the focus value of  $\beta$ ;  $[[\beta_A]] \in F(\beta)$
  - or
  - b. **Indirect Parallelism:**  $\beta_A$  together with shared presuppositions entails an antecedent LF,  $\beta_{AC}$ , ( $\beta_A \approx \beta_{AC}$ ) and  $[[\beta_{AC}]] \in F(\beta)$ .<sup>8,9</sup>

The definition of Parallelism in (16) captures the asymmetries in (14) and (15). I will illustrate this with (15) and allow the reader to verify that the informal discussion of (14) is translatable to the language of (16). Consider (15b) and take  $\beta$  to be the sentence *BILL talked to many women*.  $F(\beta)$  is now the set of structured propositions of the form *x talked to many women*. There is no antecedent in the preceding discourse which is of this form, and therefore Direct Parallelism is not satisfied. However, Indirect Parallelism is satisfied. The Discourse includes the proposition that John talked to every woman he saw and this proposition, together with some shared presuppositions, entails the LF *John talked to many women*. Because *John talked to many women* denotes a (structured) proposition of the form *x talked to many women*, it is a member of  $F(\beta)$ . Consider now (15a) and take  $\beta$  to be the sentence *JOHN talked to every woman he saw*.<sup>10</sup>  $F(\beta)$  is now the set of structured

<sup>7</sup>  $\beta$  reflexively dominates  $\alpha$  if  $\beta$  dominates  $\alpha$  or  $\beta = \alpha$ .

<sup>8</sup> There are two differences between the definition given here and the one in Rooth (1992).

1. Rooth assumes that the condition governing phonological down-stressing doesn't care about structure. More specifically, he assumes that focus values contain sets of semantic objects (devoid of structure). I assume with Wold (1995; forthcoming) that structure is crucial for Parallelism (for both down-stressing and ellipsis). For this reason, I must assume that focus values contain structural information. They could either be sets of syntactic structures (in which case the condition that  $[[x]] \in F(y)$  would be replaced by the condition that  $x \in F(y)$ ) or sets of structured meanings (see Lewis 1972). For presentational purposes, I will assume the latter.

2. I ignore the cases in which the discourse antecedent denotes a *set* of structured propositions (e.g., the case of question answer-pairs). This is done for presentational purposes and bears no significance.

<sup>9</sup> The precise details of the definition of Direct Parallelism is not crucial for the purposes of this paper. More specifically, for the purposes of this paper it would be possible to adopt other definitions of Direct Parallelism that entail isomorphism in structure (e.g., the definition in Fiengo and May (1994)). What is crucial is that Parallelism could be satisfied also indirectly (via an accommodated proposition).

<sup>10</sup> I leave it for the reader to see that other choices of  $\beta$  do not affect the result.

propositions of the form *x talked to every woman he saw*. There is no antecedent in the preceding discourse which is of this form, and therefore Direct Parallelism is not satisfied. However, in contrast to what we saw in the case of (15b), Indirect Parallelism is also not satisfied. The Discourse includes the proposition that Bill talked to many women, and this proposition does not entail (not even with plausible accommodations) an LF of the form *x talked to every woman he saw*.

In this context, we can focus on cases in which  $\beta$  is a clause-denoting constituent which is preceded by another clause denoting constituent  $\beta_A$ . We can now see that the definition of Parallelism in (16) makes a clear distinction between  $\beta$  and  $\beta_A$ . Parallelism is licensed either if  $\beta_A$  is directly parallel to  $\beta$  ( $[[\beta_A]] \in F(\beta)$ ) or if there is some proposition entailed by  $\beta_A$  (together with additional plausible assumptions),  $\beta_{AC}$ , which is directly parallel to  $\beta$ ; it is sufficient for  $\beta_{AC}$  to be entailed by  $\beta_A$ ; it is not sufficient for it to be entailed by  $\beta$ . In the next section, I intend to show that this asymmetry in the definition of Parallelism accounts for the asymmetry we have seen in (3) and (4).

But before that, we have to go over Rooth's demonstration that indirect Parallelism (as defined in (16)) is relevant not only for phonological reduction but also for phonological deletion (i.e., ellipsis). A first (theoretical) observation is that constructions that involve ellipsis have an identical focus structure to parallel constructions that involve phonological reduction. The reason for this is fairly evident. Focus must be marked by pitch accent and pitch accent is inconsistent with both phonological deletion and reduction.<sup>11</sup> For this reason, both deleted and phonologically reduced phrases cannot be Focus-marked or contain a Focus-marked constituent.<sup>12</sup> In other words, if there is a Parallelism condition

---

<sup>11</sup> Hubert Truckenbrodt (p.c.) raises the following question. What is the difference between the phonological consequences of focus marking and the phonological consequences of other syntactic objects, say the phonological instructions that are contained in a lexical item? More specifically, why is it the case that an antecedent for ellipsis allows us to ignore the latter (yielding VP deletion) but not the former (which would yield a deleted VP with a Focus marked constituent)? The problem would disappear if we view Focus not as an instruction to phonology but rather as a consequence of phonology (as in Reinhart 1995).

<sup>12</sup> An obvious question to ask is whether a constituent that contains ellipsis or down-stressing can be dominated by an F-marked constituent with pitch accent outside the domain of ellipsis/down-stressing. If this were possible, there would be a danger that it would nullify the Parallelism requirement which is otherwise imposed by the lack of pitch accent in the elided/down-stressed constituent. However, I believe that various observations in Schwarzschild (in press) indicate that even if such a case were possible, there is no real danger that Parallelism will be nullified. Specifically, I believe that, although couched in a different view of focus, Schwarzschild presents strong arguments







*every teacher<sub>y</sub> [ $Q_x$  [ $x$  admires  $y$ ]]*, where  $Q$  is a variable ranging over possible denotations for DPs (i.e., generalized quantifiers). *A boy admires every teacher* denotes a structured proposition of the relevant form iff the object receives wide scope. Hence parallel scope assignment is ensured by Direct Parallelism. The obvious question is whether there are conditions under which Indirect Parallelism obviates this result (as it did in the case of the requirement of parallel variable binding.) In the next section I will claim that there are, and that this is the source of the asymmetry we've seen in the previous section.

However, first I would like to state a consequence of Parallelism which is relevant for the constructions that interest us. As I discussed above, the relevant constructions are those in which  $\beta$  is a clause-denoting constituent which is preceded by another clause-denoting constituent  $\beta_A$ . Furthermore, the relevant consequence of Direct Parallelism is that it ensures identical scopal positions for the DPs in the antecedent and the elided VP (i.e., structural isomorphism).<sup>15</sup> Indirect Parallelism could (under certain conditions) obviate the necessary requirement. This is stated in (23)

- (23) **Parallelism (Consequence of):** Phonological reduction or deletion is licensed only if, either:
- a. **Direct Parallelism:** The LF of a sentence that contains the elided/down-stressed material,  $\beta_E$ , is structurally isomorphic to a sentence that contains the antecedent,  $\beta_A$ .
  - or**
  - b. **Indirect Parallelism:**  $\beta_A$  (together with shared presuppositions) entails an LF,  $\beta_{AC}$ , and  $\beta_{AC}$  is structurally isomorphic to  $\beta_E$ .<sup>16</sup>

---

<sup>15</sup> Of course Parallelism requires more than structural isomorphism. It also requires identical lexical items in all positions that are not dominated by an F-marked constituent. This requirement will be irrelevant for most of the cases I will discuss, since for most cases I will only consider pairs of structures where identity of the relevant lexical items is satisfied. In the few cases in which the identity of lexical items will be relevant (e.g. (33b,34b)) I will return to the definition of Parallelism in (16).

<sup>16</sup> In the discussion that follows I will sometimes omit the mentioning of shared presuppositions and simply talk about entailments. This sloppy way of talking is motivated for the sake of brevity, and does not affect the points that are being made.

### 3.3. The Asymmetry of the ESG follows from the Asymmetry in Parallelism

We are now in a position to see that the asymmetry which is built into the definition of Parallelism can account for the asymmetry in (3) and (4).

- (3) a. Some boy admires every teacher and Mary does, too.  $(\exists > \forall) *(\forall > \exists)$   
 b. Mary admires every teacher and some boy does, too.  $(\exists > \forall) (\forall > \exists)$
- (4) a. Some boy admires every teacher and Mary *admires every teacher*, too.  
 $(\exists > \forall) *(\forall > \exists)$   
 b. Mary admires every teacher and some boy *admires every teacher*, too.  
 $(\exists > \forall) (\forall > \exists)$

The basic idea is that Economy is blind to Parallelism (as assumed in Chapter 2). For this reason, Economy blocks QR over the subject in both the second conjunct of (a) and the first conjunct (b). If Direct Parallelism were the only way to satisfy Parallelism, Inverse Scope would be blocked in both (a) and (b). The contrast between (a) and (b) is related to Indirect Parallelism. In (b), but not in (a), Indirect Parallelism allows Inverse Scope in the Scopally Informative conjunct even though Inverse Scope is not allowed in the Scopally Uninformative conjunct.

Now, to the actual details. Given Economy Inverse Scope is not allowed in the Scopally Uninformative conjuncts in both (a) and (b). Let's see what happens if the Scopally Informative conjuncts receive Inverse Scope. The obvious observation is that  $\beta_A$  and  $\beta_E$  are not structurally isomorphic and Direct Parallelism is not satisfied. This is seen in (3'  $(\forall > \exists)$ ):

- (3'  $(\forall > \exists)$ )
- a.  $*\beta_A$ : every teacher  $\lambda y$  (some boy  $\lambda x$  (x admires y)) and  
 $\beta_E$ : Mary  $\lambda x$  every teacher  $\lambda y$  (x admires y)<sup>17</sup>
- b.  $\beta_A$ : Mary  $\lambda x$  every teacher  $\lambda y$  (x admires y) and  
 $\beta_E$ : every teacher  $\lambda y$  (some boy  $\lambda x$  (x admires y))

<sup>17</sup>As mentioned in Chapter 2, I assume that non-subject quantifiers always move short-distance for interpretability (type mismatch; see Heim and Kratzer 1998). However, this assumption is not crucial for anything I say in this chapter.

The options that are made available by Indirect Parallelism are less obvious. The goal is to show that in (b), but not in (a), Indirect Parallelism is satisfied (or, in other words, that the distribution of \* is justified). Let's start with (3'b). The idea is that if Mary admires every teacher, it follows that for every teacher there is a girl who admires that teacher (namely Mary). In other words,  $\beta_A$  entails the LF,  $\beta_{AC}$ , which is structurally isomorphic to  $\beta_E$ . Therefore, Indirect Parallelism is satisfied. This is summarized in (3'b<sub>( $\forall>\exists$ )</sub>).

(3'b<sub>( $\forall>\exists$ )</sub>) Direct Parallelism is not satisfied. However Indirect Parallelism is satisfied:

1. ( $\beta_A$ ) Mary  $\lambda x$  every teacher  $\lambda y$  ( $x$  admires  $y$ )  $\approx >$   
 $(\beta_{AC})$  every teacher  $\lambda y$  (some girl  $\lambda x$  ( $x$  admires  $y$ ))
2. every teacher  $\lambda y$  (some girl  $\lambda x$  ( $x$  admires  $y$ ))  
 is an appropriate antecedent for  
 every teacher  $\lambda y$  (some boy  $\lambda x$  ( $x$  admires  $y$ ))

Now let's move to (3'a). The reason Indirect Parallelism cannot help (3'a) is related to the asymmetry of the condition. In (3'a), we need the entailment relation to go in the other direction from the one which was utilized in (3'b). In other words, we need an LF with wide scope for the subject to follow from the LF *every teacher  $\lambda y$  (some boy  $\lambda x$  ( $x$  admires  $y$ ))*. However, there is no (obvious) LF that could play this role. (See 3.4., where more is said about the set of LFs that are available for accommodation.) An LF with a ( $\forall>\exists$ ) interpretation does not entail any (obvious) LF in which the relative scope of the two quantifiers (associated with the two argument positions) is reversed. This is summarized in (3'a<sub>( $\forall>\exists$ )</sub>).

(3'a<sub>( $\forall>\exists$ )</sub>) Direct Parallelism is not satisfied. Indirect Parallelism is not satisfied either.

There is no LF entailed by *every teacher  $\lambda y$  (some boy  $\lambda x$  ( $x$  admires  $y$ ))* which can serve as antecedent for *Mary  $\lambda x$  every teacher  $\lambda y$  ( $x$  admires  $y$ )*

In particular,

*every teacher  $\lambda y$  (some boy  $\lambda x$  ( $x$  admires  $y$ ))*  $\not\approx >$

*some boy  $\lambda x$  every teacher  $\lambda y$  ( $x$  admires  $y$ )*

This account of the asymmetry makes no reference to the relative structural position of the first and the second conjunct, and thus does not run into any of the problems discussed in 3.1.1. Furthermore, the account relies crucially on the semantic properties of the particular quantificational expressions (specifically, their ability to contribute to various entailment







My aim is to show that this time Indirect Parallelism cannot obviate the result in either order (in other words that both (12'a<sub>more than 3 > not</sub>) and (12'b<sub>more than 3 > not</sub>) are ruled-out). Let us start with (12'a<sub>more than 3 > not</sub>). The question is whether there is an LF,  $\beta_{AC}$ , which is entailed by  $(\beta_A)$  *More than 3 languages  $\lambda x$  (not (K.H. speaks x))* and which is structurally isomorphic to  $(\beta_E)$  *yes (R.P. speaks more than 3 languages)*. There is definitely no obvious LF of this sort. In particular, reversing the scope of the object QP and the negative operator in  $(\beta_A)$  would not yield a proposition that is entailed by  $(\beta_A)$ . Therefore, Indirect Parallelism is not satisfied:

(12'a<sub>more than 3 > not</sub>)

Direct Parallelism is not satisfied. Indirect Parallelism is not satisfied either.

There is no LF entailed by *More than 3 languages  $\lambda x$  (not (K.H. speaks x))* which can serve as antecedent for *yes (R.P. speaks more than 3 languages)*.

In particular

*More than 3 languages  $\lambda x$  (not (K.H. speaks x))*  $\not\Rightarrow$   
*not (K.H. speaks more than 3 languages)*

consider now (12'b<sub>more than 3 > not</sub>). Once again there is no proposition to accommodate in order to satisfy Indirect Parallelism. A potential proposition to accommodate would result from reversing the scope of the affirmative operator and the object QP in the antecedent. However, such a reversal would result in an illegitimate LF (an LF that violates Economy):

(12'b<sub>more than 3 > not</sub>)

Direct Parallelism is not satisfied. Indirect Parallelism is not satisfied either

There is no LF entailed by *yes (R.P. speaks more than 3 languages)* which can serve as antecedent for *More than 3 languages  $\lambda x$  (not (K.H. speaks x))*

Note that

*more than 3 languages  $\lambda x$  yes (R.P. speaks x)* is not generated by the grammar because it violates Economy.

I now move to the other case discussed in 3.1.2. in which scope disambiguation goes in both directions. The aim is to show that in (13), just as in (12), Indirect Parallelism cannot obviate the results of Direct Parallelism.

- |      |    |  |  |
|------|----|--|--|
| (13) | a. | A boy talked to every teacher and a girl did, to Jane. | $(\exists > \forall) *(\forall > \exists)$ |
|      | b. | A girl talked to Jane and a boy did, to every teacher. | $(\exists > \forall) *(\forall > \exists)$ |

As is standard by now, we focus on the case in which an SSO applies in the Scopally Informative sentence but (given Economy) not in the Scopally Uninformative sentence. The aim is to show that Indirect Parallelism cannot obviate the violation of Direct Parallelism:

(13'<sub>(v>3)</sub>)

- a.  $*\beta_A$  : every teacher  $\lambda y$  (a boy  $\lambda x$  (x talked to y))  
 $\beta_E$  : a girl  $\lambda x$  Jane  $\lambda y$  (x talked to y)<sup>19</sup>
- b.  $*\beta_A$  : a girl  $\lambda x$  Jane  $\lambda y$  (x talked to y)  
 $\beta_E$  : every teacher  $\lambda y$  (a boy  $\lambda x$  (x talked to y))

Lets start with (13'b<sub>(v>3)</sub>). The reason Indirect Parallelism is not satisfied is that there is no obvious LF,  $\beta_{AC}$ , that is entailed by the proposition that a girl talked to Jane and would be structurally isomorphic to ( $\beta_E$ ) *every teacher  $\lambda y$  (a boy  $\lambda x$  (x talked to y))*. A potential LF of this sort is one in which the scopal relationship of the two QPs in  $\beta_A$  is reversed. But, as in the case of (12b), this is not a legitimate LF since it does not obey Economy (hence is not generated by the grammar):

(13'b<sub>(v>3)</sub>)

Direct Parallelism is not satisfied. Indirect Parallelism is not satisfied either  
 There is no LF entailed by *a girl  $\lambda x$  Jane  $\lambda y$  (x talked to y)* which can serve as  
 antecedent for *every teacher  $\lambda y$  (a boy  $\lambda x$  (x talked to y))*  
 Note that  
*Jane  $\lambda y$  a girl  $\lambda x$  (x talked to y)* is not generated by the grammar because it  
 violates Economy.

The explanation of (13'a<sub>(v>3)</sub>) is at the moment incomplete. Given everything that was said up to now, we would expect (13'a<sub>(v>3)</sub>) to satisfy Indirect Parallelism. The reason is that ( $\beta_A$ ) [*every teacher  $\lambda y$  (a boy  $\lambda x$  (x talked to y))*] (with shared presuppositions) entails the LF *a boy  $\lambda x$  a teacher  $\lambda y$  (x talked to y)* and the latter is structurally isomorphic to  $\beta_E$ . If *a boy  $\lambda x$  a teacher  $\lambda y$  (x talked to y)* were an accommodatable LF (if it could stand for  $\beta_{AC}$  in the definition of Indirect Parallelism), we would predict (13a) to allow Inverse Scope. What I would like to suggest is that independent considerations block the

---

<sup>19</sup> *Jane* is outside the VP for reasons that are independent of scope. See Lasnik (1995) and Jayaseelan (1990).



I would like to propose and motivate a constraint which, although different, is inspired by Tancredi's suggestion. The constraint is based on the idea that accommodation is a last resort. As such, I suggest that it must be minimal, given what we can call "accommodation seeking material". The accommodation seeking material is the material in  $\beta_E$  which is presupposed (i.e. not dominated by a focus-marked constituent; henceforth non-F-dominated) and is nevertheless not directly available (i.e. it is absent in  $\beta_A$  [modulo Vehicle Change]). The intuition behind the proposal I am making is that only accommodation-seeking material indicates that accommodation is necessary; other material is either already present in the antecedent or F-dominated (thus indicated as "new information"):

- (24) Accommodation of  $\beta_{AC}$  must be minimal given the accommodation seeking material  $\alpha$ .
- (25)
  - a. An accommodation  $\beta_{AC}$  is minimal given  $\alpha$ , if there is no alternative accommodation to  $\beta_{AC}$ ,  $\beta'$ , such that  $\beta'$  contains  $\alpha$  and  $\beta'$  is closer to  $\beta_A$  than  $\beta_{AC}$  is.
  - b.  $\beta'$  is closer to  $\beta_A$  than  $\beta_{AC}$  is, when the accommodated material of  $\beta'$  is a proper subset of the accommodated material in  $\beta_{AC}$ .
  - c. The accommodated material of an accommodation  $\beta$  consist of the lexical material which is present in  $\beta$  and absent in  $\beta_A$ .

For the purposes of this paper, I will make use of a consequence of (24):

- (26) Accommodation of  $\beta_{AC}$  must have a trigger.
- (27) Accommodation has a trigger when  $\beta_E$  contains accommodation seeking material, i.e. when  $\beta_E$  contains non-F-dominated material which is absent in  $\beta_A$ .

This consequence follows trivially; If  $\beta_E$  contains no accommodation seeking material the minimal accommodation will be the trivial one, i.e.  $\beta_{AC} = \beta_A$ .

**3.4.1. Evidence for the constraint on accommodation:** I would like to show that this constraint makes some surprising predictions for the availability of ESG obviations. But first I would like to provide some independent evidence for the constraint. Specifically, I would like to show that both ellipsis and phonological reduction constructions which rely on Indirect Parallelism must satisfy (26).



Why, then, is sloppy identity blocked in (28b)? (26) provides the answer.  $\beta_E$  lacks non-F-dominated material which is absent in  $\beta_A$ , and thus there is no trigger for accommodation.

In (29) we see that in certain cases (26) restricts phonological reduction as well as ellipsis. (29a,b) are the phonological reduction analogs of (28a,b). (29c) shows that in order to get the sloppy identity when (26) is not satisfied, it is necessary to focus the pronoun and thus make accommodation unnecessary. (In (29c),  $F(\beta_E)$  is the set of structured propositions of the form  $x \vee I \text{ was bad mouthing } y$ .  $\beta_A$  denotes a structured proposition of the relevant form if we make the independently needed assumption that *convinced Mary* is a possible alternative to *deny*.)

- (29) a. First you convinced Mary<sub>1</sub> that I was bad-mouthing her<sub>1</sub>, and then  
[Fred<sub>2</sub>]<sub>F</sub> *believed that I was bad-mouthing him<sub>2</sub>*.  
b. \*First you convinced Mary<sub>1</sub> that I was bad-mouthing her<sub>1</sub>, and then [Fred<sub>2</sub>]<sub>F</sub>  
[denied]<sub>F</sub> *that I was bad-mouthing him<sub>2</sub>*.  
c. First you convinced Mary<sub>1</sub> that I was bad-mouthing her<sub>1</sub>, and then [Fred<sub>2</sub>]<sub>F</sub>  
[denied]<sub>F</sub> *that I was bad-mouthing HIM<sub>2</sub>*.<sup>20</sup>

To complete the argument for (26), it is important to show (empirically) that the potential  $\beta_{AC}$  (the LF *Mary<sub>1</sub> believed that I was bad-mouthing her<sub>1</sub>*) is Parallel to  $\beta_E$  in (28b, 29b). In other words, it is important to provide empirical evidence that without (26), Indirect Parallelism would predict Sloppy identity in these constructions. This evidence is provided in (30). In the construction in (30), there is no need for accommodation in order to achieve sloppy identity. Therefore, (26) is irrelevant and sloppy identity is possible.

- (30) a. [Sue<sub>2</sub>]<sub>F</sub> *believed that I was bad-mouthing her<sub>2</sub>*. However, [Jane<sub>3</sub>]<sub>F</sub>  
[denied]<sub>F</sub> *that I was <bad-mouthing her<sub>3</sub>>*.

---

<sup>20</sup> A minimal variation on (29b) can show that the triggering requirement in (26) is insufficient. In (i-ii), the non-F-dominated constituents provide a trigger which is nonetheless insufficient for the accommodation which would yield sloppy identity. The unavailability of the accommodation in (i-ii) follows from the more complete minimality condition in (24).

- (i) First John convinced Mary<sub>1</sub> that Bill was bad-mouthing her<sub>1</sub>, and then [Sue<sub>2</sub>]<sub>F</sub>  
[denied]<sub>F</sub> *that this idiot was bad-mouthing \*her<sub>2</sub>/HER<sub>2</sub>*  
(ii) First John convinced Mary<sub>1</sub> that every boy was bad-mouthing her<sub>1</sub>, and then  
[Sue<sub>2</sub>]<sub>F</sub> [denied]<sub>F</sub> *that many boys were bad-mouthing \*her<sub>2</sub>/HER<sub>2</sub>*



- b. [Sue<sub>2</sub>]<sub>F</sub> believed that I was bad-mouthing her<sub>2</sub>. However, [Fred<sub>3</sub>]<sub>F</sub> [denied]<sub>F</sub> that I was bad mouthing him<sub>3</sub>.

The evidenced is re-enforced in (31). (31a, b) are very similar to (28b, 29b), respectively. The only difference is that in (31a, b), the LF which must be accommodated in (28b, 29b) to license sloppy identity is part of the discourse. Therefore in (31a, b) accommodation is unnecessary and sloppy identity is available.

- (31) a. First John convinced Mary<sub>1</sub> that I was bad-mouthing her<sub>1</sub>, and then [Sue<sub>2</sub>]<sub>F</sub> believed that I was <bad-mouthing her<sub>2</sub>>. But lucky for me at least [Jane<sub>3</sub>]<sub>F</sub> [denied]<sub>F</sub> that I was <bad-mouthing her<sub>3</sub>>.  
 b. First John convinced Mary<sub>1</sub> that I was bad-mouthing her<sub>1</sub>, then [Fred<sub>2</sub>]<sub>F</sub> believed that I was bad-mouthing him<sub>2</sub>. But lucky for me at least [Jane<sub>3</sub>]<sub>F</sub> [denied]<sub>F</sub> that I was bad-mouthing her<sub>3</sub>.

An Independent piece of evidence for the constraint in (26) comes from the paradigm in (32). I wouldn't pretend that I have a complete understanding of this paradigm. Nevertheless, I think that it safe to take it as evidence for (26).

- (32) a. \*What John said revealed my identity. [BILL]<sub>F</sub> did, too.  
 b. Bill revealed my identity. *What* [John]<sub>F</sub> said did, too.

Consider first the unacceptability of (32a). This unacceptability is most likely the result of the elided VP and its antecedent being non-identical. [Specifically, it seems reasonable to assume that the VP *revealed my identity* is different when it takes the animate subject *Bill* and the inanimate subject *what John said*. It is possible, for example, that the theta role assigned to the subject is different in the two cases (*agent* in one case and *causer* in the other).] This non-identity makes  $\beta_A$  a bad alternative to  $\beta_E$ . ( $[[\beta_A]] \notin F(\beta_E)$ .) Why then is (32b) acceptable. I suggest that in (32b) it is possible to accommodate the LF  $\beta_{AC}$  (= *What Bill said revealed my identity*) and that  $[[\beta_{AC}]] \in F(\beta_E)$ . If this suggestion is correct, there must be something that blocks a similar accommodation in (32a) (*What John said revealed my identity*  $\approx$  *John revealed my identity*). (26) can serve this purpose, since  $\beta_E$ , which

doesn't contain novel material which is non-F-dominated has no trigger for accommodation.<sup>21</sup>

**3.4.2. Consequences for the ESG:** It is now time to see whether (26), when combined with Parallelism and Economy, makes any interesting predictions for the pattern of disambiguation in ellipsis. Consider the constructions in (33) and (34) with unmarked pronunciation. (Pitch accent is represented with an acute accent.) The ordering account from Fox (1995a) predicts anti-disambiguation in all of the constructions in (33-34). However, this predictions is not borne out in the (b) cases.

- (33) a. Jon is likely to win the lottery. Someone from New York is, too.  
 b. Jon is likely to win the lottery. Two people from New York are, too.  $(\exists > \text{likely}) (\text{likely} > \exists)$   
 $(\exists 2 > \text{likely}) *(\text{likely} > \exists 2)$
- (34) a. John likes every teacher. At least one girl does, too.  $(\exists > \forall) (\forall > \exists)$   
 b. John likes every teacher. At least two girls do, too.  $(\exists 2 > \forall) *(\forall > \exists 2)$

This contrast follows from the account provided in this chapter once the constraint on Indirect Parallelism in (26) is taken into account. To see this, we have to know what is the focus structure of  $\beta_E$ . Given that pitch-accent is on the most embedded constituent in the DP, there are various possibilities for F-marking (See Jackendoff (1972). For more recent theories of stress assignment, see Cinque (1993) Selkirk (1995) and Truckenbrodt

<sup>21</sup> It seems to me that the correct formulation of (24/26) should allow accommodation in (i) given the down-stressing of the subject. However, the present formulation is insufficient because the non-F-dominated material in (i) is present in  $\beta_A$  (modulo Vehicle Change).

- (i) a. What John said revealed my identity. I wish *he* hadn't.  
 b. What he said revealed my identity. *he* didn't plan to.

This requires a modification in the definition of accommodation-seeking material:

- (27') Accommodation has a trigger when  $\beta_E$  contains accommodation seeking material, i.e. when  $\beta_E$  contains non-F-dominated material which is not present in  $\beta_A$  in a *parallel structural position*.

It is interesting to see whether Indirect Parallelism together with the constraint on accommodation can play a role in accounting for other cases in which an antecedent and elided VP are non-identical (e.g. Passive/active counter-parts). For discussion, see Hardt (1992), Johnson (1997), Tancredil (1992: 130), Wold (forthcoming), among others.

(1995)). These possibilities can be divided into two: F-marking can target (a) the whole DP or (b) a smaller constituent which excludes the quantificational material (*someone* in (33a) and *at least one* in (34a)).

Let's consider how Economy and Parallelism determine the LF representations of (33-34) given the constraint on Indirect Parallelism in (26). I will focus on (34) and leave it for the reader to see that the account extends to (33). In section 3.3., we mentioned that anti-disambiguation is predicted for (34a). The idea was that disambiguation, which in standard cases is predicted by Economy and Parallelism, is obviated by the availability of Indirect Parallelism. ( $\beta_A$  (=John likes every teacher) entails  $\beta_{AC}$  (=at least one boy likes every teacher with Inverse Scope) and  $\beta_{AC}$  is structurally isomorphic to  $\beta_E$  (=At least one girl likes every teacher with Inverse Scope)

However, we now have to show that the constraint on Indirect Parallelism in (26) is satisfied as well. In other words, we have to show that there is some material in  $\beta_E$  which is non-F-dominated and is absent from  $\beta_A$ . This is the case in (34a), if there is narrow focus on the Common Noun *girl* (as in (34'a1)), but isn't the case if there is focus on the whole DP (as in (34'a2)).

- (34'a) 1. [ $\beta_A$  John likes every teacher]. [ $\beta_E$  At least one [girl] <sub>F</sub> does, too]. ( $\exists > \forall$ ) ( $\forall > \exists$ )  
 2. [ $\beta_A$  John likes every teacher]. [ $\beta_E$  [At least one girl] <sub>F</sub> does, too]. ( $\exists > \forall$ )  $*(\forall > \exists)$

Anti-disambiguation is predicted for (34a), because (34'a1) is a possible focus structure.<sup>22</sup>

Now consider (34b) and its two possible focus structures in (34'b). When there is narrow focus on *girls* as in (34'b1), there is no way for Parallelism to be satisfied;  $F(\beta_E)$  is the set of structured propositions of the form *Every teacher* <sub>y</sub> *at least two*  $N_x$  *x like y*, and  $\beta_A$  does not entail a structured proposition of this form; In particular,  $\beta_{AC}$  (above) does not belong to  $F(\beta_E)$  because it contains the wrong numeral expression.

- (34'b) 1. [ $\beta_A$  John likes every teacher].  $*[\beta_E$  At least two [girls] <sub>F</sub> do, too].

<sup>22</sup> As far as minimality is concerned, we have to show that the accommodated material in  $\beta_{AC}$  is minimal given the trigger (see note 20.) This is only relevant if there is a trigger (i.e. only in (34'a1)). To see whether  $\beta_{AC}$  (=at least one boy likes every teacher with Inverse Scope) is minimal we have to check whether there is some other accommodatable proposition,  $\beta'$ , which contains the trigger *at least one* and nevertheless contains less accommodated material than  $\beta_{AC}$ . (I.e., the accommodated material must consist of the words *at least one* and of nothing else.) This is obviously impossible, and I therefore conclude that  $\beta_{AC}$  is a possible accommodation in the case of (34'a1).



structured proposition in  $F(\beta_E)$ . The relevant difference between (37b) and (34b) is that in (34b) the only way to satisfy (26) is by excluding the numeral *two* from the F-marked constituent, and such a focus structure makes it impossible to satisfy Parallelism. In (37b), by contrast, it is possible to focus the numeral and not the Common Noun, thus making the accommodatable LF an appropriate alternative to  $\beta_E$ .

Let us now return to the puzzle with which we ended the previous section. In example (13) repeated below, disambiguation holds in both directions. This was not predicted by the ordering account, but followed from the proposal made in this chapter with a small residue.

- (13) a. A boy talked to every teacher and a girl did, to Jane.  $(\exists > \forall) *(\forall > \exists)$   
 b. A girl talked to Jane and a boy did, to every teacher.  $(\exists > \forall) *(\forall > \exists)$

The remaining problem was to explain why there is disambiguation in (13a). Specifically, even on the  $(\forall > \exists)$  interpretation,  $\beta_A$  entails the LF *A boy talked to a teacher* (to which Economy assigns Surface Scope) and this LF structurally isomorphic to the LF determined by Economy for  $\beta_E$ .

Once we take into account the constraint on accommodation in (26), the problem is overcome. To see this, let us look at the focus structure of (13a) in (13'a). As we can see,  $\beta_E$  doesn't contain any non-F-dominated material which is absent from  $\beta_A$ . In other words, there is no trigger for accommodation and (26) is not satisfied.

- (13'a)  $[\beta_A \text{ A boy talked to every teacher}]$ . and  $[\beta_E [a [girl]_F] \text{ did, to } [Jane]_F]$ .  $(\exists > \forall) *(\forall > \exists)$

This explanation makes a surprising prediction. Specifically, it is predicted that if we change (13a) minimally so that (26) is satisfied, the predictions of Economy will be obviated by Indirect Parallelism. (In other words, once again there will be anti-disambiguation.) Furthermore, it is predicted that (13b) will not be affected by the change, since (as demonstrated in the previous section) in (13b) there is no way for Indirect Parallelism to obviate the predictions of Economy even when the constraint in (26) is ignored.

This prediction is borne out below:

- (38) a.  $[\beta_A$  At least one Republican senator talked to every CEO from NY], and  
 $[\beta_E$  [a [Democrat]<sub>F</sub>] did, to [Bill Clinton]<sub>F</sub>].  $(\exists > \forall) *(\forall > \exists)$   
 b.  $[\beta_A$  At least one Republican senator talked to every CEO from NY], and  
 $[\beta_E$  [a [Democrat]<sub>F</sub>] did, to [a CEO from [Boston]<sub>F</sub>].  $(\exists > \forall) ?(\forall > \exists)$   
 c.  $[\beta_A$  One Democratic senator talked to a CEO from Boston], and  
 $[\beta_E$  [a [Republican]<sub>F</sub>] did, to [[every]<sub>F</sub> CEO from [NY]<sub>F</sub>].  $(\exists > \forall) *(\forall > \exists)$

(38a) is identical (in all relevant respects) to (13a). (38b) differs minimally in a way that allows the satisfaction of (26). (The non-F-dominated material *a CEO from* indicates the need for accommodation.) I suggest it is for this reason that the  $(\forall > \exists)$  interpretation is not so bad in (38b). Finally, in (38c) there is no way to satisfy Indirect Parallelism. As we've seen in the previous section, there is no LF which is entailed by  $\beta_A$  and would be isomorphic to the LF that yields the  $(\forall > \exists)$  interpretation for  $\beta_E$ . The contrast between (38b) and (38c) is very surprising from the perspective of the ordering account, and follows straightforwardly from the account proposed here.<sup>23</sup>

### 3.5. Conclusion

This chapter re-examined data which (in Fox 1995) I have taken to argue that Economy is sensitive to the demands of Parallelism. The data involved a Putative asymmetry in the nature of the ESG; a second Scopally Uninformative conjunct can disambiguate an Scopally Informative first conjunct but not vice versa. I have argued that the apparent asymmetry is not real. There are cases in which disambiguation holds in both directions (12-13) and cases in which it only holds in the opposite direction from that assumed in Fox (1995a), (38). I have suggested that in principle Economy predicts disambiguation in both directions, but that sometimes the effects of Economy are obviated by Indirect Parallelism. I have argued for this suggestion (a) by demonstrating a correlation between patterns of

<sup>23</sup> The careful reader must have noticed that, for examples such as (i), Indirect Parallelism predicts Surface Scope in  $\beta_A$  with Inverse Scope in  $\beta_E$ .

- (i) Many CEOs from NY talked to every senator. *At least one* CEOs from [Boston]<sub>F</sub> did, too.

My impression is that the prediction is false, though establishing this is not a trivial matter (see Abusch (1994), among others). My hope is either that my impression is wrong or that there is a natural way to state (24) which would avoid the problem

entailment and anti-disambiguation (Anti-disambiguation is possible only when the entailment needed for Indirect Parallelism holds.) and (b) by verifying the predictions of an independently motivated condition on Indirect Parallelism.





Binding Informative sentences are sentences which are semantically distinct under local and non-local variable binding. Binding Uninformative sentences, by contrast, are semantically identical under the two binding relationships. In the sentence in (4), for example, local binding as in (4'a), and non-local binding as in (4'b) are semantically identical.

- (4) Every boy thought that he likes his mother. *(Binding Uninformative)*
- (4') a. Every boy  $\lambda x(x \text{ thought that } x \lambda y(y \text{ likes } y\text{'s mother}))$ . *(local binding)*
- b. Every boy  $\lambda x(x \text{ thought that } x \lambda y(y \text{ likes } x\text{'s mother}))$ . *(non-local binding)*

In such Binding Uninformative sentences, it is impossible to use intuitions about meaning in order to determine binding relationships. However, I would like to suggest that we can use Parallelism for this determination. Specifically I would like to suggest that in an ellipsis environment a Binding Informative sentence can serve as a Binding Detector, which can indicate the binding properties of Binding Uninformative sentences. Specifically, using Parallelism I would like to argue for the following generalization:

- (5) Binding Uninformative sentences are restricted to local binding.

The argument for the generalization about variable binding in (5) is thus identical to one of the argument for the generalization about scope in (2). Furthermore, the two generalizations are very similar. It is therefore plausible that they have a unified explanation. More specifically, the generalization in (2) suggests that the operation that yields Inverse Scope cannot be semantically vacuous. We would like to say that the generaliation in (5) suggests that there is something that allows for non-local variable binding (preferably something similar to the operation that yields Inverse Scope) and that this 'something' cannot be semantically vacuous. The problem for unification, however, is that it is not very common to think of non-local variable binding as a result of an operation (though, see Kayne 1998). Nevertheless, I will provide a tentative statement of an economy condition that applies both to scope and to variable binding. However, this statement will be inconsistent with the details of Scope-Economy as they have been spelled-

out in chapter 2. I will leave this inconsistency as an unresolved problem.

The idea that non-local binding is impossible when it is semantically identical to local-variable binding was suggested on independent grounds by Heim (1993). Modifying work by Reinhart (1983), Heim demonstrates that certain obviation of Binding Theory (BT) can be accounted for by this suggestion. Given the logic outlined above, we can say that Binding Theory, just like Parallelism, enables us to distinguish between (interpretively identical) representation, and in doing so it vindicates the generalization in (5).

The argument that I present for the generalization in (5) relies heavily on work by Heim and Reinhart. This intellectual debt is reflected in Fox (1998a) which is what follows with minor revisions.

Heim (1993) argues for a locality condition on variable binding, which I will call Rule H. Rule H states that a variable,  $x$ , cannot be bound by an antecedent,  $\alpha$ , in cases where a more local antecedent,  $\beta$ , could bind  $x$  and yield the same semantic interpretation.<sup>1</sup> We can view Rule H as an optimality consideration. In particular, we can say that grammar prefers representations in which a variable is locally bound, and that (given this preference) Optimality blocks non-local binding. This statement, like any other statement about Optimality, can be evaluated only when accompanied by a claim about the nature of the set of candidates out of which the most optimal is chosen (the reference set). If we say that the reference set is restricted to linguistic objects that have identical semantic interpretations, we get an optimality characterization of Rule H. This partially semantic view of the reference set was taken in independent work on Economy by Golan (1993), Reinhart (1995) and Fox (1995a). The obvious question is whether Rule H and the relevant principle of Economy call for unification. This chapter has two goals. The first is to provide new evidence in favor of Rule H. The second is to draw possible guidelines for unification.

My evidence in support of Rule H is based on the observation that it can solve a puzzle having to do with the interpretation of pronouns in VP ellipsis constructions. The puzzle, which was first presented in Dahl (1974), and has since been discussed by many researchers (cf. Sag (1976), Ristad (1992), Kehler (1993) and Fiengo and May (1994)), relates to the distribution of strict and sloppy ambiguities in constructions that involve two

---

<sup>1</sup> Heim proposed Rule H as an extension (and modification) of work by Reinhart (Reinhart (1983), Grodzinsky and Reinhart (1993)). Whether it is an appropriate extension is a controversial matter which bears on many issues outside the scope of this chapter. For the purposes of this chapter it is enough that Rule H is well motivated. The role that it plays relative to the various principles proposed by Reinhart doesn't bear directly on issues discussed here. Though see note 15.

or more pronouns linked to the same antecedent. What is puzzling is that the first pronoun cannot be interpreted by strict identity if the second pronoun is interpreted by sloppy identity.<sup>2</sup> This is exemplified in (6) and (7).<sup>3</sup> If the first pronoun is interpreted by strict identity, the second one must be interpreted by strict identity as well, hence the unavailability of the interpretation in (7).

- (6) John said that he<sub>i</sub> liked his<sub>i</sub> mother.  
 (7) a. Bill did too.  
     a. Bill said that John likes John's mother. (*Strict, Strict*)  
     b. Bill said that Bill likes Bill's mother. (*Sloppy, Sloppy*)  
     c. Bill said that Bill likes John's mother. (*Sloppy, Strict*)  
     d. \*Bill said that John likes Bill's mother. (*Strict, Sloppy*)

I will show that Rule H, interacting with an independent requirement on parallelism in ellipsis constructions, provides a solution for Dahl's puzzle. This solution will in turn generate some new predictions which I will attempt to corroborate. Specifically, I will demonstrate that the following generalization holds: in an ellipsis construction, a pronoun can be non-locally bound only if non-local binding is semantically distinct from local binding, both in the elided and the antecedent VP. This generalization follows from the interaction I suggest exists between Rule H and Parallelism. Further, it is exactly parallel to another generalization which we have seen in Chapter 2: in an ellipsis construction, a quantifier can have non-local scope only if non-local scope is semantically distinct from local scope, both in the elided and the antecedent VP. I will conclude this chapter with an outline of a unified account for both generalizations.

The chapter is organized as follows. In section 4.1., I present the basics of the solution for Dahl's puzzle and the nature of the new predictions that it makes. In order to test whether the predictions hold, I have to present some background. In particular, I have to go over some of Heim's original arguments in favor of Rule H. This is done in section 4.2. In section 4.3. and 4.4., I demonstrate that the predictions of my proposal are in fact borne out. In section 4.5., I present an outline for a unification of Rule H with the semantically-based economy condition that I proposed in Fox (1995a). This unification is

---

<sup>2</sup> In the course of this chapter, it will become clear that the generalization is somewhat different.

<sup>3</sup> Indexation is meant as a notational device. This device should not be mistaken as an indication of any theoretical persuasion. For my assumptions regarding the *computational* representation of variable binding, see section 2.

inconsistent with the local implementation of this condition which I defended in chapter 2. This inconsistency is left as an open problem.

#### 4.1. The Basic Proposal

In this section I demonstrate that Rule H can be used to explain Dahl's puzzle. In particular, I demonstrate that the restriction on the interpretation of (7) (the unavailability of (7d)) follows from an interaction of two principles. One principle is Rule H. The other principle, which is needed on independent grounds, requires parallelism between an elided VP and its antecedent. The basic idea is that Rule H excludes a representation (6) which would be parallel to (7d).<sup>4</sup>

**4.1.1 Rule H:** Rule H is an optimality condition which prefers representations where a pronoun is locally bound over representations where binding is more distant. In other words, Rule H requires that a pronoun be bound by the most local antecedent available. In this respect Rule H is very similar to other locality conditions (cf. Rizzi (1990) and Aoun (1987)). However, it is also very different.

Take, for example, condition A of the binding theory.<sup>5</sup> This condition requires (under certain formulations) that an anaphor be bound by the most local antecedent available, hence the Specified Subject Condition. However, Rule H, obviously, can't require from pronouns what condition A requires from anaphors. The difference between Rule H and condition A is related to the way they interpret the word *available*. To see this, consider the way condition A would rule out a sentence such as (8). (8b) is bad because

---

<sup>4</sup> The proposal made here is very similar, though not identical, to a proposal made in Kehler (1993), which was recently brought to my attention. Kehler proposes that Dahl's observation is the result of an interaction of a copying algorithm (for ellipsis resolution) and a locality condition on linking (Kehler's (26)). However, because Kehler's locality condition is not an optimality condition, because it doesn't make reference to the grammatical and (in particular) the interpretive status of competitors, it doesn't generate the predictions that my proposal does.

<sup>5</sup> Condition A of the binding theory might not be the best example, because it is not at all clear that this condition is best understood as an optimality condition that involves comparison, though certain ideas in the literature are suggestive (for example the definition of accessible SUBJECT suggested in Chomsky (1981) and developed in an optimality-like fashion in Aoun (1987), the definition of a CFC in Chomsky (1986), or the relationship between 'blocking' and 'antecedence' in Burzio (1998)). I use condition A to illustrate the nature of Rule H only because it, too, relates to anaphora (broadly construed).

there is an available representation, (8a), in which the anaphor is bound by a more local antecedent.

- (8) a. John said that Bill liked himself.  
           b. \*John said that Bill liked himself.

Consider now the way Rule H applies to pronouns. (9b) is not blocked by (9a). The reason is that Rule H compares only representations which are semantically equivalent. (9b) and (9a) are not semantically equivalent and are therefore not compared by Rule H.

- (9) a. John said that Bill liked his mother.  
           b. John said that Bill liked his mother.

Condition A requires that an anaphor be bound by the most local antecedent available for well-formed binding. Rule H requires that a pronoun be bound by the most local antecedent available for well-formed binding with a *designated interpretation*. The two conditions are stated in (10) and (11) respectively.

- (10) Condition A: An anaphor,  $\alpha$ , can be bound by an antecedent,  $\beta$ , only if there is no closer antecedent,  $\gamma$ , such that it is possible to bind  $\alpha$  by  $\gamma$ .
- (11) Rule-H: A pronoun,  $\alpha$ , can be bound by an antecedent,  $\beta$ , only if there is no closer antecedent,  $\gamma$ , such that it is possible to bind  $\alpha$  by  $\gamma$  and get the same semantic interpretation.

Rule H does not compare the representations in (9). However, it does compare the possible representations of (6) in (6'); these representations are semantically equivalent and (6'a) blocks (6'b)<sup>6</sup>

- (6') a. John said that he liked his mother.  
           b. John said that he liked his mother.

<sup>6</sup> The notation I use is borrowed from Higginbotham (1983). Of course, in order to give more substantial content to my claim that (6'a) and (6'b) yield the same interpretation, we have to go over the system of interpretation for Higginbotham's notation. Heim (1993) develops this system in detail, and I will touch upon it in section 2.

**4.1.2 Parallelism:** Parallelism is a general constraint which ensures that in ellipsis constructions the antecedent and elided VP receive parallel interpretations. (See chapter 3 and references therein.) Here I will only be concerned with the implications of Parallelism for the interpretation of pronouns. It is well known, since Ross (1967), that pronouns in an elided VP must receive parallel interpretations to the parallel pronouns in the antecedent VP. Thus, in a sentence such as (12), the pronoun in the antecedent and in the elided VP must receive the same referential value.

(12) Mary likes him and Sue does too <likes him>.

However, sameness in reference isn't always required. If the antecedent of the pronoun is within the first conjunct (and if certain other conditions hold, see Fiengo and May (1994) and Tomioka (1995)), Parallelism can be achieved in another way. In a sentence such as (13), for example, Parallelism can be achieved either by letting the pronouns have an identical referential value, as represented in (13'a) ("strict identity"), or by letting them be linked to their respective antecedents by identical dependencies, as represented in (8'b) ("sloppy identity"). Other options, such as (13'c), are ruled out by Parallelism.

(13) John likes his mother. Bill does too.

	Same refer. value	Identical depend.	No parallelism
(13')	a. John likes his mother. <sup>7</sup> <u>                    </u>	b. John likes his m. <u>                    </u>	c. * John likes his m. <u>                    </u>
	Bill likes John's mother.	Bill likes his m. <u>                    </u>	Bill likes Phil's m.

The descriptive generalization is stated in (14).

(14) NPs in the elided and antecedent VP must either (a) have the same referential value or (b) be linked by identical dependencies.

We would like this descriptive generalization to follow from a general theory of Parallelism, and indeed most theories are designed to capture the generalization. However,

---

<sup>7</sup> Here, and in other places in the chapter, I use an r-expression in order to represent identity in referential value. I use this merely as a notational device. Note, however, that Fiengo and May propose that a pronoun can actually be replaced by an r-expression in the elided VP (by vehicle change).

I believe that none of the theories capture the generalization in the appropriate manner.

Consider, for example, theories such as those of Sag (1976) and Williams (1977). These theories assume that the generalization follows from a requirement of semantic identity between the antecedent and the elided VPs. Under this assumption, strict identity, as in (8'a), is the result of accidental coreference. In other words, the claim is that the strict/sloppy ambiguity is the result of an ambiguity in the sentence that includes the antecedent VP. However, Dahl (1973) has demonstrated that the identity requirement is too strong. Specifically, it predicts that in a discourse which involves more than one elided VP, the interpretation will either be strict across-the-board or sloppy across-the-board. Mixed readings such as those in (15) are predicted to be unavailable.<sup>8,9</sup>

- (15) a. Smithers<sub>i</sub> thinks that his<sub>i</sub> job sucks. Homer does, too <think that Homer's job sucks>. However, Homer's wife doesn't <think that Homer's job sucks>.  
 b. Smithers<sub>i</sub> thinks that his<sub>i</sub> job sucks. Homer does, too <think that Homer's job sucks>. However, Marge doesn't <think that Homer's job sucks>.

I do not know how to derive the generalization in (14) from a principled theory. I will therefore elevate it to the status of a principle:

- (16) NP Parallelism: NPs in the antecedent in the elided VP must either:  
 (a) have the same referential value. (**Referential-Parallelism**) or  
 (b) be linked by identical dependencies. (**Structural-Parallelism**)

---

<sup>8</sup> The proposal about Parallelism that I made in chapter 3 suffers from the same problem since it appeals to structured propositions. Rooth's theory which appeals to unstructured propositions is too weak since it predicts that an antecedent sentence which is Scopally Uninformative (as well as one which is Binding Uninformative) will never disambiguate the ellipsis sentence (contrary to what we've seen in the previous chapter). See Fox (1995a: note 67). As far as I can see, the same problem holds for other proposals that don't make reference to the structure of the antecedent or the ellipsis sentence (e.g., Dalrymple et. al. (1991)).

<sup>9</sup> Fiengo and May (1994) assume that the mixed reading depends on the referential properties of the subject in the third clause. More specifically, they assume that when there is sloppy identity between the first pair of VPs, strict identity is possible in the second pair only if the subject of the third sentence contains an NP which is coreferential with the subject of the second sentence. Although this is generally correct, I don't think that it should be captured by a formal constraint. This can be seen in (15b) where the mixed reading is available for educated people (those who know that Marge is Homer's wife). My suspicion is that the general tendency is to have all three sentences parallel to each other. This would block a mixed reading given the descriptive generalization in (14). The only way to get a mixed reading is by dividing the sentences into two groups and checking Parallelism independently for each group. The referential properties of the subject help in the grouping.

**4.1.3 Deriving Dahl's observation:** Suppose that Parallelism applies to the output of Rule H. This means that (6) and (7) above will each be subject to Rule H, and that optimality considerations will be blind to Parallelism. This derives Dahl's observation. The basic idea is that Rule H doesn't allow the second pronoun in (6) to be bound by the matrix subject. However, the second pronoun in (7d) is bound by the matrix subject. Therefore (6) and (7d) cannot be parallel.

Let's see in more detail how, given Rule H, Parallelism distinguishes between (7a-c), on the one hand, and (7d) on the other. Recall that Rule H excludes (6'b) given the availability of (6'a) (and given that the two are semantically indistinguishable). Now, consider the representations that are given to (7a-d) by Rule H. These are presented in (7'a-d). What we have to show is that, given (6'a), (7'a-c) obey Parallelism, while (7'd) doesn't.

- (6') a. John said that he likes his mother.  
           └───┘ └───┘  
 b. \*John said that he likes his mother.  
           └───┘ └───┘
- (7') a. Bill said that John likes his mother.                   *(obeys Parallelism, with (6'a))*  
           └───┘ └───┘  
 b. Bill said that he likes his mother.                   *(obeys Parallelism, with (6'a))*  
           └───┘ └───┘  
 c. Bill said that he likes John's mother.               *(obeys Parallelism, with (6'a))*  
           └───┘  
 d. Bill said that John likes his mother.               *(doesn't obey Parallelism with (6'a);*  
           └───┘   *would obey Parallelism with (6'b))*

Each of the NPs in the VPs of the sentences in (7') must obey Parallelism with respect to the parallel NPs in (6'a). Consider first (7'a). Here the NP *John* obeys referential-parallelism with the first pronoun in (6'a). The pronoun in (7'a) is linked by an identical dependency to that which links the parallel pronoun in (6'a) to its antecedent, and thus obeys structural-parallelism.<sup>10</sup> Consider now (7'b). Here it is easy to see that both pronouns obey structural-parallelism with the parallel pronouns in (6'a). In (7'c) the first pronoun obeys structural-parallelism with the first pronoun in (6'a) and the NP *John* obeys

<sup>10</sup> This pronoun obeys referential-parallelism as well. For ease of presentation, I will ignore referential-parallelism in cases in which structural-parallelism holds.



referential-parallelism with the second pronoun in (6'a). Consider now (7'd). Here the NP *John* obeys referential-parallelism with the first pronoun in (6'a), but the pronoun in (7'd) doesn't obey Parallelism at all. It receives a different referential value from that of the second pronoun in (6'a), and is linked by a different dependency. The only way in which (7'd) could obey Parallelism would be if somehow non-local binding for (6), as in (6'b), were allowed. (7'd) would be parallel to (6'b) by virtue of referential-parallelism for the NP *John* and structural-parallelism for the pronoun *his*.

**4.1.4 An extension :** Before I go on to specify sophisticated predictions of the account presented here, let me present one prediction which might not be unique to this particular account, but is nevertheless interesting (and will be important later on). The prediction is that Dahl's puzzle can re-emerge with judgments of acceptability. The basic idea is that we can make (7), under its four possible readings, into the antecedent for ellipsis, and turn (6) into the sentence that includes the elided VP. The prediction is that when a sentence with a meaning like that of (7d) will be the antecedent for ellipsis, the result will be unacceptable. Of course, for the experiment to work, we have to find a way of keeping the interpretation of (6) constant. This is achieved by manipulating independent factors having to do with agreement and quantification.

Consider the paradigm in (17).<sup>11</sup>

- (17) a. Every boy said that Mary liked her dog.

Well, Mary did too

<said that she liked her dog>

- b. Every boy said that he liked his dog.

Well, Mary did too

<said that she liked her dog>

<sup>11</sup> Some speakers require parallelism in the morphological properties of the pronouns in the antecedent and the elided VPs (i.e., they do not accept sloppy identity in an ellipsis construction such as *John likes his mother and Mary does, too*. For such speakers, the judgments in (17), and in many other places, are not very clear. However, this problem is overcome if the elided VPs are replaced with phonologically reduced VPs.

c. Every boy said that he liked Mary's dog.

Well, Mary did too  
 <said that she liked her dog>

d. \* Every boy said that Mary liked his dog.

Well, Mary did too  
 <said that she liked her dog>

\* by Parallelism  
 \*by Rule H

Let's examine the pronouns in the elided VP beginning with (17a). The first sentence of (17a) is the equivalent of (7a) and the second is the equivalent of (6). There is no ambiguity here because of the gender switch. Parallelism between the sentences of (17a) is obeyed, just as it was in the pair <6, 7a>; the first pronoun in the elided VP obeys referential-parallelism, and the second pronoun obeys structural-parallelism. Consider now (17b). Here, just as in the pair <6,7b>, structural-parallelism is obeyed for both pronouns. Once again, there is no ambiguity; the "sloppy-identity" interpretation is forced, because quantifiers don't have a referential value, and only structural-parallelism can be obeyed. Consider now (17c). Here, as in the pair <6,7c>, structural-parallelism is obeyed for the first pronoun in the elided VP and referential-parallelism is obeyed for the second pronoun. Consider now (17d). As in the pair <6,7d>, there is no parallelism. The reason is that the pronoun in the first sentence of (17d) does not have a referential value, and therefore referential-parallelism can't hold. Further, structural-parallelism doesn't hold because the pronoun in the first sentence is linked to the matrix subject and the parallel pronoun in the second sentence is linked to the embedded subject. The only way in which Parallelism could hold would be if Rule H didn't apply and long distance binding were allowed in the second sentence of (17d).

**4.1.5 New predictions:** The explanation of Dahl's observation was crucially dependent on the idea that non-local binding was impossible (for (6) and for the second sentence of (17d)) due to Rule H. This generates a prediction; in environments in which Rule H allows for non-local binding, Dahl's puzzle would be obviated. There are two environments in which Rule H allows for non-local binding. One environment is where local binding and non-local binding yield different semantic interpretations. Another

environment is where local binding is impossible (for syntactic reasons). In both cases, local-binding and non-local-binding are not compared by Rule H. We therefore predict that in both cases sentences such as (17d) and (7d) will be licensed. Before we test for these predictions, we must understand in what range of cases local and non-local binding yield different semantic interpretations (besides trivial cases, such as (9) above). In other words, we have to understand how the lines we have drawn in the representations above match up to interpretive notions.

#### 4.2. More on Local and Non-local Binding

In the previous section we have assumed a system in which the bound variable interpretation of (18) could (in principle) be the result of two distinct representations, as in (18'). This system differs from the standard system of indexation in which there is nothing on the bound variable which designates who (out of a set of coindexed elements) is its binder. Therefore, the standard system of indexation provides only one representation for the intended interpretation of (18), given in (18'').

(18) Every boy knows that he likes his mother.

(18') a. Every boy knows that he likes his mother.

b. Every boy knows that he likes his mother.

(18'') Every boy<sub>i</sub> knows that he<sub>i</sub> likes his<sub>i</sub> mother.

The question is whether there are any reasons to assume representations such as those in (18'). Heim (1993) argues that there are.<sup>12</sup> Consider the sentence in (19).

(19) Every boy knows that only he likes his mother.

Focus on the interpretations in which the two pronouns are bound variables. There are two such interpretations. To see this, consider what property must hold of every boy for the sentence to be true. Suppose John is a boy. For (19) to be true, John, like any other boy,

---

<sup>12</sup> See Higginbotham (1983) and Partee and Bach (1984) for additional arguments in favor of representations such as those in (18').



(19') is the need to account for the ambiguity of sentences such as (19). Another, perhaps more important, motivation is that any account of the semantics of variable binding involves something like lambda abstraction. Once lambda abstraction is admitted, it turns out that representations such as (19'') for (19) and (18''') for (18) 'come for free'. An account of variable binding would need to justify ruling out any of these representations. The representations in (18') and (19') are just a shorthand for the lambda notations and as such are justified.

Assume, then, that representations such as those in (18') are, in principle, justified. Now we should ask, what is the motivation for Rule H? We have seen that when local and non-local binding yield different interpretations, as in (19), both are licensed. Rule H states that when they yield the same interpretation, as in (18), non-local binding, (18b), is blocked. What could be the motivation for this statement? Obviously, the motivation cannot come from interpretation alone. Rule H states that a certain representation is impossible when it is interpretively identical to another representation. In order to find evidence for Rule H, we need a way to distinguish between representations which are interpretively identical. Interpretation alone will be of no help. What we need is to utilize our knowledge of non-interpretive mechanisms that might be sensitive to the difference between the two representations.<sup>13</sup> Heim suggests Binding Theory as such a mechanism.<sup>14</sup>

Consider sentence (20). This sentence cannot have the interpretation of sentence (21). The question is why. Standard wisdom is that Binding Theory can provide the answer. Under a system that utilizes indexation, as in (18''), the only way for (20) to get the interpretation of (21) would be by the representation in (20'). In this representation, the pronoun *him* is bound in its Governing Category in violation of Principle B. However,

---

<sup>13</sup> I discuss a similar problem in Fox (1995a). There the issue was how to find evidence for a principle that allows QR only if QR affects interpretation. There the idea was to look for grammatical mechanisms, other than interpretation, that might be sensitive to the position of a quantifier. Among these were grammatical constraints such as Parallelism and the Coordinate Structure Constraint. Here the idea is to look for mechanisms that are sensitive to the distinction between local and non-local binding. Here, too, an obvious thing to look at is grammatical constraints. As we will see, Heim argues that Binding Theory is one such constraint. I will argue that Parallelism is another. In fact, Parallelism argues for Rule H in exactly the same way that it argued for the economy principle of Fox (1995a). See section 4.4.

<sup>14</sup> Heim doesn't describe the role that Binding Theory plays in her proposal in exactly these terms. The main goal of her chapter was not to present evidence for Rule H. Rather, her aim was to show that arguments in favor of an optimality consideration proposed in Reinhart (1983) (Rule I, cf. Grodzinsky and Reinhart (1993)) can be extended to arguments in favor of Rule H.

given the system we have argued for above, this simple answer cannot be maintained. Sentence (20) can get the interpretation of (21) with both representations in (20''). (20''a) violates Condition B, but (20''b) does not.

- (20) John said that he likes him.  
 (20') John<sub>i</sub> said that he<sub>i</sub> likes him<sub>i</sub>.  
 (20'') a. John said that he likes him.  
           b. John said that he likes him.  
 (21) John said that he likes himself.

This problem was noticed by Higginbotham (1983), who proposed to deal with it by complicating Binding Theory. Heim suggests keeping Binding Theory intact and utilizing Rule H instead. Specifically, she suggests that the ill-formedness of (20), under the intended interpretation, follows from an interaction of two principles. One principle is Rule H; the other is Principle B. Rule H is an optimality principle which blocks (20''b), given the availability of (20''a) (and given that the two are interpretively identical). Binding theory applies to the output of Rule H and rules out (20''a).

This account makes a new prediction. Namely, it predicts that in cases in which local and non-local binding yield a different interpretation, sentences similar to those in (20) will be good with a representation along the lines of (20''b). This representation will allow an obviation of Principle B. To see that this is the case, consider the sentences in (22).

- (22) a. Everybody hates Lucifer. In fact, Lucifer knows very well that only he (himself) pities him.  
       b. Everybody hates every devil. In fact, every devil knows very well that only he (himself) pities him.<sup>15</sup>

---

<sup>15</sup> Rule H is a modification of Reinhart's non-coreference rule. As such, it should also account for BT(C) obviations in constructions such as (i).

- (i) Every body hates John. Only he himself likes John.

Such cases are problematic if we assume that proper names denote individuals. Under this assumption, (i) contains no variables to which Rule H can apply. Irene Heim (p.c.) points out that this problem will not arise if we assume that proper names denote variables with the presupposition that the variable is identical to a given individual (*John* denotes a variable *x* and presupposes that *x* is identical to some individual JOHN). If an r-expression is c-commanded by a potential antecedent, Rule-H will force variable binding (unless non-variable binding is semantically distinct). In turn, variable binding from an A-position is blocked by BT(C). In (i), variable binding is semantically distinct from

In (22) Principle B is obviated. The explanation is straightforward. The existence of the focus particle *only* brings about an interpretive difference between the cases of local and non-local binding, along the lines of (19). Therefore, Rule H licenses non-local binding. Non-local binding doesn't violate Principle B and the sentences are okay. In fact, it is possible to tell from the interpretation that only non-local binding is possible in (22). The sentences in (22) are well-formed only under the interpretation of non-local binding as in (23). The interpretation of local binding, as in (24), is ruled out by Principle B. This could be seen by comparing (22) with (25). The sentences in (25), in contrast with those in (22), can have the interpretations in (24).

- (23) a. Lucifer knows that only he (himself) pities him.  
           Lucifer  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } x))) \Rightarrow$   
           Lucifer knows that I don't pity him.
- b. Every devil knows that only he (himself) pities him.  
           Every devil  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } x))) \Rightarrow$   
           Every devil knows that I don't pity him.
- (24) a. Lucifer knows that only he (himself) pities him.  
           Lucifer  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } y))) \Rightarrow$   
           Lucifer knows that I don't pity myself.
- b. Every devil knows that only he (himself) pities him.  
           Every devil  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } y))) \Rightarrow$   
           Every devil knows that I don't pity myself.
- (25) a. Lucifer knows that only he pities himself.  
           Lucifer  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } y))) \Rightarrow$   
           Lucifer knows that I don't pity myself.
- b. Every devil knows that only he pities himself.  
           Every devil  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } y))) \Rightarrow$   
           Every devil knows that I don't pity myself.

Notice that Heim's explanation for the ill-formedness of (20) (under the intended interpretation) is almost identical to the explanation that I proposed in the previous section

---

coreference and the latter is licensed.

for Dahl's puzzle.

(20) John said that he likes him.

Heim suggests that the ill-formedness of (20) follows from a certain interaction of Rule H and Binding Theory. I suggest that an identical interaction holds between Rule H and Parallelism and that this interaction accounts for Dahl's puzzle. Parallelism, just like Binding Theory, applies to the output of Rule H. The ordering is crucial for both accounts. Rule H must be oblivious to the requirements of Binding Theory and Parallelism, otherwise the fact that local binding is violated by these constraints could serve as justification for non-local binding.

Binding Theory and Parallelism provide the kind of evidence that one can expect to find for Rule H. Both are constraints that are sensitive to the difference between local and non-local binding, and both show that when the two representations are interpretively identical non-local binding is not licensed.

Of course, in order to be convinced that the accounts are correct, we have to show that in cases in which local and non-local binding yield different semantic effects, Binding Theory and Parallelism show us that non-local binding is possible. Heim showed that this is the case for Binding Theory. In the next section we will show that this is also the case for Parallelism.

#### **4. 3. When Local and Non-Local Binding Yield Different Interpretations**

In the previous section we have seen a case in which local and non-local binding yield different semantic interpretations. We have seen that in this case Principle B was obviated, thus supporting Heim's general approach. The aim of this section is to show that whenever a difference in interpretation brings about an obviation of Principle B, it also brings about an obviation of Dahl's puzzle. This is one of the two predictions which the account of Dahl's puzzle makes and which were spelled out at the end of section 1.

Consider the contrast between (26) and (27). In (26), just as in (20), local and non-local binding yield the same semantic effect. Non-local binding is ruled out by Rule H, and subsequently local binding is ruled out by Principle B. In (27), by contrast, Principle B is obviated. The reason is, presumably, that local binding, (27'a), and non-local binding, (27'b), yield different semantic interpretations.



(26) John said that he likes him.

(27) Every boy said that John likes him. Even John, himself, said that he likes him.

(27') a. ... Even John, himself, said that he likes him.

Even (John,  $\lambda x (x \text{ said that } x \lambda y (y \text{ likes } y))$ )

b. ... Even John, himself, said that he likes him.<sup>16</sup> (he = John)

Even (John,  $\lambda x (x \text{ said that he } \lambda y (y \text{ likes } x))$ ) (he = John)

The difference in interpretation between the representations in (27') is similar, though not identical, to the difference we have seen in (22). The similarity stems from the existence of a focus particle. Focus particles, such as *only* (in (22)) and *even* (in (27)), affect interpretation via reference to the focus value of the constituent with which they compose (see Rooth (1985)). When there are no focus particles, the difference in the focus value of certain constituents doesn't affect the interpretation of the sentence. However, when focus particles are around, focus values become relevant for interpretation. Focus values, in turn, can be affected by the kind of dependencies that exist between pronouns and their antecedents. In particular, they sometimes care about the difference between local and non-local binding. For this reason, the presence of a focus particle has the potential of making the difference between local and non-local binding relevant for interpretation. We have seen this happen in the case of *only* in (22). Let us now see how it happens in the case of *even* in (27).

The difference between (22) and (27) is directly linked to the different contributions that *only* and *even* make to the interpretations of sentences in which they are contained. *Only* contributes to interpretation by affecting truth conditions whereas *even* contributes to interpretation by affecting presuppositions (See Rooth (1985) and Wilkinson (1996), among others). A sentence with the structure of (28a) is true just in case there is no person,  $x$ , other than John, such that  $\phi(x)$ . A sentence with the structure of (28b) is true just in case  $\phi(\text{John})$  is true. However, it makes an additional requirement on presuppositions. For a

<sup>16</sup> The representation in (20'b) differs from the representation in (13''b) in that the intermediate pronoun is not bound by the matrix subject, but rather enters a relationship of coreference with it. We see that both representations obviate BT(B). The difference between them is not relevant for Dahl's puzzle, under the account presented here, since both involve non-local binding by the matrix subject and, therefore, both should allow obviation of Dahl's puzzle.

sentence with the structure of (28b) to be uttered, it must be agreed that John is one of the least likely people to have the property  $\lambda x\phi(x)$ .

- (28) a. Only (John,  $\lambda x\phi(x)$ )  
b. Even (John,  $\lambda x\phi(x)$ )

We can now be precise about the difference in the interpretations of (27'a) and (27'b). The difference has to do with the contribution that *even* makes to interpretation. In other words, it has to do with presuppositions. In (27'a) the presupposition is that John is one of the least likely people to have the property  $\lambda x (x \text{ said that } x \lambda y (y \text{ likes } y))$ . This entails, for instance, that I, for one, am probably more likely than John to say that I like myself. In (27'b) the presupposition is that John is one of the least likely people to have the property  $\lambda x (x \text{ said that he } \lambda y (y \text{ likes } x))$ , where  $\text{he}=\text{John}$ . This entails, for instance, that I, for one, am probably more likely to say that he, that is John, likes me. There is obviously a formal difference in the presuppositions of the two sentences. This difference licenses non-local binding. The availability of non-local binding, (27'b), allows (27) to obviate Principle B.<sup>17</sup>

Given this analysis, we predict that Dahl's puzzle will be obviated in an environment such as that of (27). That this is the case is demonstrated by the contrast in (29). (29a) is bad because non-local binding is ruled out in the second sentence and, as a consequence, Parallelism is not maintained. (29b) is okay because in the second sentence local and non-local binding yield different semantic interpretation. Non-local binding is allowed, and Parallelism is maintained.

- (29) a. \*Every adult thought that the little kid liked him. The little kid did too. <thought that the little kid liked the little kid>  
b. Every kid thought that the little kid liked him. Even the little kid, himself, did.

Notice that the second sentence in (29a) might also be ruled out by Principle B when applied to the elided VP (see, Kitagawa (1992), Ristad (1992), Fiengo and May

<sup>17</sup> Kai von Stechow and Irene Heim (p.c.) pointed out a problem that arises at this point. If the meaning of *even* licenses non-local variable binding why is the same not true for the meaning of the sentential particle *too*.

It might be possible to address this problem by assuming a local application of Rule-H. The basic idea would be that in the case of *too* local variable binding is determined before the semantics of the focus particle are taken into account. However, it remains to be seen whether the details of such a suggestion can be worked-out in a natural way.

(1994), among others). Therefore, (30) might serve as a better minimal pair.

- (30) a. \*Every boy<sub>i</sub> said that Mary liked his<sub>i</sub> dog. Well, Mary did too  
           <said that she liked her dog>                    (= (10d))  
       b. Every person<sub>i</sub> said that Mary liked his<sub>i</sub> dog. Even Mary, herself, did  
           <said she liked her dog>

Consider now the contrast in (31). Once again, the explanation of the contrast is that in (31b) non-local binding yields a different semantic effect than that of local binding (the difference has to do with the semantic contribution of the focus particle *only*. By now, this difference should be clear to the reader, and I will not go over it.). Non-local binding is, thus, licensed by Rule H, and Principle B is obviated.

- (31) a. John believes that he likes him.  
       b. Almost no one believes that John likes him. Probably John is the only one who believes that he likes him.  
       c. Almost no one believes that John likes him. Probably John is the only one who does <believes that he likes him>.

Once again we predict that Dahl's puzzle would be obviated as well. This is demonstrated in (31c), and by the contrast in (32)

- (32) a. \*No boy<sub>i</sub> believes that Mary likes his<sub>i</sub> mother. However, she does.  
       b. No one<sub>i</sub> believes that Mary likes his<sub>i</sub> mother. Probably, Mary is the only one who does.

Consider now the sentences in (22) and their two conceivable representations in (23-24), repeated below. This case is different from (31) with respect to the position of the focus particle.<sup>18</sup> However, as we have seen in the previous section, this is also an environment in which local and non-local binding yield a different semantic effect and in which Principle B is obviated.

- (22) a. Everybody hates Lucifer. In fact, Lucifer knows very well that only he (himself) pities him.  
       b. Everybody hates every devil. In fact, every devil knows very well that only he (himself) pities him.

<sup>18</sup> It is also different in that the first pronoun is linked to the matrix subject via direct binding and not via coreference (see note 16).

- (23) a. Lucifer knows that only he (himself) pities him.  
 Lucifer  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } x))) \Rightarrow$   
 Lucifer knows that I don't pity him.
- b. Every devil knows that only he (himself) pities him.  
 Every devil  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } x))) \Rightarrow$   
 Every devil knows that I don't pity him.
- (24) a. Lucifer knows that only he (himself) pities him.  
 Lucifer  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } y))) \Rightarrow$   
 Lucifer knows that I don't pity myself.
- b. Every devil knows that only he (himself) pities him.  
 Every devil  $\lambda x (x \text{ knows that only } (x, \lambda y (y \text{ pities } y))) \Rightarrow$   
 Every devil knows that I don't pity myself.

As we see by the contrasts in (33) and in (34), this is an environment in which Dahl's puzzle is obviated as well.

- (33) a. \*Every boy believes that Mary likes his mother. Mary does too <believe that she likes her mother>.
- b. Every boy believes that only Mary likes his mother. Mary does too <believe that only she likes her mother>.<sup>19</sup>
- (34) a. \*John believes that he likes his mother. Bill does too <believe that John likes Bill's mother>.
- b. John believes that only he likes his mother. Bill does too <believe that only John likes Bill's mother>.

As in the case of (22b), we can tell by the interpretation that only non-local binding is possible in (33b) and in (34b). Thus, in (33b) the belief that is attributed to Mary is the belief that Mary is the only person who likes Mary's mother, not the belief that she is the only person who likes her own mother. The same is true of the belief that is attributed to Bill in (34b). Once again this follows from the proposal; local binding would violate Parallelism.

<sup>19</sup> As in the case of (22), we can tell by the interpretation that only non-local binding is possible. This is explained by the fact that local binding would violate Parallelism.

#### 4.4. When Local Binding Is Impossible

In the previous section, I have argued that when local and non-local binding yield different semantic interpretations, Dahl's puzzle is obviated. This was predicted by the explanation I offered for Dahl's puzzle in section 1. The basic idea was that whenever there is a choice, Rule H prefers local binding to non-local binding. One place where there isn't a choice (given the way the comparison set is constructed) is where local and non-local binding yield different interpretations. As mentioned in section 1, there is also another place where there isn't a choice: if local binding is impossible for grammatical reasons, non-local binding would be the only candidate in the reference set, and would thus be licensed. In this section, I will demonstrate that also in this place Dahl's puzzle is obviated.

I will discuss two principles that under certain circumstances might force non-local binding. One principle is that which requires c-command for binding; the other is the theta criterion. I will assume that these principles differ from Parallelism and from Binding Theory in that they exclude elements from the comparison set.<sup>20</sup>

**4.4.1 C-command:** C-command is a prerequisite for Binding. Consider a sentence such as (35). In this sentence, the matrix subject must bind both of the pronouns, as in (35'a), since the local binding, in (35'b), is impossible (for lack of c-command).<sup>21</sup> Rule H licenses (35'a) since (35'b) is an illegitimate competitor. We therefore predict that when we embed a sentence such as (35) in an ellipsis construction, we will not see the effects of Dahl's puzzle. That this is the case is demonstrated in Fiengo and May (1994:156) (see also

---

<sup>20</sup> This distinction is similar in nature to Chomsky's distinction between principles that cause a derivation to crash and principles that can rule out a converging derivation (such as the principles of Binding theory). It is possible that the difference between the theta criterion and the c-command requirement on the one hand and Binding Theory and Parallelism on the other is that the former are prerequisites for interpretation. C-command might be necessary for the interpretation of a bound variable relationship (in fact, it is, under many semantic theories). Satisfying the argument structure of a predicate (the theta criterion) is also necessary for achieving an interpretation. If the comparison set is composed of representations with a designated interpretation, it follows, trivially, that all representations in the reference set must be interpretable.

<sup>21</sup> Notice that even if c-command were not a prerequisite for binding, it is not clear that in (35) the linearly closer antecedent is more local under the appropriate measure of locality. If, as some assume, locality is defined by c-command ( $\alpha$  is more local to  $x$  than  $\beta$ , just in case  $\alpha$  c-commands  $x$  and is c-commanded by  $\beta$ ), then, even if c-command were not a prerequisite for binding, Dahl's puzzle would be obviated (since both the representations in (35') would be licensed, neither involving more local binding than the other).

Kehler (1993)). It is further demonstrated by the minimal pairs in (36) and in (37).

(35) Bill said that all of his friends like his mother.

(35') a. Bill said that all of his friends like his mother.

b. \*Bill said that all of his friends like his mother.

(36) a. Bill said that he likes his mother.

John did too <\*say that Bill likes John's mother>

b. Bill said that all of his friends like his mother.

John did too <said that all of Bill's friends like John's mother>

(37) a. No boy said that Mary likes his mother.

\*Mary, however, did <said she likes her mother>

b. No boy said that all of Mary's friends like his mother.

Mary, however, did <said that all of her friends like her mother>

**4.4.2 The theta criterion:** Some component of grammar must require that a binding relationship will hold between a moved constituent and its trace—call it the theta criterion.<sup>22</sup> Consider the implications for a sentence such as (38). In particular, consider the two representations in (38'). (38'a) involves non-local binding of the trace, and does not violate the theta criterion. (38'b) involves local binding of the trace but violates the theta criterion (c.f. Rizzi (1986)).<sup>23</sup> Given the assumption that theta-criterion violators are not legitimate competitors for Rule H, (38'a) is licensed by Rule H.

(38) John seems to himself [t to be a genius]

<sup>22</sup> Noam Chomsky (p.c.) points out that under natural assumptions regarding the formation of A-chains, a representation such as the one in (38'b) is meaningless. Therefore, nothing special needs to be stated in order to account for the fact that (39'b) is the only representation in the reference set.

<sup>23</sup> Rizzi suggested his chain-formation algorithm in order to account for the ill-formedness of strong-crossover in many environments. In particular, he suggested that the algorithm requires that (38'b) will be the representation of (38) and it predicts that the sentence will be ruled out by the theta criterion. (38) poses a problem for Rizzi's suggestion, as noted in Chomsky (1986b). The support for the algorithm comes from the ungrammaticality of the Italian parallels of (38).

Rizzi's algorithm is very similar to Rule H. In particular, if we would assume that the theta criterion, like Binding Theory and Parallelism, applies to the output of Rule H, we would get Rizzi's results. In the spirit of Reinhart (1998), this could be the proposal for SCO in the case of A-bar movement.

- (38') a. John seems to himself [t to be a genius]  
           b. John seems to himself [t to be a genius]

Now consider (39). Given the fact that (39'a) is a violation of the theta criterion, only (39'b) and (39'c) compete with respect to Rule H. (39'b) is preferred. The prediction is that if (39) is embedded in an ellipsis construction, Dahl's puzzle will be obviated.

- (39) John seems to himself [t to be smarter than his classmates]
- (38') a. John seems to himself [t to be a genius]  
           b. John seems to himself [t to be a genius]
- (39') a. John seems to himself [t to be smarter than his classmates]  
           b. John seems to himself [t to be smarter than his classmates]  
           c. John seems to himself [t to be smarter than his classmates]

To see that this is the prediction, consider an ellipsis construction such as that in (40a). In the first sentence of this construction, the pronoun *his* is bound by the chain <John, t>. Therefore, the parallel pronoun of the second sentence can be bound by the chain <Bill, t> without violating Parallelism. To see that the prediction is borne out, consider the contrasts in (40) and in (41).<sup>24</sup>

- (40) a. John seems to himself [t to be smarter than everyone in his class].  
           Bill does too <seems to John [t to be smarter than everyone in Bill's class]>.  
       b. John believes [himself to be smarter than everyone in his class].  
           \*Bill does too <believes John to be smarter than everyone in Bill's class>.<sup>25</sup>
- (41) a. [No girl in the committee]<sub>i</sub> seems to John to be smarter than her<sub>i</sub> classmates.  
           John, however, does.  
       b. \*[No girl in the committee]<sub>i</sub> believes John to be smarter than her<sub>i</sub> classmates.  
           John, however, does.

<sup>24</sup> To assist you with the judgments in the (a) cases, consider a context in which John is searching in each class for the smartest person; say, to form an elite club.

<sup>25</sup> Note that there is no independent problem with strict identity for reflexives in ECM constructions. (See Fiengo and May (1994) and Fox (1993) for some discussion.)

## 4.5. Rule H and Scope-Economy

In this section, I would like to show that my argument in favor of Rule H is identical to one of the arguments that I presented in favor of Scope-Economy (in Fox (1995a) and in another form in chapter 2). Further, I would like to show that the generalization that I present in this chapter is very similar to the generalization presented in Fox (1995a), and that, in fact, there is a simple statement that encompasses both of these generalizations. Finally, I would like to suggest an outline for a potential unification of Rule H and Scope-Economy. However, as mention in the outset, this unification will be inconsistent with the local definition of Scope-Economy for which I argued in chapter 2.

**4.5.1 The Ellipsis Binding Generalization:** In the previous sections, I have argued for Rule H on the basis of the fact that it provides a solution for Dahl's puzzle. Stated somewhat differently, the argument was based on the following generalization:

- (42) **The Ellipsis Binding Generalization (EBG):** In an ellipsis construction, a pronoun can have a non-local antecedent only if binding by a local antecedent will yield a different interpretation, both in the sentence that includes the elided VP and in the sentence that includes the antecedent VP.

The generalization is partially exemplified by the paradigm in (43). In (43a), local and non-local binding yield a different interpretation in the sentence that includes the *elided* VP, but not in the sentence that includes the *antecedent* VP; therefore, according to the EBG, non-local binding is impossible (this is Dahl's puzzle). In (43b, c) non-local binding yields a different interpretation in both sentences, and it is allowed according to the EBG.

- (43) a. John said that he liked his mother. Bill did too \*<said that John liked Bill's mother>.  
 b. John said that Mary liked his mother. Bill did too <said that M. liked B's mother>.  
 c. John said that only he liked his mother. Bill did too <said that only J. liked B's mother>.

The explanation of the EBG was based on an interaction between Rule H and Parallelism. Rule H applies before Parallelism. If in both sentences non-local binding is semantically distinct from local binding, the two options are not compared by Rule H. Non-local binding is allowed in both sentences, and Parallelism can be maintained. If,



however, non-local binding is semantically indistinguishable from local binding in one of the sentences, then Rule H doesn't allow non-local binding in that sentence, and consequently Parallelism doesn't allow non-local binding in the other sentence.

**4.5.2 The Ellipsis Scope Generalization:** In Fox (1995a), I have presented an identical argument for a semantically based Economy principle. There, the argument was based on the Ellipsis Scope Generalization:

- (44) **The Ellipsis Scope Generalization (ESG):** In an ellipsis construction, a quantifier can have non-local scope only if local scope will yield a different interpretation, both in the sentence that includes the elided VP and in the sentence that includes the antecedent VP.

The generalization is partially exemplified by the paradigm in (45). In (45a, b), local and non-local scope yield a different interpretation in the sentence that includes the antecedent VP, but not in the sentence that includes the elided VP; therefore, by the ESG, non-local scope is impossible. In (45c,d) non-local scope yields a different interpretation in both sentences, and it is allowed according to the ESG.

- |      |    |  |                       |                        |
|------|----|--|-----------------------|------------------------|
| (45) | a. | Some boy admires every teacher. Mary does too.       | $(\exists > \forall)$ | $*(\forall > \exists)$ |
|      | b. | Some boy admires every teacher. Every girl does too. | $(\exists > \forall)$ | $*(\forall > \exists)$ |
|      | c. | Some boy admires every teacher. Some girl does too.  | $(\exists > \forall)$ | $(\forall > \exists)$  |
|      | d. | Some boy admires every teacher. Many girls do too.   | $(\exists > \forall)$ | $(\forall > \exists)$  |

The explanation of the ESG was based on an interaction between Parallelism and a semantically based economy principle. Economy applies before Parallelism. If in both sentences, non-local scope is semantically distinct from local scope, the two options are not compared by Economy. Non-local scope is allowed in both sentences, and Parallelism is maintained. If, however, non-local scope is semantically indistinguishable from local scope in one of the sentences, then Economy doesn't allow non-local scope in that sentence, and consequently Parallelism doesn't allow non-local scope in the other sentence.

**4.5.3 An outline of a unification:** The ESG and the EBG are strikingly similar. In fact, the only difference between them is that whenever the EBG talks about local or non-local binding, the ESG talks about local or non-local scope. Also the account of the two generalizations seems almost identical. Here, the only difference is that whenever Rule H

prefers local binding to non-local binding, Economy prefers local scope to non-local scope. The question is, can we unify the two accounts? Can we state a natural generalization that will encompass both the ESG and the EBG, and can we state a single optimality condition which will prefer both local to non-local binding and local to non-local scope?

In order for a unification to succeed, we need to find a feature that is shared by local binding and local scope and that distinguishes them from their non-local counterparts. I would like to suggest, tentatively, that the feature is related to length of dependencies. In a structure such as (46a), which involves local scope, there is a shorter dependency, than that in (46b), which is the result of a longer instance of QR. The structures are identical in all other respects. In a structure such as (47a), which involves local binding, there is a shorter dependency than that in (47b), which involves non-local binding. Once again, the structures are identical in all other respects.

- (46) a.  $[_{IP} XP... [_{vP} \text{Everyone} [_{vP} \text{likes } t]]]$   
           b.  $[_{IP} \text{Everyone} [_{IP} XP... [_{vP} \text{likes } t]]]$
- (47) a. John said he likes his mother  
           b. John said he likes his mother

Suppose that we had the optimality condition in (48). This condition prefers (46a) to (46b) and (47a) to (47b). If we assume that the comparison set for this condition includes only linguistic objects with the same semantic interpretation and if we assume that Parallelism applies to the output of this condition, we account for both the ESG and the EBG, which can now be stated together as in (49).

- (48) Prefer (representations with) shorter dependencies.<sup>26</sup>

<sup>26</sup> This unification raises a problem for Quantifier Lowering. In Fox (1995a) as well as chapter 2, I argue that economy prefers to reduce instances of Quantifier Lowering. This preference makes perfect sense under a derivational approach in which Optimality considerations attempt to reduce derivational effort. Quantifier Lowering, just like Quantifier Raising, involves derivational effort and is expected to be avoided if possible. Under the representational approach suggested in (48), the attempt to reduce instances of Quantifier Lowering makes little sense. Quantifier Lowering has the effect of reducing the length of dependencies and thus should be preferred under (48). See Fox (1995c) where I

- (49) **The Ellipsis Dependency Generalization (EDG):** In an ellipsis construction, a non-local dependency can exist only if more local dependencies yield different interpretations, both in the sentence that includes the elided VP and in the sentence that includes the antecedent VP.

#### 4.6. Conclusion

This chapter provides new arguments for an economy condition which has been postulated in Heim (1993), Rule H. Rule H is sensitive to certain aspects of interpretation and is oblivious to Parallelism. As such, Rule H is very similar to Scope-Economy. I provided the guidelines of a potential unification for the two economy conditions. However, this unification is stated as a global economy condition and is therefore at the moment inconsistent with the local algorithm that I have advocated for Scope Economy in chapter 2. This inconsistency is left as an open problem.

## **Part II: Binding Theory and the Representation of Scope**

In chapter 2-3, various mechanisms were utilized to test the predictions of Scope-Economy. Special attention was paid to the idea that Parallelism can help us detect the structure of Scopally Uninformative sentences. The same logic was applied in chapter 4 to the investigation of Binding Uninformative sentences. The result of this investigation revealed an economy condition on variable binding (VB-Economy) which is very similar in its nature to Scope-Economy.

As mentioned, VB-Economy was suggested on independent grounds in Heim (1993). Modifying work by Reinhart (1983), Heim demonstrates that certain obviation of Binding Theory (BT) can be accounted for by the postulation of VB-Economy. The argument for VB-Economy from BT is identical in its logic to the argument from Parallelism. BT, just like Parallelism, enables us to distinguish between Binding Uninformative representations, and in doing so it vindicates the predictions of Economy.

Stated somewhat differently, what we've learned up to this point is that Parallelism can be used to study scope, and that Parallelism and BT can be used to study variable binding. The obvious question to ask is whether BT can be used to study scope as well.

A necessary condition for a positive answer is that BT applies at LF. In chapter 5-6, I will argue that at least condition C of BT (BT(C)) applies at LF and only at LF. As such, this condition provides a very powerful tool for studying the covert structure of quantification. Based on this tool, I will argue (in chapter 5) that Scope Reconstruction is the result of "literal" reconstruction in the syntax and (in chapter 6) that A-bar chains are converted to operator variable constructions under an economy condition (suggested in Chomsky (1993) in a somewhat different form) which I will call OV-Economy. This economy condition allows deletion of material from the copy at the tail of an A-bar chain, only if such deletion is needed for interpretability.

Assuming that Binding Theory applies at LF, we would like to see whether it can be used to further test the predictions of Scope-Economy. As we will see, OV-Economy makes it impossible to test the predictions of Scope-Economy based on BT(C). The case of BT(A), which I address in the end of chapter 6, is different. It seems that although normal cases of A-bar movement don't affect BT(C), they can affect BT(A). Although this fact is ill-understood, it can be utilized to test the predictions of Scope-Economy and once again the predictions are borne out.

## Chapter 5: Condition C and Scope Reconstruction<sup>1</sup>

In the early days of generative grammar, overt movement was characterized as involving a disparity between the position at which an item is interpreted and the position at which it is overtly realized. However, since the 60s, it has been known that this is not strictly true. Although certain interpretive properties of a moved constituent are determined in the base position, other properties can be determined at the landing site.<sup>2</sup> Specifically, although aspects of interpretation having to do with predicate argument relations -- with theta assignment -- are determined at the base position, aspects of interpretation having to do with scope and variable binding are determined at the landing site (see Chomsky (1977) and references therein.)<sup>3</sup> This has led to a different view of the grammar's architecture. Under this view, overt displacement affects meaning as well as sound.<sup>4</sup> Structures involving movement serve as the input for both the conceptual and the articulatory systems. In the articulatory system, the base position is ignored. However, in the conceptual system, both positions are taken into account, and the interpretive properties are distributed among them according to the division outlined above.

This new picture itself had to be revised given the observation that the effects of movement on scope and variable binding are not obligatorily present; there are cases where the semantic effects of movement (predicted by the new picture) are "undone" (henceforth, cases of "Scope Reconstruction"). The attempts to deal with Scope Reconstruction, and the resulting accommodation to the picture, can be divided into two types. The first type of accommodation assumes that Scope Reconstruction is the outcome of semantic procedures. What is claimed is that the interpretive principles can deal with movement chains in at least two ways. One way results in an interpretation in which scope is determined at the base position; the other, in which scope is interpreted at the landing site. The second type of accommodation assumes that Scope Reconstruction is already determined in the syntax. In

---

<sup>1</sup>Many of the arguments in section 5.2.-5.3. were made independently in Romero (1997).

<sup>2</sup>Some suggestions that this is the case were already present in Chomsky (1957). For a collection of many of the original arguments, see Jackendoff (1972).

<sup>3</sup>Other aspects of interpretation which are affected by movement (e.g. topic/focus) are outside the scope of this work.

<sup>4</sup>This new view (together with the postulation of traces) paved the road for an account of scopal ambiguity in terms of movement operations that are invisible to phonology (such as Quantifier Raising).

other words, it is claimed that the structures that serve as the input to semantic interpretation (the structures of LF) determine whether or not there is Scope Reconstruction.

This chapter argues for various interactions between Condition C of the Binding Theory (henceforth, BT(C)) and Scope Reconstruction. More specifically, it is argued that BT(C) indicates the position in a chain at which a scope bearing element is interpreted. This correlation is mysterious under the semantic approach to Scope Reconstruction. However, it follows straight-forwardly from the syntactic approach under the assumption that Binding Theory is sensitive to LF positions. As such, it argues that Scope Reconstruction is syntactic and that Binding Theory applies at LF.

The structure of the chapter is as follows. In section 5.1, I will elaborate on the two approaches to the phenomenon of Scope Reconstruction and establish that a specific correlation between this phenomenon and BT(C) effects is predicted by the syntactic approach. In sections 5.2 and 5.3, I will demonstrate that the prediction is born out, for A-bar and A movement respectively. Finally, in section 5.4, I will discuss the ramifications of the arguments in favor of syntactic reconstruction for the status of the semantic mechanism of type shifting.

### 5.1. Semantic vs. Syntactic Accounts of Scope Reconstruction

It is well known that overt movement can affect scope. This has been established in many ways. The simplest examples are probably from scrambling languages (see Hoji, 1985).<sup>5</sup> However, we can also see this by English examples such as (1).<sup>6</sup> In (1b), the embedded

---

<sup>5</sup>For some reason, which is not completely clear to me, the ability to affect scope is restricted to short-distance scrambling. See Tada (1993).

<sup>6</sup>The English examples that are standardly invoked to make this point are simpler. For example, Jackendoff (1972), argues that movement affects scope based on contrasts such as (i): In (ia) *many* prefers to have wide scope relative to negation; in (ib) the preference goes in the other direction.

(i) a. Many arrows didn't hit the target.  
b. The target wasn't hit by many arrows.

The examples demonstrate that overt movement affects interpretive preferences. However, they do not demonstrate that overt movement yields interpretations that would be unavailable otherwise. Given the availability of covert scope shifting operations, both sentences in (i) are ambiguous. In order to demonstrate that scope is affected by movement, overt movement must do something that QR cannot do. In other words, the demonstration is dependent on the existence of constraints which apply to QR and not to overt movement. In (1) the constraint might follow from an account of the "clause-boundedness" of QR. There might be additional constraints on specific types of quantifiers

object is overtly displaced and the result of this displacement allows it to receive wide scope relative to another scope-bearing element (the existential quantifier). This scope relation would have been impossible without overt movement, as demonstrated in (1a).<sup>7</sup>

- (1) a. John seems to a (#different) teacher [ t to be likely to solve every one of these problems].  $(\exists > \forall) *(\forall > \exists)$ <sup>8</sup>  
 b. [Every one of these problems] seems to a (different) teacher [ t to be likely t to be solved t by John.  $(\exists > \forall) (\forall > \exists)$

However, overt movement does not obligatorily affect scope. This has been known at least since May (1977). Consider constructions that involve successive cyclic raising such as those in (2). In these constructions the scope of the moved quantifier *can* be affected by movement. However it need not; the scope may be construed in the base position or in any of the intermediate landing sites.

- (2) a. Someone from New York is very likely t to win the lottery.  
 b. Someone from New York seems t to be very likely t' to win the lottery.  
 c. Many soldiers seem t to be very likely t' to die in the battle.

Take (2a), which is two-ways ambiguous. One interpretation results from the quantifier taking scope in the final landing site. For the sentence to be true under this interpretation, there must be a person from NY who is very likely to win the lottery (e.g., a person who bought a sufficient number of tickets to make winning a likely outcome). Under the second interpretation, in which the quantifier has scope in the position of *t*, the truth conditions are

---

(e.g. monotone decreasing, see McCawley (1998: 618-628)) which will yield a similar argument. (See Liu (1990), Beghelli (1993) and Beghelli and Stowell (1995) for a detailed investigation of the properties of different quantificational elements.)

<sup>7</sup>A similar point can be made when we consider variable binding. In (ia), the universal quantifier cannot bind a variable which is outside of its scope. In (ib), overt movement gives the quantifier wider scope and allows it to bind the variable.

- (i) a. \*The teacher is expected by his<sub>1</sub> mother [ t to encourage every boy<sub>1</sub> ].  
 b. Every boy<sub>1</sub> is expected by his<sub>1</sub> mother [ t to be encouraged t by the teacher ].

<sup>8</sup>The impossibility of wide scope for the universal quantifier can be seen by the ungrammaticality of (1a) when *different* receives a bound interpretation as in *a different guard is standing on top of every building*. We can further demonstrate the impossibility of the  $(\forall > \exists)$  scope relation by considering cases in which the alternative scope relation results in an interpretation which is cognitively anomalous, e.g.: # *This soldier seems to someone to be likely to die in every battle*. or # *The ball seems to a boy to be under every shell*. (c.f. *Every shell seems to a (different) boy to be over the ball*.)

much less demanding; they merely require that there be enough ticket buyers from NY to make it likely that the city would yield a winner.

The examples in (2) demonstrate the availability of Scope Reconstruction. This availability can be further demonstrated in ways that are not based on any assumptions regarding the interpretation of modality. The demonstration is based on the constructions in (3, 4), partially due to Lebeaux (1994).<sup>9</sup>

- (3) a. [At least one soldier]<sub>i</sub> seems (to Napoleon) [t<sub>i</sub> to be likely to die in every battle].
- b. [At least one soldier]<sub>i</sub> seems to himself<sub>i</sub> [t<sub>i</sub> to be likely to die in every battle].
- c. [At least one soldier]<sub>i</sub> seems to his<sub>i</sub> commanders [t<sub>i</sub> to be likely to die in every battle].
- (4) a. One soldier is expected (by Napoleon) [t to die in every battle].
- b. One soldier<sub>i</sub> is expected by his<sub>i</sub> commander [t<sub>i</sub> to die in every battle].

In the (a) sentences in (3,4) the universal quantifier in the embedded clause can take scope over the matrix subject ( $\forall > \exists$ ). In other words, the sentences can be understood with the soldiers varying with the battles. The sentences can also receive an interpretation in which the matrix subject takes wide scope ( $\exists > \forall$ ). This is the implausible interpretation which asserts the existence of a single soldier who is expected to die in all of the battles.

One could imagine that the source of the ambiguity in the (a) sentences is the availability of long distance Quantifier Raising (QR). The universal quantifier can move by QR over the existential quantifier, and the optionality of this movement, one might think, is the cause of the ambiguity. However, this is most likely not the case. The (b, c) sentences show that the ambiguity in the (a) sentences requires Scope Reconstruction. If QR was sufficient to yield the ambiguity in the (a) sentences, we would expect to find the same kind of ambiguity in the (b,c) sentences. However, the latter sentences are unambiguous. Their meaning is restricted to the implausible interpretation which results from assigning wide

---

<sup>9</sup> Similar data is noted in Aoun (1982), attributed to Rizzi. I changed Lebeaux's examples slightly. The first change has to do with the choice of lexical items. I tried to make the interpretation resulting from wide scope for the matrix subject ( $\exists > \forall$ ) conflict with world knowledge. This change makes the unavailability of the alternative scopal relation in (b,c) ( $\forall > \exists$ ) very striking. The second change is the addition of (3c) and (4b) which contain a bound pronoun (rather than a reflexive) within the dative PP (see also Hornstein (1994:160)). This change is meant to explain why I don't draw the same conclusions that Lebeaux does (see next note).



scope to the existential quantifier. While this restriction is unaccounted for under the assumption that QR alone is the source of the ambiguity in the (a) sentences, the assumption that Scope Reconstruction is needed accounts for the restriction straightforwardly. In the (b) and (c) sentences, the existential quantifier must bind a variable in a position outside its scope as determined by Scope Reconstruction. Hence Scope Reconstruction is impossible.<sup>10</sup> We must assume that the  $(\forall > \exists)$  interpretation of the (a) sentences results from a combination of Scope Reconstruction and short distance QR. The matrix subject receives scope in the position of  $t$  and the universal quantifier receives scope above this position (via QR).

Consider next the contrast between the sentences in (5,6). Some speakers find the (a) sentences slightly marginal. This is accounted for by Weak Crossover (WCO) under the assumption that QR involves A-bar movement. However, the (b) sentences are acceptable. This contrast is explained if we assume that QR is not necessary to get wide scope for the universal quantifier. This assumption, in turn, is explained by the availability of Scope Reconstruction in the (b) sentences.

- (5) a. ??[Someone from his<sub>1</sub> class]<sub>2</sub> shouted to every professor<sub>1</sub> [PRO<sub>1</sub> to be careful].  
 b. [Someone from his<sub>1</sub> class]<sub>2</sub> seems to every professor<sub>1</sub> [t<sub>2</sub> to be a genius].
- (6) a. ??[His<sub>1</sub> father]<sub>2</sub> wrote to every boy<sub>1</sub> [PRO<sub>1</sub> to be a genius].  
 b. [His<sub>1</sub> father]<sub>2</sub> seems to every boy<sub>1</sub> [t<sub>2</sub> to be a genius].

**5.1.1. Syntactic Accounts of Scope Reconstruction:** Under the syntactic accounts of Scope Reconstruction, the ambiguous sentences in (2-4) are disambiguated at LF. Under one disambiguation the Quantifier Phrase is in its surface position and binds a variable in the trace position. Under other disambiguations, the Quantifier Phrase is in one of the intermediate trace positions (or, alternatively, in a position that binds such positions<sup>11</sup>). This is illustrated with the two LF structures of (2a) given in (2a')

---

<sup>10</sup>Lebeaux (1994) draws a more radical conclusion. Given the unavailability of  $(\forall > \exists)$  in (3b), he concludes that BT(A) must be satisfied at LF. Although the conclusion seems plausible to me, I don't think the example bears on it. All we can argue for on the basis of (3b), as is perhaps clearer from (3c) and (4b), is that  $(\forall > \exists)$  in the (a) sentences requires Scope Reconstruction. The facts follow with the addition of a (virtual) tautology that a quantifier cannot bind a variable outside of its scope.

<sup>11</sup> Such binding would be the result of a Quantifier Lowering operation as in May (1977). See Chomsky (1995: 327).

- (2a') LF<sub>1</sub>. [Someone from NY]<sub>1</sub> is very likely [t<sub>1</sub> to win the lottery]  
 LF<sub>2</sub>. is very likely [[Someone from NY] to win the lottery].

Such syntactic reconstruction could be achieved by various mechanisms. Among these mechanisms is quantifier lowering suggested for A-movement by May (1977) and the copy theory of movement suggested for A-bar reconstruction by Chomsky (1993) and applied by many to A-movement (see, among others, Hornstein (1994)). However, at the moment I will abstract away from the details of the implementation (see, however, 6.5-6.6). What matters right now is the property that all the syntactic accounts share; they all assume that Scope Reconstruction involves an LF structure in which the Quantifier Phrase is literally in the reconstructed position. This account of Scope Reconstruction extends straightforwardly to all the cases discussed including the cases of variable binding. I.e., under this account it is straightforward to explain the fact that the (b) sentences in (5,6) do not show a WCO effect.

**5.1.2. Semantic Accounts of Scope Reconstruction:** Semantic accounts assume that syntactic reconstruction is not necessary for Scope Reconstruction. (See Chierchia (1995a), Cresti (1995) and Rullmann (1995).) In other words, they assume that there is a semantic mechanism that yields the two interpretations of sentence such as (2a) from a structure with no syntactic reconstruction such as LF<sub>1</sub> (in (2a')). The existence of such a semantic mechanism is tacitly assumed in the scope principle suggested by Aoun and Li (1994) and explored by Frey (1989)<sup>12</sup>, Kitahara (1994) and Krifka (1998), among others.

The semantic nature of this mechanism has been developed explicitly within frameworks which assume that a semantic type is associated with each syntactic expression. A further assumption is that the sister of a moved constituent is interpreted as a function which can be expressed with  $\lambda$  abstraction over a variable in the trace position. The question is, of course, a variable of what type. The assumption is that in cases in which a Quantifier Phrase undergoes movement, the variable can either range over individuals (i.e. be interpreted as a variable of type *e*) or over generalized quantifiers -- that is, second order predicates (i.e. be interpreted as a variable of type  $\langle et, t \rangle$ ). For an incomplete illustration, consider the two options for interpreting LF<sub>1</sub>. (For a more complete

---

<sup>12</sup>Thanks to Uli Sauerland for conveying to me the contents of this work (in German) which I unfortunately am unable to read.

discussion, see Heim and Kratzer (forthcoming).) These two options are represented in (2'') where  $x$  ranges over individuals and  $Q$  ranges over generalized quantifiers.

- (2'') a. [Someone from NY]  $\lambda x$  (is very likely [ $x$  to win the lottery])  
 b. [Someone from NY]  $\lambda Q$  (is very likely [ $Q$  to win the lottery])

In (2''a), where the variable is of type  $e$ , the sister of the moved Quantifier Phrase is interpreted as a function from individuals to truth values (type  $et$ ). Since the moved Quantifier Phrase is of type  $\langle et, t \rangle$ , the Quantifier Phrase takes its sister as argument. It is easy to see that the resulting interpretation is one in which the existential quantifier has scope over the modal verb. In (2''b), the sister of the quantifier is interpreted as a function from generalized quantifiers to truth-values (type  $\langle \langle et, t \rangle, t \rangle$ ). In this case the Quantifier Phrase is the argument of its sister, and it is easy to see (once we consider lambda-conversion) that the modal verb receives wide scope.<sup>13</sup>

The semantic account of Scope Reconstruction extends straightforwardly to all the cases discussed besides the cases of variable binding via reconstruction (the (b) sentences in (5,6)). A few further assumptions need to be made in order for semantic reconstruction to allow for variable binding (into the reconstructed element). I will not go over the assumptions and instead refer the reader to Engdahl (1986), Sternefeld (1997), Chierchia (1995b) and Sharvit (1997).

**5.1.3. Distinguishing the two Accounts via BT(C):** Consider the structural configuration in (7), in which linear precedence represents c-command. If BT(C) is sensitive to LF structures, the syntactic account predicts that Scope Reconstruction would be impossible in (7). This prediction is stated in (8).

(7) [QP ... $r$ -expression<sub>1</sub>...] <sub>2</sub>.....pronoun<sub>1</sub>.... $t_2$

(8) **Scope Reconstruction feeds BT(C):** Scope Reconstruction should be impossible in the structural configuration in (7).

---

<sup>13</sup>Note that this semantic account, just like the syntactic account, captures straightforwardly the virtual tautology that Scope Reconstruction to a position of a trace is incompatible with the binding of a variable outside the c-command domain of this trace. For example, this account is on a par with the syntactic account with respect to the contrasts in (3-4). See notes 9 and 10.

In order for this prediction to follow under the semantic account of Scope Reconstruction, one would have to assume that condition C makes reference to the semantic type of traces and that the LF in (7) is ruled out iff the semantic type of the trace is  $\langle et, t \rangle$ . (See Sternefeld (1997).) This, however, cannot be considered an explanation. The necessary assumption is a post hoc stipulation which doesn't tell us why things are the way they are. In particular, it would be just as plausible to make the opposite assumption (i.e., that BT(C) rules out (7) when the semantic type of the trace is  $e$ ) from which it would follow that Scope Reconstruction is obligatory in (7).<sup>14</sup>

Under the syntactic account of Scope Reconstruction (8) is *explained*. Condition C receives the simple definition based on constructions for which the structural analysis is not debated (constructions without movement). Under the natural assumption that an interpretive principle (such as Binding Theory) is sensitive to LF structures, (8) follows. Therefore, if we can show that the prediction holds, we will have an argument in tandem for both the syntactic account and the assumption that BT(C) applies at LF. In the following sections I will demonstrate that the prediction holds and that indeed we have the argument.<sup>15</sup>

---

<sup>14</sup>Gennaro Chierchia (personal communication; attributing the idea to work in progress by Yael Sharvit) points out that the correlation in (8) could follow from Semantic Reconstruction if (certain) BT(C) effects are viewed as the results of a preference for variable binding over coreference (as in Reinhart's Rule I, see Grodzinsky and Reinhart (1993)). In particular, he suggests that if the semantic type of the trace yields a "reconstructed" interpretation, variable binding is possible and thus coreference is ruled out. Although this is conceivable, I think that many questions need to be answered before we know whether this is a real alternative. For example, what is the status of a basic semantic type such as  $\langle et, t \rangle$  which allows for scope reconstruction with no variable binding reconstruction (see 2''b)? While it is possible for Semantic Reconstruction to predict that BT(C) will be affected by variable binding reconstruction, is there a non-stipulative way to predict that it will be affected by simple Scope Reconstruction (via a trace of type  $\langle et, t \rangle$ )? Another question relates specifically to A-bar movement; is there a way for the proposed alternative to account for the fact that under normal circumstances QR (as well as *wh-movement*) does not obviate BT(C) effects (chapter 6)?

<sup>15</sup>Romero (1996) presents additional arguments against the semantic account. In particular, she spells out the modifications in the Binding Theory that the semantic approach would require and displays their stipulative nature. Furthermore, she develops additional unwelcome results related to the interpretation of sluicing.

## 5.2. A-bar Reconstruction

In this section I will show that the prediction in (8) holds for A-bar movement. The section has three parts. Section 5.2.1 will focus on *how many* questions and will contain an elaboration on data from Heycock (1995). Section 5.2.2 will explore standard constituent questions. The discussion will be based on data from Lebeaux (1990) which uses variable binding as diagnostic of Scope Reconstruction. Section 5.2.3 will continue the discussion of standard constituent questions. This time, indirect binding via adverbs of quantification will serve as the diagnostic for Scope Reconstruction.

Before we begin, a point of caution is in order. It is well known that certain cases of A-bar movement cannot bleed BT(C) (van Riemsdijk and Williams (1981), Freidin (1986) and Lebeaux (1988)). This inability holds independently of Scope. Consider for example Lebeaux's pair in (9).

- (9) a. [Which argument that John<sub>1</sub> made] did he<sub>1</sub> believe t?  
 b. ??/\*[Which argument that John<sub>1</sub> is a genius] did he<sub>1</sub> believe t?

(9b) is bad independently of Scope Reconstruction. The discussion in this section will focus on cases such as (9a) in which A-bar movement *does* bleed BT(C). We will see that in these cases bleeding is incompatible with Scope Reconstruction. In other words, we will see that if an A-bar construction of the type in (7) is acceptable, it is disambiguated in favor of the non-Scope-Reconstructed interpretation.

**5.2.1. How Many questions - an elaboration on Heycock (1995):** Consider *how many* questions of the sort in (10). A plausible analysis of such questions assumes that the *Wh*-phrase *how many NP* has two parts. One part consists of the *Wh*-word *how* (which could be paraphrased as *what n*) and the other consists of the DP *many NP* (see Frampton (1991), Cresti (1995), Rullman (1995)). Roughly speaking a *how many* question asks for an integer *n*, such that *n* many people satisfy a certain property. This is illustrated in (10') and schematized in (11).

- (10) [How many people]<sub>1</sub> did you meet t<sub>1</sub> today.

- (10') How *n*: *n* many people λ<sub>t</sub> you met t today.  
*What is the number n, s.t. there are n many people that you met today*

- (11) [How many NP]<sub>1</sub>  $\phi(t_1)$   
 How  $n$ :  $n$  many NP  $\lambda t \phi(t)$ <sup>16</sup>

Consider now what happens when scope-sensitive elements intervene between the final landing site of the moved constituent and its trace position. In such a case, the scope of the quantificational DP, *many NP*, can be construed either above or below this element (see, among others, Longobardi (1987), Cresti (1995)). This is demonstrated by the question in (12).

- (12) How many people did Mary decide to hire.

Two readings:

- a. *many* > *decide*:  
*What is the number  $n$ , s.t.*  
*there are  $n$  many people  $x$ , s.t.*  
*Mary decided to hire  $x$ .*
- b. *decide* > *many*:  
*What is the number  $n$ , s.t.*  
*Mary decided to hire  $n$  many people*

The two readings of the question become visible once we consider situations in which they would demand different answers.<sup>17</sup> Consider the following scenario:

(S1) After a day of interviews, Mary finds 7 people who really impress her, and she decides to hire them. None of the other people impress her. However she knows that she needs more than 40 people for the job. After thinking about it for a while she decides to hire 50 -- the 7 that she likes and 43 others to be decided by a lottery.

---

<sup>16</sup>Obviously, stating the semantic type of  $t$  would beg the question we are investigating. The claim that Scope Reconstruction is syntactic (for which I argue) amounts to the claim that the trace left by DPs always ranges over individuals (of type  $e$ ). (See section 5.4.)

<sup>17</sup>As Kai von Stechow (personal communication) has pointed out, the two readings become even more visible once we insert the adverb *already*. This adverb tends (though doesn't have to) mark the scope of the quantificational DP. Thus, *How many people did Mary already decide to hire* prefers the interpretation in (12a) and would, thus, receive 7 as its natural answer under (S1). The reader who has problems with the scope judgments that follow might try to make use of the adverb.

It is clear that under (S1) there are two appropriate answers for (12). One answer is seven, which corresponds to the interpretation of (12a), and the other, which corresponds to the meaning of (12b), is fifty.<sup>18</sup>

The ambiguity of (12) shows that the DP *many NP* can undergo Scope Reconstruction in *how many* questions.<sup>19</sup> In this section I will present evidence from Heycock (1995) and expand on it to demonstrate that such reconstruction obeys the predictions in (8). The evidence will argue that Scope Reconstruction in *how many* questions should be dealt with by syntactic mechanisms.

Consider the sentences in (13-15).<sup>20</sup> In none of these sentences is BT(C) an issue. The sentences, however, differ minimally in a way that Heycock exploits to test the prediction in (8).<sup>21</sup> The semantics of the embedded predicates in the (a) sentences forces,

---

<sup>18</sup>Noam Chomsky (personal communication) points out a possible caveat. Tense is not represented in (12). However, if we take tense into account, it is conceivable that the two interpretations would result from variability in the interpretation of Tense (with no recourse to scope). More specifically, the two answers (given S1) could correspond to the two moments in time in which Mary made her decisions.

However, I think there are good reasons to believe that there is a pragmatic principle which forces the answer to correspond to the later moment in time. Take a sentence parallel to (12) which is not scopally ambiguous such as *Which people did Mary decide to hire?* The proposal about tense would incorrectly predict a similar ambiguity in this sentence. I believe the prediction is incorrect because of the pragmatic principle. Without any clear context, it makes no sense to give an answer that would correspond to tentative decisions that Mary made along the way.

Note that not much bears on this pragmatic explanation. The arguments in this section can be restated with the verb *want*, which is ambiguous in the present tense (e.g. How many people does Mary want to hire). Furthermore, the French examples in note 18 make it pretty clear that the ambiguity in (12) has to do with scope. Thanks to Kai von Steiger for a helpful discussion of this issue.

<sup>19</sup>As noted in Dobrovie-Sorin (1992) (and Heycock (1995)), French has overt forms which disambiguate in favor of Scope Reconstruction. While the (a) sentences in (i) and (ii) are equivalent to their English counterparts, the (b,c) sentences are unambiguous (with narrow scope for *many*). Thus, given S1, (ib,c) have fifty as their only answer. (Marie Claude Boivin, personal communication.)

(i) a. Combien de personnes Marie a-t-elle décidé d'engager?  
 b. ? Combien Marie a-t-elle décidé d'engager de personnes?  
 c. Combien Marie a-t-elle décidé d'en engager?  
 (ii) a. Combien de livres Marie a-t-elle décidé d'acheter?  
 b. ? Combien Marie a-t-elle décidé d'acheter de livres?  
 c. Combien Marie a-t-elle décidé d'en acheter?

<sup>20</sup>David Pesetsky suggested the use of the pair in (13) for the exposition of Heycock's results.

<sup>21</sup> I am not completely sure whether what follows is entirely faithful to Heycock. The reason for this uncertainty is that Heycock alternates between an account of the ambiguity in (12) in terms of scope and an account in terms of a notion of referentiality (that will

or at least highly prefers, Scope Reconstruction. The (b,c) sentences, by contrast, are compatible with the non-reconstructed reading.<sup>22</sup> The sentences in (b,c) are ambiguous along the lines of (12), whereas the (a) sentences are limited to the interpretation in which the DP *many NP* has narrowest scope.

- (13) a. [How many stories] is Diana likely to invent t? (*likely* > *many*;  
\**many* > *likely*)  
b. [How many stories] is Diana likely to re-invent t? (*likely* > *many*;  
*many* > *likely*)
- (14) a. [How many houses] does John think you should build t? (*think* > *many*;  
\**many* > *think*)  
b. [How many houses] does John think you should re-build t? (*think* > *many*;  
*many* > *think*)  
c. [How many houses] does John think you should demolish t? (*think* > *many*;  
*many* > *think*)
- (15) a. [How many papers that he<sub>1</sub> writes]  
does John<sub>1</sub> think t will be published?<sup>23</sup> (*think* > *many*;  
\**many* > *think*)  
b. [How many papers that he<sub>1</sub> wrote]  
does John<sub>1</sub> think t will be published? (*think* > *many*;  
*many* > *think*)

To see the contrast between the (a) and the (b) sentences, let's focus on (13b). Consider what an interpretation without Scope Reconstruction would be like. Such an interpretation would presuppose that there are certain stories such that Diana is likely to invent them. However, such a presupposition is virtually a contradiction; we think about the objects of invention as being created at the time of invention and we, therefore, can't talk about these objects at earlier moments, hence the weirdness of *#John will invent this story*, *#Which of these stories is John likely to invent?*

---

extend to an explanation of why VPs must reconstruct). Under the referentiality account that Heycock suggests, "non-referential" phrases (whatever exactly this means, see Cinque (1991), Rizzi (1990), Frampton (1991)) must show connectedness effects (i.e. BT Reconstruction). The reason we get connectedness effects in the (b) sentences is that the DP is interpreted non-referentially. However, once we consider the structural configurations in (22) below we will see that scope is not only the clearer notion, it is also the empirically appropriate one.

<sup>22</sup>The sentences in (15b,c) are actually three-ways ambiguous. However, for the moment, we can ignore the intermediate scope (*think* > *many* > *should*).

<sup>23</sup>The sentence should be understood with a future interpretation for the tense in the relative clause. Wide scope for *many* would require John to have thoughts about specific papers. However, at the time of thinking there are no papers to have thoughts about; the papers will come to exist only in the future.



As Heycock notes, this difference allows us to test the prediction in (8). Consider the sentences in (16-18). Only in the (a) cases, in which Scope Reconstruction is forced, do we get a BT(C) effect. (Only the (a) cases are unacceptable.)<sup>24</sup>

- (16) a. \*[How many stories about Diana's<sub>1</sub> brother] is she<sub>1</sub> likely to invent t?  
b. [How many stories about Diana's<sub>1</sub> brother] is she<sub>1</sub> likely to re-invent t?<sup>25</sup>
- (17) a. \*[How many houses in John's<sub>1</sub> city] does he<sub>1</sub> think you should build t?  
b. [How many houses in John's<sub>1</sub> city] does he<sub>1</sub> think you should re-build t?  
c. [How many houses in John's<sub>1</sub> city] does he<sub>1</sub> think you should demolish t?
- (18) a. \*[How many papers that John<sub>1</sub> writes]  
does he<sub>1</sub> think t will be published?  
b. [How many papers that John<sub>1</sub> wrote]  
does he<sub>1</sub> think t will be published?

We can also demonstrate that the unacceptability of the (a) cases is related to BT(C) when we compare these sentences with the sentences in (19). The latter sentences demonstrate that aside from BT(C) there is nothing the matter with the coindexation in the (a) sentences.

- (19) a. [How many stories about her<sub>1</sub> brother] is Diana<sub>1</sub> likely to invent t?  
b. [How many houses in his<sub>1</sub> city] does John<sub>1</sub> think you should build t?

---

<sup>24</sup>In section 6.5 I will follow Lebeaux (1988) in assuming that A-bar movement can bleed BT(C) only if an adjunct is inserted after movement. The discussion in 6.5 will, thus, imply that the PPs in (16-17) are adjuncts. This implication, which might seem problematic at first sight, is argued for in Heycock (1995). See Scütze (1995) to appreciate the hairiness of the complement adjunct distinction (at least within the nominal domain).

I would like to point out that a few speakers I've consulted find the (b) sentences in (16) and (17) slightly degraded. These speakers find the (a) sentences still worse. It seems plausible to suggest that these speakers prefer to analyze the PP as a complement. Still, they can marginally analyze it as an adjunct, which would allow BT(C) obviation in (b) but not in (a).

<sup>25</sup> There is a potential problem with (16a). The problem is that this sentence might be bad independently of Reconstruction. To see this consider the contrast in (i) and (ii). This contrast might be accounted for by postulating a PRO in the subject of the NPs in the (b) sentences along the lines of Chomsky (1986b). (See also Higgenbotham (1983) and Williams (1985, 1987)). If the correct account is along these lines, (16a) would have a BT(C) violation (in the moved position) irrespective of whether or not there is Reconstruction. This confound is overcome in (17) and (19).

- (i) a. Diana<sub>1</sub> objected to many stories about her<sub>1</sub>.  
a. \*Diana<sub>1</sub> invented many [PRO<sub>1</sub> stories about her<sub>1</sub>].
- (ii) a. Clifford<sub>1</sub> expected many lies about him<sub>1</sub> to be effective.  
a. \*Clifford<sub>1</sub> is planning to come up with many [PRO<sub>1</sub> lies about him<sub>1</sub>].



If (8) is right, there are two predictions for the configuration (7). On the one hand, QP is obliged to take scope over all of the scope-bearing elements c-commanded by the pronoun. On the other hand, such an obligation does not hold with respect to the scope-bearing elements that c-command the pronoun. This is stated in (22).

- (22) **Predictions of (8):** a. In (23), QP must take scope over the scope-bearing element  $SB^1$ .  
 b. In (24), QP need not take scope over the scope-bearing element  $SB^2$ .<sup>27</sup>

(23)  $[QP \dots r\text{-expression}_1 \dots]_2 \dots \text{pronoun}_1 \dots SB^1 \dots t_2$

(24)  $[QP \dots r\text{-expression}_1 \dots]_2 \dots SB^2 \dots \text{pronoun}_1 \dots t_2$

(13-21) demonstrate that when independent factors force QP to take narrow scope with respect to  $SB^1$ , the result is ill-formed. I will now try to show that the predictions of (22) are attested also when these factors are not active.

Compare the pairs in (25) and (26).

- (25) a. [How many slides of Jonathan's<sub>i</sub> trip to Kamchatka] did he<sub>i</sub> decide to show t at the party? (*many* > *decide*) (\**decide* > *many*)  
 b. [How many slides of his<sub>i</sub> trip to Kamchatka] did Jonathan<sub>i</sub> decide to show t at the party? (*many* > *decide*) (*decide* > *many*)
- (26) a. How many people from Diana's<sub>1</sub> neighborhood does she<sub>1</sub> think t are at the party? (*many* > *think*) (\**think* > *many*)  
 b. How many people from her<sub>1</sub> neighborhood does Diana<sub>1</sub> think t are at the party? (*many* > *think*) (*think* > *many*)

The (a) sentences in these pairs are instantiations of the structural configuration in (23), with the modal verbs *decide* and *think* standing for  $SB^1$ . In these sentences *many NP* must have wide scope over  $SB^1$  as predicted by (22). This can be seen when we compare the possible answers to (a) and to (b) under the crucial scenarios.

---

<sup>27</sup>The prediction in (b) depends on the additional assumption that there is a position for reconstruction between  $SB^2$  and  $\text{pronoun}_1$ . This assumption is probably uncontroversial for the cases discussed in this section, since for all of these cases there is a CP node intervening between the two positions. In the next sub-section, there will be evidence for more reconstruction positions (see section 5.2.2.2.).

The two readings of (25b) are paraphrased in (25'). The difference between them is illustrated by the possible answers given the scenario (S2):

(25') Two readings:

a. *many* > *decide*:

*What is the number  $n$ , s.t.*

*there are  $n$  many slides of the trip to Kamchatka  $x$ , s.t.*

*Jonathan decided to show  $x$  at the party* (answer given S2: 52)

b. *decide* > *many*:

*What is the number  $n$ , s.t.*

*Jonathan decided to show  $n$  many slides at the party* (answer given S2: 100)

(S2) Jonathan wants to show slides from his trip to Kamchatka at a party. He tries to figure out how many slides he can show within an hour. After consulting with his roommate, Uli, he decides to show 100 slides (out of the 1000 he has). Now it's time to choose the actual slides. After an hour of internal debate, he decides on 52 slides that he really likes and prepares them for display. The remaining 48 slides will be chosen at random at the time of the party.

(25a) by contrast can only have the interpretation of (25'a); the only possible answer to this question given S2 is 52.

(26b) can have either the (*think* > *many*) interpretation, the interpretation of (21c), in which Diana doesn't need to have thoughts about any particular person, or the (*many* > *think*) interpretation, in which it is presupposed that there are certain people which Diana thinks are at the party and the number of those is inquired. Only the latter interpretation is available for (26a). (I invite the reader to construct the relevant scenarios.)

(25-26) show that QP in (23) cannot take scope under  $SB^1$ . Now I would like to show that it can take scope under  $SB^2$  in (24). Furthermore, I would like to show that when  $SB^2$  and  $SB^1$  appear in the same construction (when we put (23) and (24) together), QP can have scope under the former but must take scope over the latter.<sup>28</sup>

Consider the contrast between the sentences in (27). (27a) is an instantiation of (23) and (27b) is an instantiation of (24). (*Decide* is the instantiation of both  $SB^2$  and  $SB^1$ .) As (8) (/(22)) predicts, Scope Reconstruction is possible only in (27b). We have

---

<sup>28</sup> The importance of this prediction is in demonstrating that the phenomena should be accounted for by reference to scope and not to a notion such as referentiality as is perhaps suggested by Heycock (see note 21). This is also the logic of Frampton's argument that scope (and not referentiality) is the relevant notion for the account of certain weak island phenomena (Frampton (1991)).

already concluded that Scope Reconstruction is impossible in (27a (=25a)) with the assistance of (S2). (S3) is a minimal variation on (S2) (with the changes in boldface), which demonstrates the contrast between the two questions. Under (S3) there are two possible answers to (27b), which correspond to the two possible scope relations.

- (27) a. [How many slides of Jonathan's; trip to Kamchatka] did he; decide to show t at the party? *(many > decide) (\*decide > many)*  
 b. [How many slides of Jonathan's; trip to Kamchatka] did Susi decide that he; would show t at the party? *(many > decide) (decide > many)*

(S3) Jonathan wants to show slides from his trip to Kamchatka at a party. **Susi** tries to figure out how many slides he can show within an hour. After consulting with **their** roommate, Uli, **she decides that Jonathan will** show 100 slides (out of the 1000 he has). Now it's time to choose the actual slides. After an hour of internal debate, **she** decides on 52 slides that **she** really likes and prepares them for **Jonathan's** display. The remaining 48 slides will be chosen at random at the time of the party.

Consider now (S4) which starts with (S2) and then has a short continuation:

(S4) Jonathan wants to show slides from his trip to Kamchatka at a party. He tries to figure out how many slides he can show within an hour. After consulting with his roommate, Uli, he decides to show 100 slides (out of the 1000 he has). Now it's time to choose the actual slides. After an hour of internal debate, he decides on 52 slides that he really likes and prepares them for display. The remaining 48 slides will be chosen at random at the time of the party.

After all of this Jonathan tells Susi about his two decisions, and wishes to show her the 52 chosen slides. He shows her 30 of the slides, at which point Susi gets bored and asks to go to sleep. Jonathan tells her that there are 22 more slides to see, and Susi says that she will see them during the party.

### The relevant facts:

1. Susi knows
  - a. that Jonathan decided to show 100 slides at the party.
  - b. that there are 52 slides such that J. decides to show them.
2. There are 30 slides such that Susi knows that J. decided to show them.

With the use of (S4) we can see that (28) is three-ways ambiguous; under (S4), (28) has three possible answers.

- (28) How many slides did Susi know that Jonathan decided to show at the party?

Answers:

- a. 30 (*many* > *know* > *decide*)
- b. 52 (*know* > *many* > *decide*)
- c. 100 (*know* > *decide* > *many*)

Now consider the contrast in (29). (29b) is three-ways ambiguous, just as (28) is. (29a), by contrast, is only two-ways ambiguous. In (29a), *many* can take scope either above or below *know* (= SB<sup>2</sup>). However, the scope of *many* relative to *decide* (= SB<sup>1</sup>) is fixed (*many* > *decide*); under (S4), the only possible answers to (29a) are 30 and 52. This is exactly the prediction of (22).

- (29) a. [How many slides of Jonathan's<sub>i</sub> trip to Kamchatka] did Susi know that he<sub>i</sub> decided to show *t* at the party?  
 b. [How many slides of his<sub>i</sub> trip to Kamchatka] did Susi know that Jonathan<sub>i</sub> decided to show *t* at the party?

We have thus seen that the predictions of syntactic reconstruction hold for *how many* questions. Scope Reconstruction has consequences for BT(C) in exactly the manner predicted by (8).<sup>29</sup>

**5.2.2. Variable binding - an elaboration on Lebeaux (1990):** Consider the option for variable binding in the constituent questions in (30). As is well-known, this option is available only if the trace of the *wh*-element is c-commanded by the binder of the variable. This is illustrated by the unacceptability of the questions in (31).

- (30) a. Which of his<sub>i</sub> students did every professor<sub>i</sub> talk to *t*?  
 b. Which student of his<sub>i</sub> did no professor<sub>i</sub> talk to *t*?  
 c. Which student of his<sub>i</sub> did you think every professor<sub>i</sub> talked to *t*?  
 d. Which of his<sub>i</sub> students did you think no professor<sub>i</sub> talked to *t*?
- (31) a. \*Which of his<sub>i</sub> students *t* talked to every professor<sub>i</sub>?  
 b. \*Which student of his<sub>i</sub> *t* talked to no professor<sub>i</sub>?  
 c. \*Which student of his<sub>i</sub> did you think *t* talked to every professor<sub>i</sub>?  
 d. \*Which of his<sub>i</sub> students did you think *t* talked to no professor<sub>i</sub>?

---

<sup>29</sup>Although the questions in (29) seem extremely complex, speakers' judgments converge. The scenario in (S4) was read to 5 speakers and they all agreed that 100 is an appropriate answer only to (29b) whereas 30 and 52 are possible answers to both (29a) and (29b).

**5.2.2.1: The correlation with BT(C):** Just as in the previous cases we've looked at, there are two possible approaches. Under one approach, which is due to Engdahl (1980), part of the *wh*-element is in the trace position, and this syntactic configuration allows for variable binding (Syntactic Reconstruction). Under another approach, due to Engdahl (1986), various semantic mechanisms are postulated to yield the semantic effects of variable binding without actual reconstruction (Semantic Reconstruction).

These two approaches can be distinguished by the prediction (of the syntactic approach) that Scope Reconstruction feeds BT(C). If the cases in (30) require reconstruction in the syntax, there should be consequences for BT(C). If, however, there are semantic mechanisms that allow for the interpretations in (30) without actual reconstruction, there should be no such consequences. Lebeaux's pair in (32) demonstrates that the prediction of syntactic reconstruction is correct.

- (32) a. [Which (of the) paper(s) that  $he_i$  gave to Ms. Brown $_j$ ] did every student $_i$  hope  $t'$  that she $_j$  will read  $t'$ ?  
 b. \*[Which (of the) paper(s) that  $he_i$  gave to Ms. Brown $_j$ ] did she $_j$  hope  $t'$  that every student $_i$  will revise  $t'$ ?

In both sentences in (32) part of the *wh*-phrase must undergo Scope Reconstruction to a position c-commanded by the Quantifier Phrase *every student* (the antecedent of the bound-variable *he*). In (32b), *every student* is c-commanded by the pronoun *she* which is, in turn, co-indexed with *Ms. Brown*. Therefore, in (32b) Scope Reconstruction yields a BT(C) effect. In (32a), *every student* is not c-commanded by the pronoun, *she*, and therefore there is a position for reconstruction (perhaps the position of  $t'$ ) which is within the scope of the antecedent of *he*, but not low enough to yield a BT(C) effect. The contrast is thus explained under the assumption that Scope Reconstruction feeds BT(C).

In fact, once we understand the logic of (32), we see that the conclusion regarding the relation between Scope Reconstruction and BT(C) can be demonstrated with additional examples, some of which come close to real minimal pairs. The logic is basically the logic of (7) and (8), repeated below:

- (7) [QP ...r-expression $_1$ ...]  $_2$ .....pronoun $_1$ .... $t_2$

- (8) **Scope Reconstruction feeds BT(C):** Scope Reconstruction should be impossible in the structural configuration in (7).

(32b) is bad because it is an instance of (7) which requires Scope Reconstruction to a position below the pronoun. We know that Scope Reconstruction is necessary because there is a quantifier below the pronoun that must have a variable dominated by QP within its scope. In (32a), the quantifier is above the pronoun and hence Scope Reconstruction need not bring about a BT(C) effect.

The predictions of (8) for the constructions that Lebeaux investigated can be summarized by the schemes in (33), in which the underlined blanks represent potential reconstruction positions. Instances of (33a) should be acceptable because they do not require Scope Reconstruction of the *wh*-element to a position below pronoun<sub>j</sub>; there could be reconstruction to a position between the pronoun and QP in which the variable is bound and nevertheless BT(C) is satisfied.<sup>30</sup> Instances of (33b) should be unacceptable because Scope Reconstruction of the *wh*-element must be to a position below QP<sub>i</sub> which is, in turn, below pronoun<sub>j</sub>; any form of reconstruction that would allow the variable to be bound will necessarily bring about a BT(C) effect.

- (33) a. [Which .... pronoun<sub>j</sub> .... r-expression<sub>j</sub>] .....QP<sub>i</sub> .....      .pronoun<sub>j</sub>.....\*      .  
 b. \*[Which ...pronoun<sub>j</sub> .... r-expression<sub>j</sub>] .....pronoun<sub>j</sub>...\*      .QP<sub>i</sub> .....       \*      ..

The important difference between the schemes in (33) is that in (33b), and only in (33b), does the kind of Scope Reconstruction that is needed bring about a BT(C) effect. However, there is an additional difference between the structures. In (33b), a more radical kind of Scope Reconstruction is needed. We should therefore add a control to the paradigm. We need to compare (33b) with (34) which needs the same kind of Scope Reconstruction but is irrelevant to BT(C).<sup>31</sup>

- (34) [Which ...pronoun<sub>j</sub> .... pronoun<sub>j</sub>] .....r-expression<sub>j</sub>...\*      .QP<sub>i</sub> .....      ..

<sup>30</sup>Note that the prediction holds only under the assumption that there is a reconstruction position between QP<sub>i</sub> and pronoun<sub>j</sub>. See section 5.2.2.2.

<sup>31</sup> From the tests I conducted it seems that it is really important to put pronoun<sub>j</sub> and QP<sub>i</sub> in (33) as close as possible to each other, so as to minimize the differences between the sentences with respect to the distance between the bound variable and its antecedent. It seems that when the distance is very great, (33b) and (34) involve a terrible parsing load. This parsing load makes the judgment very difficult and it is hard to detect the effects of BT(C). Thus, although (i) seems better than (32b), it seems worse than (32a).

(i) ?? [Which (of the) paper(s)that he<sub>j</sub> gave to her<sub>j</sub>] did Ms. Brown<sub>j</sub> hope  
 t' that every student<sub>i</sub> will revise t?



Even with the control, the tests continue to confirm the predictions of syntactic reconstruction. This is demonstrated in (35-37). The (a) and (b) examples are instantiations of (33a) and (33b) respectively; the (c) examples are instantiations of the control in (34). The results show that Scope Reconstruction, which is diagnosed by variable binding, feeds BT(C).

- (35) a. [Which of the books that he<sub>i</sub> asked Ms. Brown<sub>j</sub> for]  
           did every student<sub>i</sub> \_\_\_ get from her<sub>j</sub> \_\_\_\* ?  
       b. \*[Which of the books that he<sub>i</sub> asked Ms. Brown<sub>j</sub> for]  
           did she<sub>j</sub> \_\_\_\* give every student<sub>i</sub> \_\_\_\* ?  
       c. [Which of the books that he<sub>i</sub> asked her<sub>j</sub> for]  
           did Ms. Brown<sub>j</sub> \_\_\_\* give every student<sub>i</sub> \_\_\_ ?
- (36) a. [Which (of the) paper(s) that he<sub>i</sub> wrote for Ms. Brown<sub>j</sub>]  
           did every student<sub>i</sub> \_\_\_ get her<sub>j</sub> \_\_\_\* to grade?  
       b. \*[Which (of the) paper(s) that he<sub>i</sub> wrote for Ms. Brown<sub>j</sub>]  
           did she<sub>j</sub> \_\_\_\* get every student<sub>i</sub> \_\_\_\* to revise?  
       c. [Which (of the) paper(s) that he<sub>i</sub> wrote for her<sub>j</sub>]  
           did Ms. Brown<sub>j</sub> \_\_\_\* get every student<sub>i</sub> \_\_\_ to revise?
- (37) a. [Which (of the) paper(s) that he<sub>i</sub> gave Ms. Brown<sub>j</sub>]  
           did every student<sub>i</sub> \_\_\_ ask her<sub>j</sub> to read \_\_\_\* carefully?  
       b. \*[Which (of the) paper(s) that he<sub>i</sub> gave Ms. Brown<sub>j</sub>]  
           did she<sub>j</sub> \_\_\_\* ask every student<sub>i</sub> to revise \_\_\_\* ?  
       c. [Which (of the) paper(s) that he<sub>i</sub> gave her<sub>j</sub>]  
           did Ms. Brown<sub>j</sub> \_\_\_\* ask every student<sub>i</sub> to revise \_\_\_ ?

**5.2.2.2. The multitude of intermediate landing sites:** In the previous subsection, we have used the scheme in (33) (repeated below) to argue that Scope Reconstruction feeds BT(C).

- (33) a. [Which .... pronoun<sub>i</sub> .... r-expression<sub>j</sub>] .....QP<sub>i</sub> .....pronoun<sub>j</sub>....\*.  
       b. \*[Which ...pronoun<sub>i</sub> .... r-expression<sub>j</sub>] .....pronoun<sub>j</sub>...\*.QP<sub>i</sub> .....\*..

A crucial assumption of the investigation was that in (33a) there is an intermediate position for Scope Reconstruction between QP<sub>i</sub> and pronoun<sub>j</sub>. Now, I would like to point out that one could conduct an investigation in the other direction. In particular (assuming that the arguments in this chapter are compelling) one can attempt to take for granted the assumption that Scope Reconstruction feeds BT(C), and use the scheme in (33) to test what

type of intermediate Scope Reconstruction positions are available. Such an investigation is beyond the scope of this chapter. Nevertheless, I would like to point out one conclusion that would necessarily follow. This conclusion, although irrelevant at this point in the discussion, will play a crucial role in later sections. (See, in particular, 6.4. See also 2.1.)

Consider the acceptability of the (a) sentences in (36-37). Under the assumption that Scope Reconstruction feeds BT(C), the grammatical status of these sentences forces the conclusion that there is a reconstruction position between the subject and the object. To illustrate this, I repeat the examples below with the crucial reconstruction position in brackets:<sup>32</sup>

(36a) [Which (of the) paper(s) that he<sub>i</sub> wrote for Ms. Brown<sub>j</sub>]  
did every student<sub>i</sub> [\_\_\_\_] get her<sub>j</sub> \*\_\_to grade?

(37a) [Which (of the) paper(s) that he<sub>i</sub> gave Ms. Brown<sub>j</sub>]  
did every student<sub>i</sub> [\_\_\_\_] ask her<sub>j</sub> to read \*\_\_carefully?

If we believe that reconstruction should follow from the copy theory of movement (Chomsky (1993)), we have a direct argument that there must be an intermediate landing site for A-bar movement between the subject and the object. This is the landing site that I assumed in chapter 2 for obligatory QR. Furthermore, I present additional arguments for this landing site when I discuss interactions between BT(C) and Antecedent Contained Deletion (6.4). For presentational purposes I will assume that the landing site is adjunction to VP, along the lines of Chomsky (1986a).<sup>33</sup>

**5.2.3. Unselective Binding:**<sup>34</sup> In the previous sub-section we have used variable binding as diagnostic of Scope Reconstruction. We have seen that when part of a

---

<sup>32</sup>One might suggest that there are fewer reconstruction positions, and that QR over higher reconstruction positions (than the bracketed ones in the (a) sentences of (35-37)) allows for variable binding. However, binding from QRed positions should be ruled by WCO. Furthermore, without WCO it is hard to see how the (b) sentences could be ruled-out.

<sup>33</sup>Nissenbaum (1998) provides independent evidence for VP adjoined traces in A-bar movement. From a rudimentary investigation of constructions that fall under the scheme in (33) it seems to me that a stronger conclusion will follow. In particular, I believe one can construct an argument for the existence of intermediate landing sites in every maximal projection (See Fox (1998b).)

<sup>34</sup>The idea of using unselective binding as diagnostic of Scope Reconstruction was inspired by Chapter 3 of Chierchia (1995a) (although the construction I test and the

constituent *Wh*-question needs to be reconstructed for variable binding there are consequences for BT(C). In this sub-section we will see that the same holds for unselective binding.

Consider the sentences in (38) and (39). In (38) we can get an interpretation in which the indefinite is bound by the unselective binder. In (39), this interpretation is unavailable. To see what is meant, let's focus on the (a) sentences. (38a) has an interpretation under which the indefinite *an artist* is bound by the unselective binder, *usually*. Under this interpretation the question can be paraphrased as something like *which are the types of friends<sub>x</sub> s.t. you said that for most artists<sub>y</sub> x-type friends of y are available?* The question in (39a) does not have a comparable interpretation.

- (38). a. [Which friends of an artist<sub>1</sub> ]  
           [did you say t are usually<sub>1</sub> available]?  
       b. [Which of the people that an artist<sub>1</sub> meets]  
           [did you say t never<sub>1</sub> impress him<sub>1</sub>]?  
  
 (39) a. \*[Which friends of an artist<sub>1</sub> ]  
           [t said that they are usually<sub>1</sub> available]?  
       b. \*[Which of the people that an artist<sub>1</sub> meets]  
           [t said that they never<sub>1</sub> impress him<sub>1</sub>]?

The contrast follows straightforwardly from the assumption that an indefinite which is bound by an adverb of quantification must be within its scope at LF.<sup>35</sup> In (38) the trace of the *Wh*-phrase is within the scope of the adverb and the necessary configuration can be achieved via Scope Reconstruction. In (39), the trace of the *Wh*-phrase is outside the scope of the adverb and the necessary configuration for binding cannot be achieved.

The sentences in (38) and (39) contrast with respect to the availability of the intended interpretation. We can make the judgment sharper by considering constructions in which the intended interpretation is the only one available. Specifically, we will consider constructions in which independent factors require that the indefinite be bound by the adverb of quantification. In such constructions, the parallels of (38) will be ungrammatical independently of interpretation.

---

conclusion I draw are very different). Thanks to Orin Percus for help in constructing the experimental paradigms.

<sup>35</sup>Many different theories have been proposed for the binding of indefinites by adverbs of quantification (i.e. for the Quantificational Variability of indefinites). I think that all the theories share the assumption that the indefinite must be within the scope of the adverb.

Kratzer (1995a) observes that in individual-level predicates, an adverb of quantification must co-occur with an indefinite. This observation is demonstrated by the contrast in (40). (40a) is ungrammatical because there is no variable that is supplied to the adverb of quantification (vacuous quantification). (40b) is licensed because the indefinite supplies the variable.

- (40) a. \*John usually<sub>1</sub> knows French.  
b. A Moroccan<sub>1</sub> usually<sub>1</sub> knows French

Consider now the contrast in the grammaticality of the questions in (41). In (41a,b) the indefinite can get into the scope of the adverb via Scope Reconstruction. In (41c), this is impossible since the trace of the *Wh*-element is outside the scope of the adverb. (41c) is ungrammatical because binding of the indefinite is on the one hand impossible and on the other hand required (given the ban on vacuous quantification).

- (41) a. [Which residents of a French town<sub>1</sub>] [t usually<sub>1</sub> know English]?  
b. [Which residents of a French town<sub>1</sub>]  
[did you say t usually<sub>1</sub> know English]?  
c. \*[Which residents of a French town<sub>1</sub>]  
[t said that they usually<sub>1</sub> know English]?

With this much at hand, we have an additional test for the prediction that Scope Reconstruction feeds BT(C). In this test the binding of an indefinite by an adverb of quantification will serve as diagnostic of Scope Reconstruction and the consequences for BT(C) will be examined. The examination will be based on the assumption that, as far as BT(C) is concerned, an indefinite acts like an *r*-expression even when it is bound by an adverb of quantification. (See Lasnik (1976) and Chierchia (1995a).) This assumption is based on contrasts such as those in (42).

- (42) a. \*He<sub>1</sub> usually<sub>1</sub> thinks that an artist<sub>1</sub> is creative.  
b. An artist<sub>1</sub> usually<sub>1</sub> thinks that he<sub>1</sub>'s creative.  
c. His<sub>1</sub> parents usually<sub>1</sub> think that an artist<sub>1</sub> is creative.  
d. The parents of an artist<sub>1</sub> usually<sub>1</sub> think that he's<sub>1</sub> creative.

Although the case of unselective binding seems a little more complicated than the cases we have looked at before, the logic is identical. Just as before, we will look at constructions such as those in (7) and test the prediction in (8). The *r*-expression in our

test will be an indefinite, and the diagnostic for Scope Reconstruction will be the binding of this indefinite by an adverb of quantification.

(7) [QP ...r-expression<sub>1</sub>...]2.....pronoun<sub>1</sub>....t<sub>2</sub>

(8) **Scope Reconstruction feeds BT(C):** Scope Reconstruction should be impossible in the structural configuration in (7).

It seems that the prediction in (8) is correct. To see this, consider first the contrast between the sentences in (43). In (43a), the principle that bans vacuous quantification forces Scope Reconstruction. Scope Reconstruction, in turn, yields a BT(C) effect, and the sentence is ruled-out. In (43b,c), by contrast, Scope Reconstruction does not yield a violation of BT(C). In (43d), the individual level predicate is replaced by a stage level predicate. For this reason, Scope Reconstruction is not forced and BT(C) is not affected.

- (43) a. \*[Which languages spoken in the country a linguist<sub>1</sub> comes from]  
          does he<sub>1</sub> usually<sub>1</sub> know t?  
      b. [Which languages spoken in the country he<sub>1</sub> comes from]  
          does a linguist<sub>1</sub> usually<sub>1</sub> know t?  
      c. [Which languages spoken in the country a linguist<sub>1</sub> comes from]  
          do his<sub>1</sub> students usually<sub>1</sub> know t?  
      d. [Which languages spoken in the country a (certain) linguist<sub>1</sub> comes from]  
          does he<sub>1</sub> usually<sub>1</sub> enjoy speaking t?

Consider next the contrast in (44). In (44a) the position of the adverb of quantification forces Scope Reconstruction to a position c-commanded by the pronoun. In (44b), by contrast, there is a position, *t'*, which satisfies the requirements of the adverb, and yet does not yield a BT(C) effect. In (44c,d), Scope Reconstruction is forced to the position of *t*, but the result does not violate BT(C). (This indicates that the problem with (44a) is related to BT(C) and is not an independent problem with Scope Reconstruction.)<sup>36</sup>

- (44) a. \*[Which languages spoken in the country a linguist<sub>1</sub> comes from]  
          did you say that he<sub>1</sub> **usually**<sub>1</sub> knows t ?  
      b. [Which languages spoken in the country a linguist<sub>1</sub> comes from]  
          do you **usually**<sub>1</sub> say t' that he<sub>1</sub> knows t ?

---

<sup>36</sup>I am ignoring various questions regarding the necessary focus structure of the constituent questions. My hope is that the answer to these question will not affect the results reported here.

- c. [Which languages spoken in the country that he<sub>1</sub> comes from]  
did you say that a linguist<sub>1</sub> **usually**<sub>1</sub> knows t ?
- d. [Which languages spoken in the country a linguist<sub>1</sub> comes from]  
did you say that his<sub>1</sub> students **usually**<sub>1</sub> know t ?

(43) and (44), thus, provide us with an additional argument that Scope Reconstruction feeds BT(C).

### 5.3. A-Reconstruction

In the previous section, we have seen that A-bar Reconstruction feeds BT(C) and thus obeys the predictions of the syntactic account. In this section, I argue that the same is true of A-reconstruction.

Let's look at cases of A-reconstruction of the sort we introduced in section 5.1. Consider the ambiguous sentences in (45), and focus on (45a). Under one of its interpretations, the sentence would be true only if there were a particular first-year student who David believed was at the party ( $\exists > seem$ ). Under the other interpretation, there doesn't need to be such a student. The sentence would be true if David is at the party and happens to hear a conversation regarding the topics discussed in the intro-class, and if this conversation prompts him to conclude that at least one first-year student had to be in the room ( $seem > \exists$ ).

- (45) a. [A first-year student] seems to David t to be at the party.
- b. [Someone from NY] is very likely t to win the lottery.

As mentioned in section 5.1, this ambiguity should be accounted for by the availability of Scope Reconstruction to the position of the trace. If Scope Reconstruction does not occur, the resulting interpretation is ( $\exists > seem$ ) interpretation. If Scope Reconstruction occurs the resulting interpretation is ( $seem > \exists$ ).

Now we can test whether Scope Reconstruction feeds BT(C) in the case of A-movement. It seems that it does. To see this consider the sentences in (46-47). It seems that the (a) sentences are disambiguated in favor of the ( $\exists > seem$ ) interpretation. To see this focus on (46a). This sentence would be true only if David had a thought about a particular student of his. It would not be true in the situation I used for illustrating the

(*seem* >  $\exists$ ) interpretation of (45a). This follows straightforwardly from the assumption that Scope Reconstruction feeds BT(C).<sup>37</sup>

- (46) a. [A student of David's<sub>1</sub>] seems to him<sub>1</sub> t to be at the party. ( $\exists$  > *seem*) \*(*seem* >  $\exists$ )  
 b. [A student of his<sub>1</sub>] seems to David<sub>1</sub> t to be at the party. ( $\exists$  > *seem*) (*seem* >  $\exists$ )
- (47) a. [Someone from David's<sub>1</sub> city] seems to him<sub>1</sub> t to be likely to win the lottery.  
 ( $\exists$  > *seem*) \*(*seem* >  $\exists$ )  
 b. [Someone from his<sub>1</sub> city] seems to David<sub>1</sub> t to be likely to win the lottery.  
 ( $\exists$  > *seem*) (*seem* >  $\exists$ )

The contrast is more evident in (48) where the semantics of the construction requires Scope Reconstruction. (See the discussion of creation verbs in 5.2.1.)

- (48) a. For these issues to be clarified,  
 [Many more/new papers about his<sub>1</sub> philosophy] seem to Quine<sub>1</sub> [t to be needed].  
 a. #For these issues to be clarified,  
 [Many more/new papers about Quine's<sub>1</sub> philosophy] seem to him<sub>1</sub> [t to be needed].

#### 5.4. Ramifications for the Interpretation of Chains

In the previous sections we have seen that Scope Reconstruction feeds BT(C). This fact, which follows straightforwardly from syntactic accounts of Scope Reconstruction, cannot

---

<sup>37</sup>(46-47) were tested up to this point with 12 speakers. Eight of the speakers got the effect and some of them thought that the judgments were strong. Four speakers got no effect. Data similar to that which I'm presenting was reported independently in Romero (1996) and Sportiche (1996). I assume that this lends support to the reality of the effect. I also got more or less the same results using Lebeaux's observation that inverse scope in constructions such as (i) and (ii) depends on Scope Reconstruction (see section 5.1 above). I got the same kind of split. A little more than half the speakers get inverse scope only in (i).

- (i) a. At least one of his<sub>1</sub> soldiers is expected by Napoleon<sub>1</sub> to die in every battle.  
 b. One of his<sub>1</sub> soldiers is expected by Napoleon<sub>1</sub> to die in every battle.  
 (ii) a. At least one of Napoleon's<sub>1</sub> soldiers is expected by him<sub>1</sub> to die in every battle.  
 b. One of Napoleon's<sub>1</sub> soldiers is expected by him<sub>1</sub> to die in every battle.

be explained by the semantic account. We thus have an argument in favor of syntactic accounts. This argument has certain consequences for semantics. In particular, it implies that the procedures that were suggested for semantic reconstruction must be restricted.

To appreciate this implication, let's look again at the chain in (2), repeated as (49). As mentioned in section 5.1.2, plausible principles for the interpretation of chains (Heim and Kratzer (forthcoming)) yield the interpretations in (49'').

- (49) [Someone from NY] is very likely [t to win the lottery]  
 (49'') a. [Someone from NY]  $\lambda x$  (is very likely [x to win the lottery])  
       b. [Someone from NY]  $\lambda Q$  (is very likely [Q to win the lottery])

(49''b) yields the semantic effects of Scope Reconstruction without actual reconstruction. If this semantic interpretation is possible, we incorrectly predict no correlation between Scope and BT(C). We, therefore, need to rule out (49''b). We need a principle from which it would follow that a trace in a theta position is interpreted as a variable which ranges over individuals (type *e*). Space limitations do not allow me to discuss various possibilities. (See Beck (1996) for a promising proposal.) Nevertheless, I think it is important to stress the obvious consequence for semantics: If type-shifting operations are allowed in the semantic component at all, they must be restricted; something must be added in order to insure that the scope of moved constituents will be reflected in the syntax.<sup>38</sup>

## 5.5. Conclusion

In this chapter, I have argued for a correlation between BT(C) and Scope Reconstruction: whenever overt movement bleeds BT(C), Scope Reconstruction is impossible. This correlation argues that BT(C) is sensitive to the position in which overtly moved constituents are interpreted. As such, it argues that BT(C) applies at LF. This, in turn, suggests that BT(C) can be used to investigate various properties of LF representations. In this chapter, I tried to investigate the nature of Scope Reconstruction. In the next chapter, I will argue that we can also use BT(C) to investigate the nature of QR. The result of the

---

<sup>38</sup>At the moment I see two possibilities both of which seem pretty natural. The first possibility is that traces like pronouns are always interpreted as variables that range over individuals (type *e*). The second possibility is that the semantic type of a trace is determined to be the lowest type compatible with the syntactic environment (as suggested in Beck 1996).



investigation will motivate certain conclusions regarding the structure and interpretation of A-bar chains. These conclusions will help us specify the nature of the mechanisms that are involved in Scope Reconstruction.

---

## Chapter 6 - Economy and Operator Variable Formation

The correlation between Scope Reconstruction and BT(C), which I attempted to establish in chapter 5, argues in favor of two conclusions. First, it argues that Scope Reconstruction should be represented structurally (i.e., that there is syntactic reconstruction). Second, it argues that Binding Theory should be sensitive to the LF position of quantificational expressions, that is, it argues that Binding Theory applies at LF. However, it is widely believed that covert QR does not affect BT(C) (Chomsky (1981)). This can be taken as an argument that BT(C) applies at SS, as well as at LF (Lebeaux (1994)).

What I would like to claim now is that Binding Theory, or at least BT(C), applies *only* at LF. My argument will have two steps familiar from Chomsky (1993). The first step -- which was actually already taken by Chomsky and in which I will basically follow his assumptions (though not the mode of the implementation) -- argues that, contrary to initial appearance, there is a coherent story to be told in which Binding Theory applies only at LF. The second step argues that the alternative, in which Binding Theory, and specifically BT(C), applies also at other levels of representation, is empirically inferior. This second step of the argument is based on Fox (1995b).<sup>1</sup>

The first step makes crucial use of the copy theory of movement and of an economy condition which governs the way in which A-bar chains are converted to interpretable objects. More specifically, it is claimed that covert movement cannot bleed BT(C) because the relevant operator variable constructions contain a copy at the tail of the chain. The second step argues that a natural formulation of the economy condition allows us to explain an observation made in Fiengo and May (1994, henceforth F&M). The observation is that when Antecedent Contained Deletion (ACD) is involved, QR does bleed BT(C). The basic idea I will advocate is that whenever Parallelism requires an operator variable construction which lacks a copy at the tail of the chain, QR has effects on BT(C).

The structure of this chapter is the following. In 6.1., I present the first step of the argument (from Chomsky 1993). In 6.2., I discuss the way in which F&M's observation can provide us with the second step of the argument. In 6.3., I argue against F&M's account of the observation, which relies on the idea that Binding Theory applies at many levels of representation. In 6.4., I characterize a generalization which is predicted from my

---

<sup>1</sup> Chomsky (1993) also presents the second step of the argument. His argument, however, is based on BT(A) and is unrelated to scope.

account and argue that the generalization holds. In 6.5.-6.6., I return to Scope Reconstruction. Specifically, I discuss the ramifications of the copy theory of movement for the theory of the syntactic mechanisms that yield Scope Reconstruction for A and A-bar chains. Finally, in 6.7., I discuss the effects of A-bar movement on BT(A), and attempt to provide additional evidence in favor of Scope-Economy.

### 6.1. The First Step (Chomsky 1993)

Let's begin by reviewing the evidence that was taken to show that BT(C) applies at SS. Consider the contrast between (1) and (2). Under certain assumptions about the nature of covert QR (Chomsky (1977), May (1977; 1985)), the LF structures of the sentences in (1) are those in (1'). These structures are identical with respect to BT(C) to the surface structures in (2). If BT(C) applied only at LF, there would be no obvious way of accounting for the contrast. If BT(C) applied also at SS, the contrast would follow straightforwardly.<sup>2</sup>

- (1)    a. \*/?? you bought him<sub>i</sub> every picture that John<sub>i</sub> liked.  
          b. \* He<sub>i</sub> bought you every picture that John<sub>i</sub> liked.
- (2)    a. [[which picture that John<sub>i</sub> liked] [did you buy him<sub>i</sub> t]]?  
          b. [[which picture that John<sub>i</sub> liked] [did he<sub>i</sub> buy you t]]?
- (1')   a. [[every picture that John<sub>i</sub> liked] [I bought him<sub>i</sub> t]]  
          b. [[every picture that John<sub>i</sub> liked] [he<sub>i</sub> bought you t]]

**6.1.1. Chomsky's proposal:** Chomsky (1993), however, provides a way of accounting for the contrast without the assumption that BT(C) applies at SS. In particular,

---

<sup>2</sup> Given Scope-Economy, the structures in (1) involve very short QR. Therefore, it is not obvious that they pose a problem for the assumption that BT(C) applies only at LF. However, it turns out that the argument based on (1) carries over to structures for which this objection does not hold:

- (i)    a. \*/?? A different girl bought him<sub>i</sub> every picture that John<sub>i</sub> liked.  
          b. \* A different<sub>i</sub> girl wanted him<sub>i</sub> to buy every picture that John<sub>i</sub> liked.
- (ii)   a. A different girl bought John<sub>i</sub> every picture that he<sub>i</sub> liked.  
          b. A different girl wanted John<sub>i</sub> to buy every picture that he<sub>i</sub> liked.

he suggests that A-bar movement always leaves a copy and that this copy (under certain circumstances) yields a BT(C) effect, even if BT(C) applies only to the output of movement. In (1), the true output of QR is quite different from (1'). Specifically, it still has a copy of the moved constituent at the position of the trace and it is the r-expression within this copy that yields the violation of condition C. In (2) -- Chomsky claims (following Lebeaux (1988)) -- A-bar movement applies prior to the insertion of the relative clause that contains the r-expression. Therefore, in (2), the copy of the moved constituent does not yield a BT(C) effect. The difference between overt and covert movement under this proposal is not related to their respective ordering relative to Binding Theory but rather to their respective ordering relative to lexical insertion.<sup>3</sup> Covert movement is never followed by lexical insertion, and therefore never appears to get around a BT(C) violation.

As it turns out, certain cases of overt movement are similar to the cases of covert movement in that they are unable to get around a BT(C) violation. These cases are demonstrated by Lebeaux's (1988) contrast that was mentioned in the previous chapter and is repeated in (3). (See also van Riemsdijk and Williams (1981) and Freidin (1986).) Chomsky accounts for this contrast on the basis of a distinction between the timing of adjunct and complement insertion, which he also borrows from Lebeaux. According to this distinction the insertion of complements, in contrast to adjuncts, must take place prior to movement (in accordance with the extension/projection principle). From this it follows that complements, such as the boldfaced phrases in (3b), in contrast to adjuncts, such as the relative clauses in (3a), cannot bypass BT(C) via overt A-bar movement.<sup>4</sup>

- (3) a. [Which argument **that John made**] did he believe t?  
 b. ??/\*[Which argument **that John is a genius**] did he believe t?

<sup>3</sup> \*Note that the claim that overt and covert movement differ in their ordering relative to lexical insertion is strongly motivated on independent grounds. There is strong independent motivation for the claim that lexical insertion cannot follow covert operations (at least not without severe constraints). If this claim were false, it is hard to imagine how we would account for any correspondence between meaning and sound.

<sup>4</sup> The reader might wonder whether the possibility of inserting adjuncts at various points in the derivation is consistent with the observation in section 5.2. that A-bar Scope Reconstruction feeds BT(C). In 6.5, I will show that it is consistent. The basic idea is that Scope Reconstruction is the result of interpreting a large part of the copy at the base position. I will show that such an interpretation is available only if the adjunct is inserted at the base position; the option of late insertion necessarily yields the non-reconstructed interpretation.

It turns out that overt A-bar movement of certain phrases (phrases that contain complements and no adjuncts) is identical to covert movement with respect to Condition C. This weakens the argument from (1-2) that Condition C makes an overt/covert distinction. Nevertheless, this provides us only with the first stage of the argument that BT(C) applies only at LF. It is still *possible* to account for all the data under the assumption that BT(C) applies both at SS and at LF (See Lebeaux (1988, 1994)). I would now like to present the argument from Fox (1995b) that BT(C) must apply only at LF. The argument is based on an observation of Fiengo and May (1994) that certain cases of covert movement do in fact obviate BT(C). This observation cannot be accounted for under the assumption that BT(C) applies at Surface Structure.

**6.1.2. The interpretation of A-bar chains (a slight modification):** The discussion thus far hasn't spelled out the nature of the structures that get interpreted. As noted by Chomsky (1993), interpreting an operator variable construction probably requires some alterations of the copies created by movement. In particular, Chomsky suggests that the output of movement in structures such as (4), which is fully represented in (4'), undergoes a later process which forms one of the structures in (4'').

(4) Which book did Mary read t?

(4') Which book did Mary read which book?

(4'') a. Which book<sub>x</sub> did Mary read x?  
b. Which <sub>x</sub> did Mary read book x?<sup>5</sup>

Further he assumes that the structure in (4''b) is preferred to the structure in (4''a), thus accounting for the BT(C) violation in (3b). I will basically follow this assumption, but will make a small modification in the implementation. This modification will make the interpretation of the QRed structures more straightforward and will perhaps allow the

---

<sup>5</sup> It is conceivable that these structures should be interpreted via quantification over choice functions (Reinhart (1995), Kratzer (1995b), Winter (1995) and Engdahl (1980: 131-141)). However, as pointed out to me by Irene Heim, Uli Sauerland and Yoad Winter, it is not clear how such an analysis would extend to proportional quantifiers (e.g. *most*, *almost every*, etc.) This is one of the motivations for the modification that follows. See, however, Sauerland (forthcoming).

assumption to follow from general principles of Economy. Under the modification, the two structures are those in (4''').

- (4''') a. Which book<sub>x</sub> did Mary read x?  
b. Which book<sub>x</sub> did Mary read book x?

(4'''a) is interpreted standardly. For (4'''b) something novel needs to be proposed. There are various possibilities that come to mind. For concreteness, I will make a suggestion inspired by work by Rullmann and Beck (1997), Sauerland (forthcoming) and Wold (forthcoming). According to this suggestion *book x* is interpreted as a definite description *the book identical to x* yielding an interpretation paraphrasable as *which is the book, x, such that Mary read the book identical to x*.<sup>6</sup> The specifics of this proposal are not crucial for the purposes of this chapter. What is crucial is that there is a not implausible semantic method for interpreting (4'''b). Assuming this method is secure, we can state Chomsky's economy condition in the following manner:

- (5) **OV-Economy:** Given an A-bar chain,  $\alpha$ , choose the operator variable construction that is closest to  $\alpha$  given the set of interpretable options.<sup>7</sup>

An operator variable construction  $O_1$  is closer to a chain  $\alpha$  than  $O_2$  if the set of positions at the tail of  $\alpha$  that are maintained in  $O_2$  is a proper sub-set of the parallel set in  $O_1$ .

For the A-bar chain in (4'), there are two interpretable operator variable constructions (the one in (4'''a) and the one in (4'''b)). Assuming that both are interpretable, OV-Economy prefer (4'''b) to (4'''a).<sup>8</sup>

- (4''') a. \*Which book<sub>x</sub> did Mary read x? (ruled out by OV-Economy)  
b. Which book<sub>x</sub> did Mary read book x?

<sup>6</sup> Rullmann and Beck consider this possibility as a method for interpreting *wh*-in situ. Sauerland, suggests a somewhat different semantic approach which relies on the notion of a minimal choice function.

<sup>7</sup> It is an open question whether OV-Economy is sensitive to semantic interpretation (in a way similar to Scope-Economy and VB-Economy). If it is, we could replace 'the set of interpretable options' with 'the set of options that yield a given interpretation M' (with the relevant qualifications discussed in 2.5).

<sup>8</sup> For similar, though not identical, ideas see Cresti (1996).

For QR, similar issues arise. A sentence such as (6) has (6') as the output of QR, which can in turn be converted to one of the structures in (6'').<sup>9</sup> OV-Economy determines that the interpreted structure is (6''b).<sup>10</sup>

- (6) John<sub>i</sub> [<sub>VP</sub> t<sub>i</sub> likes every boy].  
 (6') John<sub>i</sub> [every boy [<sub>VP</sub> t<sub>i</sub> likes every boy]].  
 (6'') a. \*John<sub>i</sub> [every boy<sub>x</sub> [<sub>VP</sub> t<sub>i</sub> likes x]] (*ruled out by OV-Economy*)  
       b. John<sub>i</sub> [every boy<sub>x</sub> [<sub>VP</sub> t<sub>i</sub> likes boy x]]

The explanation of the inability of QR to by-pass BT(C) is the same as the explanation of the inability in the case of overt *Wh*-movement. The explanation is based on an economy principle that prefers structures in which the restrictor of the quantifier is not eliminated from the base position.

## 6.2. The Second Step (Fox 1995b)

With this much in hand, we are in a position to present the argument that BT(C) must apply only at LF. The logic of the argument is based on the nature of economy principles. These principles choose an object from a set of competitors (a reference set). If under certain circumstances, the reference set is restricted so as not to include the most optimal object, it is predicted that an otherwise unacceptable object will be licensed. In our case, we predict that (6''a) will be licensed under circumstances in which (6''b) is not a member of the reference set. Under such circumstances, QR should obviate BT(C) effects. The question is whether such circumstances exist.

In Fox (1995b) I suggest that they do. In particular, I suggest that in cases involving Antecedent Contained Deletion (ACD), the parallel of (6''b) is not licensed and the parallel of (6''a) is the only element in the reference set and hence is acceptable. As is

<sup>9</sup> I assume, based on the discussion in chapter 2, that in sentences such as (6') QR is limited to the VP level.

<sup>10</sup> Note that OV-Economy is very similar to the economy condition which prefers the least amount of pied-piping needed for interpretation (Minimize PP, Chomsky 1995). On the basis of this similarity, I suggested in Fox (1995b) a restatement of the ideas reported here in terms of Minimize PP. The reason I avoid this formulation here is because I feel more comfortable with operator variable constructions in which the whole restrictor is at the head of the chain. However, see Wold (forthcoming) where a semantic procedure is proposed for interpreting the output of QR under the assumption that only the determiner is raised.





- b. ??/\* You introduced him<sub>i</sub> to everyone that John's<sub>i</sub> mother met.
  - c. ??/\* You reported him<sub>i</sub> to every cop that John<sub>i</sub> was afraid of
- (9)
- a. You sent him<sub>i</sub> the letter that John<sub>i</sub> expected you would.
  - b. You introduced him<sub>i</sub> to everyone that John's<sub>i</sub> mother did.
  - c. You reported him<sub>i</sub> to every cop that John<sub>i</sub> was afraid you would.
- (8')
- a. You [the letter that John<sub>i</sub> expected you would write]<sub>x</sub>  
[sent him<sub>i</sub> x letter that John<sub>i</sub> expected you would write].
  - b. You [everyone that John's<sub>i</sub> mother met]<sub>x</sub>  
[introduced him<sub>i</sub> to x one that John's<sub>i</sub> mother met].
  - a. You [every cop that John<sub>i</sub> was afraid of]<sub>x</sub>  
[reported him<sub>i</sub> to x cop that John<sub>i</sub> was afraid of]
- (9')
- a. You [the letter that John<sub>i</sub> expected you would <send him x>]<sub>x</sub>  
[sent him<sub>i</sub> x].
  - b. You [everyone that John's<sub>i</sub> mother did <introduce him to x>]<sub>x</sub>  
[introduced him<sub>i</sub> to x].
  - c. You [everyone that John<sub>i</sub> was afraid you would <report him to x>]<sub>x</sub>  
[reported him<sub>i</sub> to x].

OV-Economy makes additional predictions. However, before I introduce these predictions, I would like to go over F&M's account of the contrast between (8) and (9).

### 6.3. F&M's Account

The fact that QR can obviate BT(C) seems to me to suggest that this condition applies only at LF. However, F&M (who, as mentioned, discovered the fact) draw a different conclusion. Specifically, they suggest that in the cases in which BT(C) is obviated by QR (cases such as (9)), the condition applies only at LF. However, in the cases in which BT(C) is not obviated by QR (cases such as (8)), they maintain the claim that the condition applies also at SS. More specifically, they postulate an algorithm which determines at what levels of representation BT(C) applies. The algorithm is sensitive to whether or not there is ACD, and thus captures the contrast between (8) and (9). Since I want to take the ACD facts as evidence in favor of the conclusion that BT(C) applies only at LF, I need to address the alternative.

F&M suggest that BT(C) applies at every level of representation in which it can apply. In other words, they assume that a structure is ruled out if it violates BT(C) at any

level of representation (at which BT(C) can apply). Further, they suggest that BT(C) applies to indices, and that it can apply to an index, *i*, only when all occurrences of *i* are “projected” in the structure. In addition, they assume a theory of ellipsis from which it follows that elided material is “projected” only when Parallelism is met. For the sake of this discussion, we can say that they assume an LF copying theory of ellipsis. That is to say, they assume that interpretable material which is not present at PF is also not present at syntactic levels which are the input to PF. Such material is not deleted at PF, but rather added (copied/reconstructed/projected) in the covert component.<sup>13</sup> With these assumptions, they propose to explain the contrast between (8) and (9).

BT(C) fails to apply to a certain index under one condition: when an occurrence of index is absent from the structure. Therefore, QR could affect the way BT(C) applies to a given index, only if there is an occurrence of the index such that QR is a prerequisite for its presence in the structure. Only in such a case will BT have to wait for QR in order to apply. Given F&M’s assumption about ellipsis (the assumption that the “elided” VP enters the structure only in the covert component), QR is a prerequisite for an index to enter a structure when the index is contained in an “elided” VP and when QR is necessary for the LF copying of this VP (i.e., when there is ACD). This accounts for the contrast between (8) and (9). To see this, compare (8c) and (9c):

(8c) ??/\* You reported him<sub>i</sub> to every cop that John<sub>i</sub> was afraid of

(9c) You reported him<sub>i</sub> to every cop that John<sub>i</sub> was afraid you would.

(8c) doesn’t involve ACD, and should be ruled-out by BT(C). This result is predicted by F&M in the following way. Because (8c) doesn’t involve ellipsis, all of the LF material is already present at SS. Consequently, BT(C) can, and therefore must, apply at SS. At SS, (8c) is ruled out, and QR cannot reverse the verdict. The obviation of BT(C) when ACD is involved is predicted by F&M based on the idea that the missing structure, the elided VP, contains an occurrence of the index, *i*. Therefore, BT(C) can apply only after LF-copying. LF copying, in turn, can take place only after QR. By transitivity, QR must apply before BT(C) and, consequently, can obviate this condition.

---

<sup>13</sup> It is important to point out that this assumption is crucial only for F&M’s treatment of ACD. It is not important for their other claims about ellipsis.

F&M's account relies heavily on the notion of "an index occurrence". As such, we would like to have evidence that this notion is relevant. The evidence that F&M provide is based on the contrast in (10-11).

- (10) a. You bought him<sub>i</sub> [every picture that John<sub>i</sub> wanted you to <buy him<sub>i</sub>>].  
 b. \*He<sub>i</sub> bought you [every picture that John<sub>i</sub> wanted to <buy you>].
- (11) a. You talked to him<sub>i</sub> about [every topic that John's<sub>i</sub> teacher did <talked to him<sub>i</sub> about>].  
 b. \*He<sub>i</sub> talked to you about [every topic that John's<sub>i</sub> colleague did <talked to you about>].

F&M assume that QR can always target IP. Under this assumption, the contrast between the (a) and (b) sentences is somewhat puzzling. F&M notice, however, that it follows from the notion of "index occurrence" together with the assumption that BT(C) applies to an index only once all of the occurrences of the index have been reconstructed from the antecedent VP. In the (b) sentences, in contrast to the (a) sentences, the elided VP does not contain an occurrence of the index, *i*. Therefore, BT(C) can, hence must, apply before the missing structure is filled. BT(C) applies at SS, and the (b) sentences are ruled out before QR has its say.

However, consider the sentences in (12). In these sentences there is an occurrence of the index, *i*, within the elided VP. Therefore, F&M incorrectly predict that BT(C) would be obviated by QR.

- (12) a. \*He<sub>i</sub> bought his<sub>i</sub> brother [<sub>QP</sub> everything John<sub>i</sub> wanted to <buy his<sub>i</sub> brother t>].  
 b. \*He<sub>i</sub> introduced his<sub>i</sub> brother to [<sub>QP</sub> everyone John<sub>i</sub> wanted to <introduce his<sub>i</sub> brother to t>].

The unacceptability of the sentences in (12) eliminates F&M's argument for the notion of an index occurrence.<sup>14</sup>

<sup>14</sup> An *LJ* reviewer has pointed out that, although (12) eliminates F&M's argument for index occurrences, it does not rule out their proposal entirely. Given the arguments in favor of Scope-Economy (c.f., chapter 2), it is implausible to conclude that QR targets IP in (10-12). In these cases, optional QR over the subject does not reverse the relative scope of two non-commutative quantifiers. For this reason, optional QR is not licensed and the sentences are confined to the representations that would result from obligatory QR to the VP level. In other words, even if in (10b), (11b) and (12), BT applies after QR, the

A more direct argument against F&M's proposal involves constructions with phonological reduction. These constructions behave just like ellipsis constructions with respect to BT(C):

- (13) a. You bought him<sub>i</sub> [every picture that John<sub>i</sub> wanted you to <buy him<sub>i</sub>>].  
 b. You bought him<sub>i</sub> [every picture that John<sub>i</sub> wanted you to *buy him<sub>i</sub>*].
- (14) a. You talked to him<sub>i</sub> about [every topic that John's<sub>i</sub> teacher did <talked to him<sub>i</sub> about>].  
 b. You talked to him<sub>i</sub> about [every topic that John's<sub>i</sub> teacher *talked to him<sub>i</sub> about*].

This result is unexpected under F&M's proposal.<sup>15</sup> (All occurrences of the index *i* are present at SS in the case of down-stressing.) However, it follows straightforwardly under

---

sentences should be ruled out since QR does not remove the r-expression from the c-command domain of the coindexed pronoun.

The case in (i) would be more problematic for F&M. In order to account for unacceptability of these sentences, F&M would have to claim that QR cannot target the matrix sentence. However, it can, as indicated by the option for matrix scope in (ii).

- (i) a. \*At least one boy expected him<sub>i</sub> to buy his<sub>i</sub> brother every book that John's<sub>i</sub> mother did <buy his<sub>i</sub> brother>.  
 b. \*I wanted him<sub>i</sub> to buy his<sub>i</sub> brother a book that John's<sub>i</sub> mother did <buy his<sub>i</sub> brother>.
- (ii) a. At least one boy expected John<sub>i</sub> to buy his<sub>i</sub> brother every book that his<sub>i</sub> mother did <buy his<sub>i</sub> brother>. ( $\exists > \forall$ ) ( $\forall > \exists$ )  
 b. I wanted John<sub>i</sub> to buy his<sub>i</sub> brother a book that his<sub>i</sub> mother did <buy his<sub>i</sub> brother>. ( $\exists > \text{want}$ ) ( $\text{want} > \exists$ )

My analysis of (i) will become clear in the next section.

<sup>15</sup> Another direct argument against F&M's proposal is based on an observation made independently in Merchant (1998) and Sauerland (forthcoming) that BT(C) is possible only if the r-expression is contained in the relative clause that dominates the (antecedent contained) ellipsis site:

- (i) \*I told him<sub>i</sub> [every story about John's<sub>i</sub> brother [that Mary did]].  
 (ii) I told him<sub>i</sub> [every story about Mary [that John's<sub>i</sub> brother did]].

Both authors have pointed out that the contrast would follow from OV-Economy if we assume a head internal analysis for the relative clause. Assuming a head internal analysis, Parallelism would be satisfied in an LF representation in which the material external to the relative clause is not deleted from the copy of the QRed constituent (since this material has a copy internal to the relative clause). OV-Economy, thus, determines the following LFs for (i) and (ii) (ignoring, for simplicity, the requirements of Scope-Economy):

the suggestion I made in the previous section, since (as discussed in chapter 3) constructions involving phonological deletion and reduction are subject to the same Parallelism constraints.<sup>16</sup>

## 6.4. A Predicted Generalization

Now, I would like to return to the proposal that I made in 6.2. According to this proposal, A-bar chains are converted to operator variable constructions in a manner determined by OV-Economy. Regular A-bar chains will always keep a copy of the moved constituent at the tail of the chain, and thus will never affect BT(C). However, chains that take an elided constituent outside of its antecedent (i.e., chains that are necessary for ACD resolution) are interpretable only if there is no copy at the tail of the chain. The copy is, therefore, removed, and BT(C) is obviated.

Let us consider now successive cyclic A-bar movement. Such movement creates multiple chains each of which needs to be converted to an operator variable construction. Suppose that one step of successive movement is necessary for ACD resolution, and another isn't:

- (15)  $[_{QP..r-exp.}] \dots [_{QP..r-exp.}] \cdot [_{Antecedent VP..[_{QP..r-exp.}]}]$

- (i') \**[every story about John's<sub>i</sub> father]<sub>x</sub> [that Mary told him, [x story about John's<sub>i</sub> brother]]*  
*[I told him, [x story about John's<sub>i</sub> brother]].*
- (ii') \**[every story about Mary]<sub>x</sub> [that John's<sub>i</sub> brother told him, [x story about Mary]]*  
*[I told him, [x story about Mary]].*

<sup>16</sup> Kennedy (1997) discusses the fact that there is no detectable contrast between the examples in (ii).

- (ii) a. Polly introduced him<sub>i</sub> to everyone Erik wanted her to.  
b. Polly introduced him<sub>i</sub> to everyone Erik<sub>j</sub> wanted to meet.

In Fox (1995b: 116-118), I provide an account of this fact based on the theory of Parallelism for phonological reduction. The basic idea is that (iib), in contrast to the sentence in (8), involves a case of what I call Antecedent Contained Deaccenting. The embedded VP is optionally non-F-dominated (and this is licensed given accommodation, see chapter 3). When it is non-F-dominated, Parallelism forces the otherwise uneconomical deletion at the tail of the chain. See Wald (forthcoming) for a formal presentation of this proposal.



buy [QP everything that he<sub>i</sub> thought I did <expected him<sub>i</sub> to buy t>].  
*(embedded scope; matrix ellipsis)*

- c. I  
 [QP everything that he<sub>i</sub> thought I did <buy t>]  
 expected John<sub>i</sub> to  
 [QP everything that he<sub>i</sub> thought I did <buy t>]  
 buy [QP everything that he<sub>i</sub> thought I did <buy t>].  
*(matrix scope; embedded ellipsis)*

- d. I  
 [QP everything that he<sub>i</sub> thought I did <expected him<sub>i</sub> to buy t>]  
 expected John<sub>i</sub> to  
 [QP everything that he<sub>i</sub> thought I did <expected him<sub>i</sub> to buy t>]  
 buy [QP everything that he<sub>i</sub> thought I did <expected him<sub>i</sub> to buy t>].  
*(matrix scope; matrix ellipsis)*

However, as pointed out in Larson and May (1990) and Sag (1976:73), (18'b) has no way of achieving Parallelism. We are, thus, left with (a), (c) and (d). Each of these must be converted into an operator variable construction under OV-Economy. This Economy principle chooses the most optimal operator variable construction that obeys Parallelism. In other words, each chain contains a copy at the tail of the chain unless the chains resolves antecedent containment. We thus end up with the three structures in (18'').

- (18'') a. I expected John<sub>i</sub> to  
 [QP everything that he<sub>i</sub> thought I did <buy t>]  
 buy t.  
*(embedded scope; embedded ellipsis)*
- c. I  
 [QP everything that he<sub>i</sub> thought I did <buy t>]<sub>x</sub>  
 expected John<sub>i</sub> to  
 [QP x thing that he<sub>i</sub> thought I did <buy t>]  
 buy t.  
*(matrix scope; embedded ellipsis)*
- d. I  
 [QP everything that he<sub>i</sub> thought I did <expected him<sub>i</sub> to buy t>]  
 expected John<sub>i</sub> to buy t.  
*(matrix scope; matrix ellipsis)*

In (a) there is a single instance of QR, hence a single chain; in (d) and (c) there are two chains. In all three constructions Parallelism forces a simple trace at the theta position. (d)

and (c) differ in that Parallelism requires the elimination of the intermediate copy in (d) but not in (c).

The predictions for BT(C) are obvious. In (d), both instances of QR are necessary for ACD resolution, and hence both should bleed BT(C). In (a) and (c) only one instance of QR resolves ACD and hence only one instance of QR should bleed BT(C).

**6.4.2. Testing the predictions:** Consider now the sentence in (19). This sentence is different from (18) in that it allows only matrix VP ellipsis. This is exactly what is predicted; the structures which involve embedded VP deletion (19'a and 19'c) violate BT(C), while the structure in (19'd) does not.

(19) I expected him<sub>i</sub> to buy everything that John<sub>i</sub> thought I did.

- a. \*<bought t>
- b. <expected him<sub>i</sub> to buy t>

(19') a. \*I expected him<sub>i</sub> to

[QP everything that John<sub>i</sub> thought I did <buy t>]

buy t.

(*embedded scope; embedded ellipsis*)

c. \*I

[QP everything that John<sub>i</sub> thought I did <buy t>]<sub>x</sub>

expected him<sub>i</sub> to

[QP x thing that John<sub>i</sub> thought I did <buy t>]

buy t.

(*matrix scope; embedded ellipsis*)

d. I

[QP everything that John<sub>i</sub> thought I did <expected him<sub>i</sub> to buy t>]

expected him<sub>i</sub> to buy t.

(*matrix scope; matrix ellipsis*)

The proposal predicts that QR would bleed BT(C) only if QR is long enough to get out of the c-command domain of the "dangerous" pronoun, and only if QR is needed for ACD resolution and thus requires elimination of the offending material at the tail of the chain.

That this prediction is correct can be demonstrated with additional examples. Consider the contrast between the sentences in (20). (20a) is ambiguous. (20b) requires matrix VP deletion. Once again, this is predicted; only matrix VP deletion forces the less economical structure which avoids a BT(C) violation.



- (20) a. In the end, I demanded that John<sub>i</sub> read exactly those books that he<sub>i</sub> suspected I would. a. <read t> b. <demand that he read t>  
 b. In the end, I demanded that he<sub>i</sub> read exactly those books that John<sub>i</sub> suspected I would. a. \*<read t> b. <demand that he read t>

Consider now the contrast between (21) and (22). In (21), embedded VP deletion is preferred to matrix VP deletion (for some speakers, the latter is impossible). In (22), where embedded VP deletion would bring about a BT(C) effect, the judgments are reversed.<sup>17</sup>

- (21) I sa'id that Bill<sub>i</sub> bought everything he<sub>i</sub> thought I did.  
 a. <bought t>  
 b. ? <said that he bought t>
- (22) I said that he<sub>i</sub> bought everything Bill<sub>i</sub> thought I did.  
 a. \*<bought t>  
 b. ? <said that he bought t>

In sections 5.2-5.3, we have seen that Scope Reconstruction feeds BT(C). This forces the conclusion that Binding Theory applies (also) at LF. In this section we have seen that (a) it is possible to maintain that BT(C) applies only at LF albeit what appear to be evidence to the contrary (Chomsky (1993)) and (b) this stance is virtually necessary on empirical grounds (Fox (1995b)).

## 6.5. Scope Reconstruction in A-bar Chains

In the previous section, we have seen that under normal circumstances A-bar movement fails to affect BT(C) irrespective of whether or not the moved constituent is reconstructed. This means that under normal circumstances, we do not expect A-bar movement to show an interesting correlation between Scope Reconstruction and BT(C). This might raise a question regarding the status of the correlation we have seen in section 5.2. The answer to

---

<sup>17</sup> The (a) readings require focal stress on the pronoun *I*. This should follow from independent principles (See Rooth (1992) and Tancredi (1992). See also chapter 3.). As David Pesetsky (personal communication) points out, it might be more accurate to state our correlation as a correlation between focal stress and BT(C), which is explained by the proposal I make together with an independently motivated correlation between the size of ellipsis and the site of focal stress.

this question is simple. Normal cases of A-bar movement do not affect BT(C) irrespective of reconstruction. What we've seen in section 5.2 is that when there is reconstruction the "abnormal" cases of A-bar movement behave like the normal cases. In other words, we have seen that the method which allows A-bar movement to affect BT(C) does not allow for reconstruction. Let's see how this result follows from the view of syntactic reconstruction that the copy theory of movement provides.

Under the copy theory of movement, A-bar movement can affect BT(C) only if the r-expression is within an adjunct and (modulo ACD) only if the adjunct is inserted after movement. This is illustrated schematically in (23-24).

- (23) \* $[QP \text{ .} [\text{complement} \text{ .} r\text{-expression}_1 \text{ .} ] \text{ .} ]_2$   
       ..... $\text{pronoun}_1$ ..... $[QP \text{ .} [\text{complement} \text{ .} r\text{-expression}_1 \text{ .} ] \text{ .} ]_2$
- (24) a. \* $[QP \text{ .} [\text{adjunct} \text{ .} r\text{-expression}_1 \text{ .} ] \text{ .} ]_2$       (*adjunct inserted before movement*)  
       ..... $\text{pronoun}_1$ ..... $[QP \text{ .} [\text{adjunct} \text{ .} r\text{-expression}_1 \text{ .} ] \text{ .} ]_2$   
       b.  $[QP \text{ .} [\text{adjunct} \text{ .} r\text{-expression}_1 \text{ .} ] \text{ .} ]_2$       (*adjunct inserted after movement*)  
       ..... $\text{pronoun}_1$ ..... $[QP \text{ .} ]_2$

Reconstruction, on the other hand, is achieved via the deletion of the head of the chain and the interpretation of the tail alone, as in (25)

- (25)  $QP_2$  ..... $\text{pronoun}_1$ ..... $QP_2$  ---reconstruction---->  
       ..... $\text{pronoun}_1$ ..... $QP_2$

If an adjunct is inserted after movement, reconstruction is blocked since it will not allow the adjunct to get an interpretation.<sup>18</sup> It thus follows that A-bar movement cannot affect BT(C) when the moved constituent is reconstructed.

This explanation is based on the idea that the moved constituent is fully reconstructed. However, the cases of *Wh*-movement we've looked at in section 5.2 involve forms of partial reconstruction; in these cases the *Wh*-operator is interpreted in the surface position and only a part of it is reconstructed to the base position. What I would like to do, therefore, is show that all the cases we've looked at nevertheless involve

<sup>18</sup> There are many ways to capture the idea that unrecoverable deletion of the adjunct is blocked. One possibility, among many others, is that an element can be deleted only under identity with a copy. Late insertion of an adjunct makes the head of the chain nonidentical to the tail.

reconstruction of a constituent that includes the adjunct and thus are captured by the explanation given above. Furthermore, I would like to show a case of partial reconstruction which does not necessarily include the adjunct. In this case, the prediction of the copy theory is that a violation of BT(C) would occur iff the adjunct that contains the r-expression is reconstructed.<sup>19</sup>

**6.5.1. The Cases Discussed in chapter 5:** In section 5.2 we've looked at two basic cases of reconstruction. Let's begin with the simple case discussed in section 5.2.2-5.2.3. In this case a relative clause contains a variable which must be bound in the reconstructed position. It was shown that if the relative clause contains an r-expression as well and if the r-expression is c-commanded by a co-indexed element in the reconstructed position, a BT(C) effect emerges:<sup>20</sup>

- (26) \*[Which book that he<sub>i</sub> asked Ms. Brown<sub>j</sub> for]  
did she<sub>j</sub> give every student<sub>i</sub> ?

This case follows straightforwardly from the copy theory of movement. If the adjunct is inserted after movement, there is no way for the variable to be bound. For concreteness, let's assume, following Engdahl (1980), that (26) has the following LF:

- (26') Which (choice function) f  
did she<sub>j</sub> give every student<sub>i</sub> f (book that he<sub>i</sub> asked Ms. Brown<sub>j</sub> for)?

This LF has the adjunct in the base position and thus cannot result from its late insertion.

Now let's move to the slightly more complex case which was discussed in section 5.2.1. In this case a *how many* question is separated with *how* interpreted in the surface position and *many NP* reconstructed to the base position:

- (27) How many ideas is John likely to come up with?  
How n: John is likely to come up with n many ideas?  
What is the number n s.t. John is likely to come up with n many ideas?

<sup>19</sup> This prediction would follow under any syntactic account of partial reconstruction.

<sup>20</sup> Similar considerations apply to the cases of unselective binding in which the relative clause contains a variable bound in the base position by an adverb of quantification (section 5.2.3).



- (29) [How many more ideas than what's needed for his<sub>i</sub> tenure]  
is John<sub>i</sub> planning to come up with?

This question is similar to (27) in that the semantics of the embedded creation verb forces reconstruction of the DP *N many ideas*. However, in contrast to (27), (29) is ambiguous with respect to the scopal position of the comparative quantifier. This ambiguity is represented by the two LFs in (29').

- (29') a. how n:  $(\exists N)$  N is n-more than  $(\iota M)$ [M many ideas are needed for his<sub>i</sub> tenure]  
[John<sub>i</sub> is planning to come up with N many ideas]  
(answer given S1: 60)  
b. how n: John<sub>i</sub> is planning  
 $(\exists N)$  N is n-more than  $(\iota M)$ [M many ideas are needed for his<sub>i</sub> tenure]  
[PRO<sub>i</sub> to come up with N many ideas]  
(answer given S1: 10)

To see this ambiguity consider the following situation:

- (S1) John thinks that he needs 100 ideas for tenure. He wants to come up with 110 ideas to be on the safe side (that is to say, he wants to have 10 more ideas than what's needed). However, the truth is that he needs only 50 ideas for tenure.

Under (S1) there are two possible answers to (29). One answer is 60 and the other is 10. The two answers correspond to the two LFs in (29'). If the comparative takes wide scope relative to the intensional verb *plan*, the value of the definite description  $(\iota M)$ [M many ideas are needed for his<sub>i</sub> tenure] is determined in the actual world to be 50 and the answer to the question is 60. If, on the other hand, the comparative takes narrow scope relative to the intensional verb, the value of the definite description is determined in the belief worlds to be 100 and the answer to the question is 10.<sup>22</sup>

<sup>22</sup> As pointed out to me by Irene Heim, this is not quite accurate. There are many arguments that the definite description can be in the scope of the intensional verb and yet have its value determined in the actual world. More specifically, there are arguments that certain DPs contain world variables that (like pronouns) can be bound non-locally.

However, this does not affect the argument. The argument relies on the assumption that there is no way for a definite description to be interpreted outside the scope of the intensional verb and (at the same time) to have its value determined in the worlds that the intensional verb quantifies over. I doubt that anyone would challenge this assumption.

(29) is different from (27) in that in (29) there is an adjunct *what's needed for his tenure* which is not contained within the DP *many ideas*. Therefore in (29) it is possible to reconstruct the DP without reconstructing the adjunct. We thus predict that (29) can be transformed into (30) without yielding a BT(C) effect. This prediction is born out. Furthermore, it is pretty clear that (30) has only one answer --"60"-- given the situation described in S5. In other words, it seems that (30) has only the LF in (30'a), in which the adjunct is inserted after movement.<sup>23</sup>

(30) [How many more ideas than what's needed for John's<sub>i</sub> tenure]  
is he<sub>i</sub> planning to come up with?

(30') a. how n: ( $\exists N$ ) N is n-more than (tM)[M many ideas are needed for John's<sub>i</sub> tenure]  
[he<sub>i</sub> is planning to come with N many ideas]  
(answer given S5: 60)

b. \*how n: he<sub>i</sub> is planning  
( $\exists N$ ) N is n-more than (t M)[M many ideas are needed for John's<sub>i</sub> tenure]  
[PRO<sub>i</sub> to come with N many ideas]  
(answer given S5: 10)

What we've seen in this section is that in cases of partial reconstruction late insertion of an adjunct is possible if and only if the adjunct is not contained in the reconstructed material. This is exactly what's expected under the copy theory of movement. Furthermore, this result strengthens the idea that reconstruction is syntactic. It shows us that when there is partial reconstruction, we find syntactic effects for exactly those elements that are reconstructed.

## 6.6. A Note on the A/A-bar Distinction

The conclusion that Scope Reconstruction feeds BT(C) (sections 5.2-5.3) is true for all types of movement. However, the conclusion that movement bleeds BT(C) only when there is reason to get rid of the restrictor (section 6.1-6.4) is true only with respect to A-bar movement. A-movement bleeds BT(C) with no special proviso. This well-known contrast is illustrated in (31) and (32).

<sup>23</sup> This analysis would apply also to similar examples which were problematic for Heycock (Heycock 1995: note 19). Thanks to Chris Kennedy for bringing these examples to my attention.

- (31) Standard A-bar movement fails to bleed BT(C)  
 a. ??/\*Which argument that John<sub>i</sub> is a genius did he<sub>i</sub> believe t?  
 b. \*A different person told him<sub>i</sub> about every argument that John<sub>i</sub> is a genius.
- (32) Standard A-movement bleeds BT(C)  
 Every argument that John<sub>i</sub> is a genius seems to him<sub>i</sub> to be flawless.

I don't fully understand this contrast. Nevertheless, I will state it explicitly in the following manner:

- (33) a. A-movement (optionally) leaves a simple trace.  
 b. A-bar movement obligatorily leaves a copy which is converted to an operator variable construction in accordance with economy considerations.

Given this distinction, we would like to know whether it is plausible to assume that A-reconstruction is the result of the same mechanism that was proposed for A-bar movement in the previous section. As far as I can see, there are two possibilities. One possibility is that A-movement is incapable of leaving a copy at the tail of the chain, and that therefore it must resort to another scope shifting operation such as Quantifier Lowering (as assumed in chapter 2).<sup>24</sup> Another possibility is that A-movement can optionally leave a copy and that when it does, Scope Reconstruction is available. In any event, Scope Reconstruction is reflected in the syntax and the results of section 5.3 are predicted.

For presentational purposes, I assume that Quantifier Lowering is necessary for A reconstruction. A sketch of the possible LF structures that are derived from the two types of movement is illustrated in (34-35).

- (34) Scope Reconstruction in A-movement.

SS:	Someone [that she knows]	is likely [t to win the lottery].
LF <sub>1</sub> :	Someone [that she knows]	is likely [t to win the lottery].
LF <sub>2</sub> :		is likely [Someone[that she knows]
		[t to win the lottery]].

<sup>24</sup> For an argument in favor of this possibility see Chomsky (1995). As pointed out to me by David Pesetsky and Irene Heim, the necessary stipulation about A movement could be derived from an assumption that has an air of an explanation to it, namely the assumption that copies must receive Case.

## (35) Scope Reconstruction in A-bar movement.

SS<sub>1</sub>: How many people [that she knows] (*adjunct inserted before movement*)  
is Mary likely to hire how many people [that she knows].

LF<sub>1,1</sub>: How<sub>n</sub> is Mary likely to hire n many people [that she knows].

LF<sub>1,2</sub>: How<sub>n</sub> n many people [that she knows]<sub>x</sub>  
is Mary likely to hire x people [that she knows]

SS<sub>2</sub>: How many people [that she knows] (*adjunct inserted after movement*)  
is Mary likely to hire how many people.

LF<sub>2</sub>: How<sub>n</sub> n many people [that she knows]<sub>x</sub>  
is Mary likely to hire x people.

For A-movement, there is one SS and two LFs that differ depending on whether or not there is Quantifier Lowering. The latter yields what we have called Scope Reconstruction, and has the consequences we have discussed for BT(C) (section 5.3). For A-bar movement, there are two SSs that differ depending on whether the adjunct is inserted before or after movement. Only SS<sub>2</sub> can bleed BT(C). However, only SS<sub>1</sub> can bring about Scope Reconstruction (hence the consequences in section 5.2).<sup>25</sup>

## 6.7. Condition A and Scope-Economy

In the previous sections, I have argued that BT(C) applies at LF and only at LF. If this conclusion is correct, BT(C) can be used to investigate various covert properties of LF representations. One obvious question to ask is whether BT(C) can be used to test the predictions of Scope-Economy.

Consider the structural configurations in (36) and (37), where linear precedence represents structural prominence.

- (36) a. [<sub>DPi</sub> Someone] ....pronoun<sub>i</sub>.... [<sub>DPj</sub> every... r-exp<sub>i</sub>...]  
b. [<sub>DPi</sub> Many boys] ....pronoun<sub>i</sub>.... [<sub>DPj</sub> every... r-exp<sub>i</sub>...]

- (37) a. [<sub>DPi</sub> Bill]....pronoun<sub>i</sub>.... [<sub>DPj</sub> every... r-exp<sub>i</sub>...]  
b. [<sub>DPi</sub> Every boy]....pronoun<sub>i</sub>.... [<sub>DPj</sub> every... r-exp<sub>i</sub>...]

<sup>25</sup> SS<sub>2</sub> has only one LF in which *many* has wide scope. SS<sub>1</sub> has two LFs that differ in the scope they assign to *many*





anaphor contained in a quantifier could be bound long-distance as long as QR can appropriately extend the governing domain of the anaphor. If this expectation turns out to be correct, we could use BT(A) to see what the restrictions on QR are.

Consider the contrast between (41) and (42). In (41), just as in (40), BT(A) is satisfied. However, the sentences are acceptable only when the embedded clause is interpreted with Inverse Scope. This suggests that the sentences in (41) are the covert analogs of (40). In other words, it suggests that in these sentences covert QR can target the embedded object and thus extend the domain in which the reciprocal needs to be bound so as to include the matrix subject. (The embedded object QP can move by QR over the embedded subject, and the output of QR is identical (in the relevant respects) to the output of Wh-movement in (40).)

- (41) a. The two rivals hoped that someone would hurt (every one of) each-other's operations.  $*(\exists > \forall)$   $(\forall > \exists)$   
 b. John and Bill hoped that an inspector would supervise (every one of) each other's buildings.  $*(\exists > \forall)$   $(\forall > \exists)$   
 c. John and Bill hoped that many inspectors would supervise (every one of) each other's buildings.  $*(\exists > \forall)$   $(\forall > \exists)$
- (42) a. \*The two rivals hoped that Bill would hurt (every one of) each-other's operations.  
 b. \*John and Bill hoped that Mary would supervise (every one of) each other's buildings.  
 c. \*John and Bill hoped that every inspector would supervise (every one of) each other's buildings.

In (42), BT(A) is not satisfied. The question is why? Scope-Economy provides an answer. In (42), in contrast to (41), optional QR (over the embedded subject) does not reverse the relative scope of two scopally non-commutative quantifiers and as such is ruled-out by Scope-Economy. The resulting LF thus violates BT(A).

Obviously there are more predicted interaction between BT(A) and Scope-Economy. More specifically, we can state a predicted generalization which is very similar in its logic to other generalizations discussed in chapter 2. Unfortunately, many of the sentences are too long and complicated to yield reliable judgments. Nevertheless, I find the contrast between (41) and (42) encouraging, and would like to take it as further evidence in favor of Scope-Economy and in favor of the assumption that BT applies at LF.

## 6.8. Conclusion

To sum up the previous two chapters, we have seen evidence that the predictions of BT(C) come out right only if we assume that this condition applies to the structures that get interpreted. On the one hand, we have seen that (if Quantifier Lowering exists) BT(C) must "see" the output of this LF operation (section 5.3). On the other hand, we have seen that BT(C) doesn't see the SS input to the LF operation of Quantifier Raising (section 6.2-6.4.). The reason it looks as though BT(C) inspects a pre-QR structure relates to a special property of A-bar chains: movement leaves a copy which can be eliminated only when necessary (under Antecedent Contained Deletion). This special property provides a syntactic account of A-bar reconstruction (section 6.6) which, in turn, explains the correlation with BT(C) (section 5.2). Reconstruction may work differently for A- and A-bar movement. Nevertheless, in both cases it is reflected at LF. For this reason, the semantic mechanism of type-lifting must be restricted (section 5.4).

Finally we considered the fact that BT(A), in contrast to BT(C), is affected by A-bar movement (6.7.). Although we do not understand this difference, we can use it to study scopal properties of Scopally Uninformative sentences. The result of this investigations yields new evidence in favor of Scope-Economy.

## References

- Aoun, Joseph (1985). *A grammar of anaphora*. Cambridge, MA: MIT Press.
- Aoun, Joseph. (1982). "On the Logical Nature of the Binding Principles: Quantifier Lowering, Double Raising of *There* and the Notion Empty Element," in Pustejovsky, J. and P. Sells (eds.), *Proceeding of NELS 12*, GLSA, UMass, Amherst.
- Aoun, Joseph and Li, Yen-hui Audrey. (1993). *Syntax of Scope*. Cambridge: MIT-Press.
- Abusch, Dorit. (1994), "The Scope of Indefinites," *Natural Language Semantics*, 2.2. 83-135 .
- Adger, David (1995) "Meaning, Movement and Economy," in Raul Aranovich, William Byrne, Susanne Preuss, and Martha Senturia (eds.) *The Proceedings WWCFL 13*, Stanford Linguistics Association, CSLI, pp. 451-466.
- Asher, Nicholas, Daniel Hardt and Joan Busquets (1997) "Discours, Parallelism, Scope and Ellipsis," *SALT 7*.
- Baker, Charles L. (1970) "Notes on the Description of English Questions: The Role of an Abstract Question Morpheme," *Foundations of Language* 6, 197-219.
- Barss, Andy (1986). *Chains and Anaphoric Dependencies*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Barwise, Jon and Robin Cooper (1981). "Generalized Quantifiers and natural language," *Linguistics and Philosophy* 4: 159-219.
- Beck, Sigrid (1996). "Wh-construction and transparent Logical Form," unpublished doctoral dissertation, Tübingen.
- Beghelli, Filippo (1993). "A Minimalist Approach to Quantifier Scope," *Proceedings of NELS*.
- Beghelli, Filippo and Tim Stowell (1995). "The Direction of Quantifier Movement" ms. UCLA .
- Belvadi, A. (1989) "Quantified NPs and Pronouns," Unpublished ms., University of Rochester.
- Bobaljik, Jonathan (1998) "The Autonomy of Syntax and the Typology of Movement," *GLOW Newsletter*, 40, 22-23.
- Cinque, Guglielmo (1990) *Types of A-bar Dependencies*, MIT Press.
- Chierchia, Gennaro (1995a) "Individual-Level Predicates as Inherent Generics," in G. N. Carlson and F. J. Pelletier, eds. *The Generic Book*, The University of Chicago Press.
- Chierchia, Gennaro (1995b). *Dynamics of Meaning: Anaphora Presupposition and the Theory of Grammar*, The University of Chicago Press.
- Chomsky, Noam (1957). *Syntactic Structures*, The Hague: Mouton.
- Chomsky, Noam (1977). *Essays on Form and Interpretation*, Amsterdam: Elsevier North-Holland.
- Chomsky, Noam (1979) *Rules and Representations*, New-York: Columbia U.P.
- Chomsky, Noam (1981). *Lectures on Government and Binding*, Dordrecht: Foris.
- Chomsky, Noam (1986a) *Barriers*, MIT Press.
- Chomsky, Noam (1986b), *Knowledge of Language*, New-York: Praeger.
- Chomsky, Noam (1989), "Some Notes on Economy of Derivations and Representations," printed as chapter 2 of Chomsky 1995.
- Chomsky, Noam (1992), "Explaining Language Use," *Philosophical Topics*, 20:1: 205-231.
- Chomsky, Noam (1993) "A Minimalist Program for Linguistic Theory," in *The View From Building 20*, Cambridge, MA: MIT Press (re-printed as chapter 3 of Chomsky 1995).

- Chomsky, Noam, (1994) *Bare Phrase Structure*, Cambridge, MA: MITWPL.
- Chomsky, Noam (1995) *The Minimalist Program*, Cambridge, MA: MIT Press.
- Chomsky, Noam (1998a) "Some Observations on Economy in Generative Grammar," in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky, eds. *Is the Best Good Enough*, MIT Press and MITWPL, 115-128.
- Chomsky, Noam (1998b) "Minimalist Inquiries: the framework," ms. MIT.
- Chomsky, Noam and Howard Lasnik (1992) "The theory of Principles and Parameters," In J. Jacobs and A. von Stechow, W. Sternefeld, and T. Vennemann, eds. *Syntax : An international handbook of contemporary research*, Berlin: de Gruyter.
- Chomsky, Noam & Howard Lasnik (1993). Principles and Parameters Theory. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research*, Berlin: Walter de Gruyter.
- Chung, Sandra (1982) "Unbounded Dependencies in Chamorro Grammar," *Linguistic Inquiry*, Volume 13, 1, 39-77.
- Collins, Chris (1993) *Topics in Ewe Syntax*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Cooper, Robert (1979) "Variable Binding in Relative Clauses," in F. Guenther & S. J. Schmidt eds.: 131-169.
- Cormack, Annabel (1984) "VP anaphora: Variables and scope," In F. Landman and F. Veltman *Varieties of Formal Semantics*, Foris, Dordrecht.
- Cinque Guglielmo (1982). "Constructions allowing left dislocation, 'Connectedness,' Move- $\alpha$  and the ECP, ms., MIT, and Università di Venezia.
- Cinque, Guglielmo (1993), "A Null Theory of Phrase and Compound Stress," *Linguistic Inquiry* 24, 239-298.
- Cresti, Diana (1995) "Extraction and Reconstruction," *Natural Language Semantics*,
- Cresti, Diana (1996) "Economy and the Scope of Amount Phrases," *the proceedings of WCCFL 14*..
- Cresti, Diana (1997) *Specific Indefinites*, MIT Doctoral dissertation.
- Davidson, Donald (1986) "A Nice Derangement of Epitaphs," in E. Lepore, ed. *Truth and Interpretation*, Oxford: Blackwell.
- Dahl, Östen (1974) "How to open a sentence: abstraction in natural language," In *Logical Grammar Reports*, no. 12, University of Göteborg.
- Dalrymple, M., S. Shieber, and F. Pereira (1991) Ellipsis and Higher-Order Unification," *Linguistics and Philosophy*, 14:399-452.
- Diesing, Molly, (1992) *Indefinites*, Cambridge, MA.: MIT press.
- Dobrovie-Sorin, Carmen (1992) *The Syntax of Romanian: Comparative studies in Romance*, Dordrecht: Foris.
- Enç, Mürvit (1986) "Towards a Referential Analysis of Temporal Expressions," *Linguistics and Philosophy*, 9: 405-426.
- Engdahl, Elisabet (1980) *The Syntax and Semantics of Questions in Swedish*, doctoral dissertation, Umass, Amherst: GLSA.
- Engdahl, Elisabet (1986) *Constituent Questions: The Syntax and Semantics of Questions with Special Reference to Swedish*, D. Riedel.
- von Fintel, Kai (1995) "A Minimal Theory of Adverbial Quantification," unpublished ms., MIT.
- Fiengo, Robert and Robert May (1994) *Indices and Identity*, MIT Press.
- Fox, Danny (1993). Chain and binding—A modification of Reinhart and Reuland's 'Reflexivity'. Ms., MIT.
- Fox, Danny (1995a). "Economy and Scope," *Natural Language Semantics*, 3: 283-341.

- Fox, Danny (1995b). Condition C effects in ACD. In R. Pensalfini and H. Ura (eds.) *Chapters on minimalist syntax*, MITWPL 27. Cambridge, MA: MIT Working Chapters in Linguistics, pp. 105–119.
- Fox, Danny (1995c). “Economy and Parsing,” paper presented in *IATL* and in the *Jersey Syntax Circle*.
- Fox, Danny (1998a). “Locality in Variable Binding,” in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky, eds. *Is the Best Good Enough*, MIT Press and MITWPL, 129–156.
- Fox, Danny (1998b). “Successive Scyclic Movement: Evidence from Reconstruction and Antecedent Contained Deletion,” *GLOW Newsletter*, 40: 36–37.
- Fox, Danny (in press). “Reconstruction, Binding Theory and the Interpretation of Chains,” to appear in *Linguistic Inquiry*.
- Fox, Danny and Uli Sauerland (1996) “Scope Illusions,” *NELS* 26.
- Frampton, John (1991). “Relativized Minimality: A review,” *The Linguistic Review*, 8:1–46.
- Frey, Werner (1989). Syntaktisch Bedingungen für die Interpretation, unpublished doctoral dissertation, Universität Stuttgart.
- Friedin, Robert (1986). “Fundamental Issues in the Theory of Binding,” in B. Last, ed. *Studies in the Acquisition of Anaphora*, Dordrecht: Reidel.
- Goodall, Grant, (1987) *Parallel Structures in Syntax*, Cambridge, England: Cambridge UP.
- Golan, Yael, (1993) “Node Crossing Economy, Superiority and D-linking,” unpublished ms. Tel-Aviv University.
- Grodzinsky, Yosef and Tanya Reinhart (1993) “The Innateness of Binding and Coreference,” *Linguistic Inquiry*, 24, 1: 69–102.
- Grimshaw, Jane (1997) “Minimal Projections, Heads, and Optimality,” *Linguistic Inquiry*.
- Hardt, Daniel (1992) “VP Ellipsis and Semantic Identity,” ,” *Proceedings of SALT 2*, Ohio State University, 145–162.
- Heim, Irene (1982). *The semantics of Definite and Indefinite Noun Phrases*, Doctoral dissertation, University of Massachusetts, Amherst.
- Heim, Irene (1987). “Where Does the Definiteness Restriction Apply?” In E. Reuland and A. ter Meulen, eds. *The Representation of (In)definiteness*, 21–42, MIT Press.
- Heim, Irene (1991), “Articles and Definiteness,” Published in German as “Artikel und Definitheit,” in A.v. Stechow & D. Wunderlich (eds.), *Semantics. An International Handbook of Contemporary Research*, Berlin: de Gruyter, 1991.
- Heim, Irene (1993). Anaphora and semantic interpretation: A reinterpretation of Reinhart’s Approach. Sfs-Report-07-93, University of Tübingen.
- Heim, Irene and Angelika Kratzer (1998), *Semantics in Generative Grammar*, Blackwell, Oxford.
- Hendriks, Herman (1993), *Studied Flexibility: Categories and Types in Syntax and Semantics*, Ph.D. dissertation, University of Amsterdam.
- Heycock, Caroline (1995). “Assymetries in Reconstruction,” *linguistic Inquiry*, 26:4, 547–570.
- Higginbotham, James and Robert May (1981) “Questions, Quantifiers and Crossing,” *The Linguistics Review* 1, 41–80.
- Higginbotham (1983). “Logical Form, Binding and Nominals,” in *Linguistic Inquiry*, 14: 395–420
- Hintikka, Jukka (1977) “Quantifiers in Natural Languages: Some Logical Problems,” *Linguistics and Philosophy*, 1:153–172.
- Hirschbühler, Paul, (1982) “VP-Deletion and Across-the Board Quantifier Scope,” In J. Pustejovsky and P. Sells, eds., *Proceedings of NELS 12*, 132–139. GLSA,

- University of Massachusetts, Amherst.
- Hoji, Hajime. 1985. *Logical Form Constraints and Configurational Structure in Japanese*. Seattle: University of Washington dissertation.
- Hornstein, Norbet (1995). *Logical Form: From GB to Minimalism*, Blackwell Publishers, Mass.
- Huang, C.-T. James, (1993) "Reconstruction and the Structure of VP: Some Theoretical Consequences, *Linguistic Inquiry* 24:103-138.
- Jackendoff, Ray (1972). *Semantic Interpretation in Generative Grammar*, MIT Press.
- Jacobson, Pauline (1992) "Antecedent Contained Deletion in a Variable Free Semantics," *Proceedings of SALT 2*, Ohio State University, 347-366.
- Jacobson, Pauline (1997) "Where (if anywhere) is transderivationality located?," in P. Culicover and L. McNally (eds.), *The Limits of Syntax*, Academic Press, 1997, pp. 193-214.
- Jayaseelan, K. A., (1990) "Incomplete VP Deletion and Gapping," *Linguistics Analysis* 20: 64-81.
- Johnson, Kyle (1997). Ellipsis paper, ms. University of Massachusetts, Amherst.
- Johnson, Kyle and Satoshi Tomioka (1997)
- Kamp, Hans (1981). "A Theory of Truth and Semantic Representation," In J. Groendijk et al., eds., *Formal Methods in the Study of Languages*, Amsterdam: Mathematical Center.
- Kayne, Richard (1981) "Two notes on the NIC," in A. Belletti, L. Brandi and L. Rizzi (eds.) *Theory of Markedness in Generative Grammar. Proceedings of the 1979 GLOW conference*, Pisa: Scuola Normale Superiore, 317-346.
- Kayne, Richard (1995) *The Anti-symmetry of Syntax*, MIT Press.
- \*Kayne, Richard (1998)
- Keenan Ed and Faltz (1985) *Boolean Semantics for NL*, D. Reidel, Dordrecht.
- Kehler, Andy (1993). A discourse copying algorithm for ellipsis and anaphora resolution. *Proceedings of EACL*.
- Kennedy, Christopher, (1997) "Antecedent Contained Deletion and the Syntax of Quantifiers," *linguistic Inquiry*.
- Kitagawa, Yoshihisa (1991). Copying identity. *Natural Language and Linguistic Theory* 9: 495-536.
- Kitahara Hisatsugu (1994), "Raising Quantifiers without Quantifier Raising," ms. Harvard University.
- Koizumi, Masa (1994) "Layered Specifiers," In M. González, ed. *Proceedings of NELS 24*, University of Massachusetts, Amherst: 255-269.
- Kratzer, Angelika (1995a), "Stage-Level and Individual-Level Predicates," in G. N. Carlson and F. J. Pelletier, eds. *The Generic Book*, The University of Chicago Press.
- Kratzer, Angelika (199?) "Scope or Pseudo Scope? Are There Wide Scope Indefinites," to appear in F. Ham and A. Von Stechow (eds.), *Proceeding of Recent Developments in the Theory of Natural Language*, Universität Tübingen.
- Krifka (1998). "Scope Inversion under the Rise-Fall Contour in German," *Linguistic Inquiry*, 29:1 75-112.
- Laka, Itziar (1993) *Negation in Syntax: On the Nature of Functional Categories and Projections*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Lakoff, George (1970) "Repartee," *Foundations of Language* 6. p. 389-422.
- Lakoff, George (1986) "Frame Semantic Control of the Coordinate Structure Constraint," *CLS* 22: 152-167.
- Lappin, Shalom (1993) "Ellipsis Resolution at S-structure," In A. J. Schafer, ed.

- Proceedings of NELS 23*, University of Massachusetts, Amherst: 255-269.
- Lasnik, Howard (1972) *Analysis of Negation in English*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Lasnik, Howard (1976) "Remarks on Coreference," *Linguistic Analysis*, 2,1.
- Lasnik, Howard (1993) "Lectures in Minimalist Syntax," Cambridge, MA: MITWPL.
- Lasnik, Howard (1995) "A Remark on Psuedo-Gapping," In R. Pensalfini and H. Ura (eds.) *Chapters on minimalist syntax*, MITWPL 27. Cambridge, MA: MIT Working Chapters in Linguistics, pp. 143-164.
- Lasnik, Howard and Memoro Saito (1991). "On the Subject of Infinitives," In L. M. Dobrin, L. Nichols, and R. M. Rodriguez, eds., CLS 27. Part 1: *The General Session*, Chicago Linguistic Society, University of Chicago.
- Lasnik, Howard and Memoro Saito (1992) *Move  $\alpha$* , Cambridge, MA.: MIT Press.
- Larson, Richard, (1990) "Double Objects Revisited: Reply to Jackendoff," *LI* 589-632.
- Larson, Richard and Robert May (1990). "Antecedent Containment or Vacuous Movement: Reply to Baltin," *Linguistic Inquiry*, 21, 103-122.
- Lebeaux, David (1988). "Language Acquisition and the Form of the Grammar," Doctoral dissertation, University of Massachusetts, Amherst.
- Lebeaux, David (1990). "Relative Clauses, Licensing and the Nature of Derivation," *Proceedings of Nels 20*, GLSA, UMass, Amherst, 318-332.
- Lebeaux, David (1994). "Where Does Binding Theory Appley?" ms., University of Maryland.
- Lewis, David (1972) "General Semantics," In D. Davidson and G. Harman, eds. *Semantics for Natural Language*, Dordrecht: D. Reidel.
- Longobardi, Giuseppe (1987). "Extraction from NP and the Proper Notion of Head Government," In A. Giorgi and G. Longobardi (eds.), *The Syntax of Noun Phrases*. Cambridge University Press, pp. 57-112.
- Ludlow, Peter and Steven Neale (1991) "Indefinite Descriptions: In Defence of Russell," *Linguistics and Philosophy* 14, 171-202.
- Marantz, Alec (1994) "A late Note on Late Insertion," in Y. -S. Kim et. al. (eds.), *Explananation in Generative Grammar*, Hankuk Publishing Co., Seoul, 396-413.
- May, Robert (1977) "The Grammar of Quantification," Ph.D. dissertation, MIT.
- May, Robert (1988) "Ambiguities of Quantification," *LI*, 19:118-135.
- May, Robert, (1985) *Logical Form, Its Structure and Derivation*, MIT press.
- McCloskey, James (1996) "Subjecthood and Subject Positions," unpublish ms. USCS.
- McCawley, James (1988) *The Syntactic Phenomena of English*, The University of Chicago Press.
- Merchant, Jason (1998) "On the Extent of Trace Deletion in ACD," Unpublished ms, Uil/OTS and UC-Santa Cruz.
- Milsark, Gary, (1974) *Existential Sentences in English*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Moltmann, Friederike, (1992) *Coordination and Comparatives*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Moltmann, Friederike and Anna Szabolcsi (1994) "Scope Interaction with Pair-List Quantifiers," *Proceedings of NELS 24*, University of Massachusetts, Amherst: 381-395.
- Muadz, H., (1991) *Coordinate Structures: A Plannar Representation*, Phd. dissertation, University of Arizona, Tuscon.
- Musan, Renate (1995). *On the Temporan Interpretation of Noun Phrases*, MIT dissertaion, MITWPL.



- Nissenbaum, Jon (1998) "Movement and Derived Predicates: Evidence from Parasitic Gaps," to appear in *The Proceedings of WCCFL 17*.
- Oka T. (1993) *Minimalism in Syntactic Derivation*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Partee, Barbara, (1989) "Binding Implicit Variables in Quantified Contexts", *CIS* 25.
- Partee, Barbara & Emmon Bach (1984). Quantification, pronouns and VP anaphora. In J. Groenendijk, T. Janssen, M. Stokhof (eds.), *Truth, Interpretation and Information*, Dordrecht: Foris, 99–130.
- Pesetsky, David, (1995) *Zero Syntax*, MIT press.
- Pesetsky, David, (in prep.) *Syntax at the Edge: Optimality Effects in Sentence Grammar*
- Phillips, Colin (1996) *Order and Structure*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Pica, Pier and William Snyder (1994) "Weak Crossover, Scope, and Agreement in a Minimalist Framework," to appear in *The Proceedings of WCCFL 13*.
- Poesio, Massimo and Alessandro Zucchi (1992) "On Telescoping," *Proceedings of SALT 2*, Ohio State University, 347-366.
- Reinhart, Tanya (1977) *The Syntactic Domain of Anaphora*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Reinhart Tanya (1981) "A second COMP position," *Proceedings of the 1979 GLOW conference*, Estratto, Scuola Normale Superiore Di Pisa.
- Reinhart, Tanya (1983) *Anaphora and Semantic Interpretation*, Croom Helm, London.
- Reinhart, Tanya (1987) "Specifier and Operator Binding," in Reuland E. and A. Ter Meulen eds.: 130-167.
- Reinhart, Tanya (1991) "Non-Quantificational LF," In A. Kasher (ed.) *The Chomskian turn*,
- Reinhart, Tanya, (1994) "Wh-in-situ in the Framework of the Minimalist Program", Utrecht, Holland: OTS.
- Reinhart, Tanya (1995) *Interface Strategies.*, Utrecht, OTS.
- Reinhart, Tanya, (1998a) "Wh-in-situ in the Framework of the Minimalist Program", *NALS*.
- Reinhart, Tanya (1998b), "Covaluation," Paper presented in SALT VII.
- Reinhart, Tanya, (1998b) "Wh-in-situ in the Framework of the Minimalist Program", *NALS*.
- Riemsdijk, Henk van, and Edwin Williams (1981). "NP-Structure, *The Linguistic Review* 1:171-217.
- Ristad, Eric (1992). *Computational Structure of Natural Language*. Ph.D. dissertation, MIT. Distributed by MIT Working Chapters in Linguistics.
- Rizzi, Luigi (1986). On chain formation. In Hagit Borer (ed.), *The grammar of pronominal clitics*, Syntax and Semantics 19.
- Rizzi, Luigi (1990), *Relativized Minimality*, MIT Press.
- Roberts, Craige (1987) *Modal Subordination, Anaphora and Distributivity*, Doctoral dissertation, University of Massachusetts, Amherst.
- Rodman, Robert (1976) "Scope Phenomena, 'Movement Transformations,' and Relative Clauses," In B. Partee ed., *Montague Grammar*, Academic Press, New-York: 165-176.
- Romero, Maribel (1996). "The (Cor)relation between Scope Reconstruction and Connectivity Effects," ms. UMass Amherst.
- Rooth, Mats (1985). *Association with focus*. Ph.D. dissertation, University of Massachusetts at Amherst.

- Rooth, Mats (1992) "Ellipsis redundancy and Reduction Redundancy," In S. Berman and A. Hestvik, eds., *Proceedings of Stuttgart Ellipsis Workshop*. Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29. IBM Germany, Heidelberg.
- Rooth, Mats and Barbara Partee, "Conjunction, Type Ambiguity and Wide Scope 'or'," In D.P. Flickinger, M. Macken and N. Wiegard, eds., *Proceedings of WCCFL 1*, 353-362, Stanford University.
- Ross, John R. (1967). *Constraints on Variables in Syntax*. Ph.D. dissertation, MIT.
- Rullman, Hotze (1995) *Maximality in the Semantics of wh-constructions*, Phd, Umass, Amherst: GLSA.
- Rullman, Hotze and Sigrid Beck (1997). "Reconstruction and the Interpretation of *Which*-Phrases," to appear in the proceedings of the Tübingen Reconstruction Workshop.
- Ruys, Eddy (1993) *The Scope of Indefinites*, OTS Dissertation Series, Utrecht.
- Sag, Ivan (1976) *Deletion and Logical Form*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Sauerlan, Uli (1997) "Schwarzschild's theory of Focus," handout from the LF Reading Group, MIT.
- Sauerland, Uli (forthcoming) *The Making and Meaning of Chains*, Doctoral dissertation, MIT.
- Selkirk, Elisabeth (1995), "Sentence Prosody: Intonation, Stress, and Phrasing," in J. Goldsmith (ed), *The Handbook of Phonological Theory*, 550-569, Oxford: Basil Blackwell.
- Schütze, Carson, T. (1995), "PP Attachment and Agreement," in Schütze, C., J. Ganger and K. Broihier (eds.) *Papers on Language Processing and Acquisition*, MITWPL, 26, MIT.
- Schwarzschild, Roger (in press), "Givenness and Optimal Focus," to appear in *Natural Language Semantics*.
- Sczabolcsi, Anna (1997) *Ways of Scope Taking*, Kluwer Academic Publishers, Dordrecht.
- Sharvit, Yael (1997) *The Syntax and Semantics of Functional Relative Clauses*, Doctoral dissertation, Rutgers University.
- Sportiche, Dominique (1996). "A-Reconstruction and Constituent Structure," handout of a talk given at Université de Montréal, October 1996.
- Sternefeld, Wolfgang (1997) "Fancy Binding," Ms., Universität Tübingen.
- Tada, Hiroaki (1993). *A/A-bar partition in derivation*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Takano (1995). "Predicate Fronting and Internal Subjects," *Linguistic Inquiry* 26: 327-340.
- Tancredi, Chris, (1992) *Deletion, Deaccenting and Presupposition*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Taraldsen, Tarald, (1981) "The Theoretical Interpretation of a Class of Marked Extractions," In A. Belletic, L. Brandi, and L. Rizzi (eds.), *Theory of Markedness in Generative Grammar, Proceedings of the 1979 Glow Conference*, 475-516. Scuola Normale Superiore, Pisa.
- Tomioka, Satoshi (1995). On the mismatch between variable binding and sloppy identity. To appear in the proceedings of WCCFL.
- Tomioka, Satoshi, (1995) "[Focus]<sub>F</sub> Restricts Scope: Quantifier in VP Ellipsis," to appear in *The Proceedings of SALT*.
- Tsai, W., (1994) On Economizing the Theory of A-bar Dependencies, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.
- Truckenbrodt, Hubert (1995) *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*, Doctoral dissertation, MIT, Cambridge, MA: MITWPL.

- Ura, Hiroyuki, (1994) "Varieties of Raising and the Feature-Based Phrase Structure Theory, MIT Occasional Papers in Linguistics, #7, Cambridge, MA. MITWPL.
- Westerstahl, Dag (1994) "Self-Commuting Quantifiers," unpublished ms. Stockholm University.
- Wilder, Chris (1997) "Phrasal Movement in LF: de re readings, VP Ellipsis and Binding," *Proceeding of NELS 27*, 425-440.
- Wilkinson, Karina (1996). The scope of *even*. To appear in *Natural Language Semantics*.
- Williams, Edwin (1977) "Discourse and logical form," *Linguistic Inquiry* 8.101-139.
- Williams, Edwin (1978) "Across-the-board rule application," *Linguistic Inquiry* 9: 31-43.
- Williams, Edwin (1985) "PRO in NP". *Natural Language and Linguistic Theory* 3: 277-295.
- Williams, Edwin (1987) "Implicit Arguments, the Binding Theory, and Control". *Natural Language and Linguistic Theory* 5: 151-180.
- Winter, Yoad (1995). "On the Formalization of Choice Functions as Representing the Scope of Indefinites," the Proceedings of *ESSLI 7*, Conference on Formal Grammar, Barcelona.
- Wold, Dag (1995). "Identity in Ellipsis: Focal Structure and Phonetic Deletion," Manuscript, MIT.
- Wold, Dag (forthcoming). Doctoral dissertation, MIT.
- Zubiareta, Maria Luisa (in press) *Prosody, Focus and Word Order*, to appear in MIT Press.