MOVE α, SCOPE, AND RELATIVIZED MINIMALITY

by

Masanori Nakamura

A Thesis Submitted to
the Faculty of Graduate Studies and Research
in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts

Department of Linguistics
McGill University
Montreal, Québec

To my grandparents
Nobuyuki and Setsu

ABSTRACT

This thesis deals with two aspects of operators within the framework of Government and Binding theory; (i) how they are assigned their scope, and (ii) how they are licensed. In an attempt to answer these questions, the relation of Move α (such as scrambling, NP-movement, and *wh*-movement) to the scope of operators and the licensing of *wh*-elements, negative polarity items, and adverbs are examined. It is argued that scope assignment is dictated by the Scope Principle and the Empty Category Principle. It is also argued that licensing of operators is determined by the Feature-Dependent Item Criterion. These principles and criterion make use of the concept of Government Theory Compatibility, which is built into Relativized Minimality. It is suggested that this concept should be characterized in terms of a set of lexical features. The approach advocated here accounts for the interpretative and distributional behavior of operators without recourse to parameterization of LF principles.

RÉSUMÉ

Ce mémoire traite de deux aspects des opérateurs dans le cadre de la théorie du Gouvernement et du liage, à savoir: i) comment on en attribue la portée et ii) comment ils sont identifiés. Afin de répondre à ces questions, j'étudierai les rapports entre les règles transformationnelles «déplacer α» (notamment le «scrambling», le déplacement des syntagmes nominaux et des interrogatifs de type WH) et le domaine des opérateurs et l'identifcation des interrogatifs WH, des expression de polarité négative et des adverbes. Je soutiendrai que l'attribution du portée des opérateurs est régie par le Principe de la portée et par le Principe des catégories vides. Je soutiendrai par ailleurs que l'identification des opérateurs est régie par le «Feature-Dependent Item Criterion». Ces critères et principes font appel à la notion de compatibilité énoncée dans la Théorie du gouvernement, qui fait partie de la Minimalité relativisée. Je me propose de caractériser ce concept au moyen d'un ensemble de traits lexicaux. La méthode proposée ici caracterise l'interprétation et la distribution des opérateurs sans avoir recours à la paramétrisation des principes de la forme logique (LF).

ACKNOWLEDGEMENTS

This thesis is a tiny outcome of the project I initiated on the damp evening when I arrived at Dorval Airport. It is with great pleasure that I acknowledge people who have helped me directly or indirectly.

First of all, I wish to express my sincere gratitude to Lisa Travis, my advisor, for hours of insightful, clarifying, and challenging discussions, her encouragement, and her kindness, which were all indispensable for me to write this thesis. She was generous enough to keep her office door open even when she was on sabbatical. I also wish to express my appreciation to Mark Baker for his helpful suggestions on some portions of this thesis and for his precise instruction which helped me to understand linguistic theory.

I am grateful to Alan Juffs and Edward Ikeda for their useful advice, academic or non-academic, their judgements, and, above all, their friendship. Many thanks are due to the Ikedas for their hospitality I received during my stay at their place in Toronto. I am also indebted to Mihoko Zushi for her comments, personal library, and kindness.

My study at McGill University has been made possible by a scholarship from the Ishizaka Foundation. I am most grateful to its committee members and staff for giving me an invaluable opportunity to study in Canada.

Very special thanks go to my parents, grandparents, and sister for their constant understanding and support.

Last but certainly not the least, thanks to my friends back in Japan for their friendly and encouraging correspondences in various ways.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

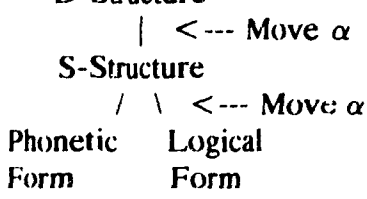
In this thesis, we are concerned primarily with the scopal and distributional behavior of operators. The syntax and semantics of operators have attracted a considerable amount of attention in linguistic theory (cf. Katz and Postal (1964), Jackendoff (1972), May (1977, 1985) among many others). One recurrent question has been: In what way and to what extent does syntax contribute to semantic interpretation? This is the general question we would like to address in the discussions to follow. Our hope is to shed some light on its related issues.

This introductory chapter has three aims; (i) to briefly mention the theoretical framework adopted here, (ii) to set the goals of this thesis, and (iii) to describe how this thesis is organized.

1.1. Theoretical Framework

The theoretical framework presupposed throughout this thesis is the Government-Binding (GB) Theory first laid out in Chomsky (1981), which may be now referred to more appropriately as the Principles-and-Parameters Approach (see Chomsky (1986a, 1989)). Under this approach, it is assumed that children are endowed with Universal Grammar (UG) which consists of principles and a finite number of parameters and that core grammars of individual languages are acquired by setting the parameters.

Here we will not attempt to give a detailed picture of the theory (see Chomsky (1981) and subsequent work). Instead, we will only sketch the model of grammar, which is relevant to the discussions that follow. It is assumed that grammar includes four different components, as is illustrated below:

(1)


D-structure is a level of representation at which the thematic relations among items are directly represented (cf. Chomsky (1981), Baker (1988) etc.). D-structure is mapped into S-structure by the application of "Move α"[1]. Instantiations of Move α between D-structure and S-structure include NP-movement, *Wh*-movement, scrambling, head movement and so forth. S-structure is a level of representation which mediates between D-structure and the "interface" levels, i.e., Logical Form (LF) and Phonetic Form (PF). LF is a level of representation at which the logical relations such as the scope of operators and predication relationships are explicitly represented and is related to S-structure again by Move α. The primary case of Move α in the mapping from S-structure to LF is Quantifier Raising (QR) (see May (1977, 1985)).[2] PF is a level of representation at which phonological structures are directly expressed. In this thesis, we are concerned exclusively with the right side of the grammatical model depicted in (1).

1.2. Goals

Given the model of grammar in (1), a question arises as to whether Move α has any impact on scope interpretation of Quantificational Phrases (henceforth QPs) or operators of various kinds since a D-structure representation undergoes Move α before it reaches LF, where logical relations are represented and interpreted.

Evidence for the claim that movement does affect scope interaction (contra the Katz-Postal Hypothesis, see Katz and Postal (1964)) seems ample. An example which clearly shows that syntactic movement proliferates scope interpretations comes from Japanese. As is well known, Japanese allows scrambling.[3] For example, the sentences in (2) are both perfectly grammatical:[4]

(2) a. Tsutomu-ga sono sake-o nonda (koto)
-NOM that -ACC drank fact
'Tsutomu drank that sake.'
b. Sono sake-o_i Tsutomu-ga t_i nonda (koto)
that -ACC -NOM drank fact
'That sake, Tsutomu drank t_i.'

It has been noted in the literature (Kuroda (1970), Hoji (1985, 1986) among others) that scrambling of a QP over another one induces scope ambiguity. Although (4a), the Japanese counterpart of the ambiguous English (3), is unambiguous, its scrambled version (4b) is ambiguous like (3):

(3) Someone loves everyone.

(4) a. Dareka -ga daremo -o aisiteiru (koto)
someone-NOM everyone-ACC love fact
'Someone loves everyone.'
b. Daremo $-o_i$ dareka -ga t_i aisiteiru (koto)
everyone-ACC someone-NOM love fact
'Everyone$_i$, someone loves t_i.'

(4a) has only one interpretation on which there is a person such that he or she loves each member of the group of people denoted by *daremo* 'everyone', whereas (4b) has not only this interpretation but also the one on which for each member of *daremo*, there is some person who loves him or her.

Given the fact that Move α in general has influence on scope phenomena, the following two questions come to mind immediately:

(5) a. What is the principle governing scope relations among operators?
b. What is the locality principle governing LF movement of operators?

The principle in (5a) should be able to capture not only the kind of cross-linguistic variation in scope interaction observed between (3) and (4a) but also the attested correlation between syntactic movement and scope such as the one in (4). The locality principle in (5b) is supposed to play a vital role in fixing the scope of operators. To see the point, compare (3) with the following example:

(6) Someone thinks that Robert loves everyone.

In sharp contrast to (3), (6) is interpreted unambiguously. In particular, the QP in the embedded clause is construed as having scope narrower than the matrix predicate. Therefore, we must conclude that in (6) the kind of QR of the lower QP that allows (3) to yield ambiguity is blocked by some locality principle.

There is another important question to be addressed when we consider the scope of operators. It has been long observed that scope-taking items like *wh*-elements, negative polarity items, and adverbs are dependent on their licensers. Consider the examples in (7)-(9) ((7) is again from Japanese, and (9) is taken from Jackendoff (1972:50)):[5]

(7) a. Dare-ga Mikiko-ni [Satoshi-ga deisuisita to] hanasimasita ka?
who-NOM -DAT -NOM got plastered COMP told Q
'Who told Mikiko that Satoshi got plastered?'
b. *Dare-ga Mikiko-ni [Satoshi-ga deisuisita ka] hanasimasita?
who-NOM -DAT -NOM got plastered Q told
'Who told Mikiko whether Satoshi got plastered?'

(8) a. Richard did not see anyone.
b. *Anyone did not see Richard.

(9) a. Stanley easily ate his Wheaties.
b. *Easily Stanley ate his Wheaties.

(7)-(9) contain the *wh*-element *dare* 'who', the negative polarity item *anyone*, the adverb *easily* respectively. It must be the case that the operators in the (a) examples are licensed by their appropriate licensers and thus well-formed, whereas those in the (b) examples are not licensed and thus ungrammatical. Therefore, the question is:

(10) What is the mechanism governing licensing of (certain) operators?

This thesis is an attempt to provide answers to the questions in (5) and (10). We also hope to gain some insight into the properties of UG by answering them.

1.3. Organization

The organization of this thesis is as follows. Chapter 2 addresses the questions in (5). In an attempt to answer them, the relation of Move α to the scope of operators is

explored in some detail. In Section 2.1., the Scope Principle and the Minimal Binding Requirement are introduced. In Section 2.2., scrambling in Japanese is examined in connection to scope interaction. There, it is claimed that the Minimal Binding Requirement can be subsumed under the Empty Category Principle, given a modified version of Relativized Minimality. Section 2.3. examines NP-movement in relation to scope interpretation. In Section 2.4., the interaction between a *wh*-phrase and a QP is considered, which leads us to a revision of the Scope Principle. The revised Scope Principle is made sensitive to Government Theory Compatibility, which is incorporated into Relativized Minimality.

Chapter 3 deals with the question in (10). In Section 3.1., the *Wh*-Criterion and the Neg-Criterion are introduced. In Section 3.2., the licensing of *wh*-elements, negative polarity items, and adverbs is discussed. There, the Feature-Dependent Item Criterion, which exploits the new concept of feature government, is presented. Feature government is simply defined in terms of m-command and Relativized Minimality.

Chapter 4 is devoted to the issue of how Government Theory Compatibility should be characterized. In Section 4.1., "inner island" phenomena are briefly considered. It is pointed out that the modified version of Relativized Minimality is superior to Rizzi's (1990) original version. In Section 4.2., an appropriate notion of Government Theory Compatibility is sought. It is suggested that Government Theory Compatibility should be defined on the basis of lexical features. In Section 4.3., the Scope Principle advanced in Chapter 2 is reexamined.

In Chapter 5, the main claims of this thesis are summarized, and a few eminent residual questions are briefly mentioned.

The overall conclusion of this thesis, which we believe is quite natural, is that scope interpretation is determined to a significant degree by the interaction of lexical properties of operators (Government Theory Compatibility in particular) and syntax (e.g. principles like the Empty Category Principle).

FOOTNOTES TO CHAPTER 1

1. Or more generally, "Affect α". See Lasnik and Saito (1984, 1992).

2. Other cases argued for in the literature, for example, include *wh*-movement (Huang (1981, 1982)), head movement (Baltin (1991)), and Left Dislocation (Pritchett (1990)).

3. See Saito (1985) for extensive discussion of scrambling in Japanese as an instance of Move α.

4. *Koto* 'fact' is added to the end of the sentence to avoid the unnaturalness which stems from the lack of a topic in the matrix sentence. The English translations for the Japanese sentences to follow are in most cases literal. We ignore *koto* in them.

5. In colloquial Japanese, the Q-morpheme *ka* may be replaced by *no* (usually, *ka* is used with polite verb forms, while *no* is used with non-polite verb forms) or may be omitted in the matrix clause. This option is not available for the embedded clause. When the Q-morpheme is omitted, marked rising intonation is required at the end of the sentence. Thus, sentences like (7b) can be made acceptable with such intonation.

CHAPTER 2

MOVE α AND SCOPE

2.0. Introduction

This chapter discusses in some detail the relation between Move α and the scope interaction among operators. We address the following questions raised in Chapter 1:

(1) a. What is the principle governing scope relations among operators?
b. What is the locality principle governing LF movement of operators?

In Section 2.1., the Scope Principle proposed by Aoun and Li (1989) and Ernst (1991) as an answer to (1a) and the Minimal Binding Requirement proposed by Aoun and Li (1989) in reply to (1b) are introduced and considered. Then, three instances of S-structure Move α, i.e., scrambling, NP-movement, and *wh*-movement, are discussed in turn. In Section 2.2., the effects of scrambling on scope interaction in Japanese are examined. There, it is argued that the Minimal Binding Requirement can be subsumed under a modified version of Rizzi's (1990) Relativized Minimality and thus ultimately under the Empty Category Principle (ECP). Section 2.3. examines NP-movement in relation to scope interpretations and shows that a wide range of data can be accounted for by the Scope Principle. Section 2.4. considers the scope interaction between a *wh*-phrase and a QP, which motivates a revision of the Scope Principle.

Our answers to (1a,b) are: (i) the revised Scope Principle which incorporates the notion of Government Theory Compatibility, used in Relativized Minimality, into its formulation, and (ii) the ECP, which utilizes Relativized Minimality (in addition to barriers), respectively. Thus, the main contention of this chapter is that Relativized Minimality (Government Theory Compatibility, to be precise) plays a crucial part in fixing the relative scope of operators.

2.1. The Scope Principle

This section introduces the Scope Principle proposed in the recent literature (Aoun and Li (1989), Ernst (1991)).[1] In connection to this principle, two claims are made; (i) chains in the definition of the Scope Principle should be defined in terms of antecedent government (cf. Chomsky (1986b), Rizzi (1990)), and (ii) the relative scope relations determined by the Scope Principle cannot be further computed in a transitive fashion.

2.1.1. Aoun and Li (1989)

Aoun and Li (1989) argue that the cross-linguistic differences in scope interaction of QPs between English and Mandarin Chinese can be explained by the two requirements in (2) and (3):

(2) Minimal Binding Requirement (MBR)
Variables must be bound by the most local potential antecedent (A'-binder).[2]

(3) The Scope Principle (SP)
A quantifier A has scope over a quantifier B in case A c-commands a member of the chain containing B.

"Potential antecedent (A'-binder)" in (2) is defined as follows:

(4) A qualifies as a potential A'-binder for B iff A c-commands B, A is in an A'-position, and coindexing of (A,B) would not violate any grammatical principle.[3]

Aoun and Li suggest that either c-command in (5a) or m-command in (5b) may be adopted for "c-command" for their purposes:

(5) a. α c-commands ß if α and ß do not dominate each other and the first branching node dominating α also dominates ß.
b. α m-commands ß if α and ß do not dominate each other and the first maximal node dominating α also dominates ß.

They follow May (1985) in assuming that IPs and VPs are typical adjunction sites for QR at LF.[4]

Let us see how the SP works using the examples from English and Japanese presented in the previous chapter. They are repeated below:

(6) Someone loves everyone.

(7) a. Dareka -ga daremo -o aisiteiru (koto)
someone-NOM everyone-ACC love fact
'Someone loves everyone.'
b. Daremo $-o_i$ dareka -ga t_i aisiteiru (koto)
everyone-ACC someone-NOM love fact
'Everyone$_i$, someone loves t_i.'

Remember that (6) and (7b) are ambiguous, while (7a) is unambiguous.

Let us assume, as is standard by now, that English has Subject Raising, i.e., subjects are generated within VP and then raised to SPEC of IP (cf. Fukui and Speas (1986), Kitagawa (1986), Kuroda (1988) among others). Let us further assume, as Aoun and Li do for Chinese, that in Japanese, this raising operation is not available.[5] We will presume with Nakayama and Koizumi (1991) that subjects in Japanese are generated outside VP or, more specifically, in SPEC of IP.

The MBR coupled with this assumption will give (6) and (7a) the LF representations in (8) and (9) respectively (in the LF representations that follow, the elements associated by a connecting line form a chain):[6,7]

(8) $[_{IP}$ someone$_i$ $[_{IP}$ x_i $[_{VP}$ everyone$_j$ $[_{VP}$ t_i loves $x_j]]]]$

(9) $[_{IP}$ dareka$_i$-ga $[_{IP}$ x_i $[_{VP}$ daremo$_j$-o $[_{VP}$ x_j aisiteiru$]]]]$

In (8) *someone* c-commands and thus may take scope over *everyone*. In addition, *everyone* c-commands a member of the chain containing *someone*, that is, the trace t_i. Therefore, *everyone* may have scope over *someone* by virtue of the SP. In (9) *dareka* 'someone' c-commands and thus takes scope over *daremo* 'everyone'. *Daremo* cannot have wide scope since it does not c-command a member of the chain containing *dareka*. The scope contrast follows.

Crucially, the MBR forbids LF representations for (7a) like the following:

(10) *[$_{IP}$ dareka$_i$-ga [$_{IP}$ daremo$_j$-o [$_{IP}$ x_i [$_{VP}$ x_j aisiteiru]]]]

In (10) the most local potential antecedent for the variable x_i is *daremo*, given Aoun and Li's assumption that QP variables, as opposed to *wh*-variables, are not constrained by Condition C of the Binding Theory and thus can be A-bound (see footnote 3). The coindexation of *daremo* with the variable x_i results in the variable x_j being A-bound by the variable x_i, but this is allowed. The variable x_i, however, is not bound by *daremo*, violating the MBR. If (10) were legitimate, it would be wrongly predicted that (7a) should be ambiguous since *daremo* c-commands the variable of *dareka* in (10).

Let us put (7b) aside for a moment. We will come back to it in Section 2.2.2.1.

2.1.2. Ernst (1991)

Ernst (1991) proposes a generalized version of Aoun and Li's (1989) SP:

(11) The Scope Principle (SP)
An operator A has scope over an operator B in case A c-commands a member of the chain containing B.

(11) is intended to cover not only QPs but also such "semantic" operators as adverbs and modals. For instance, it explains the ambiguity of examples like (12) involving raising:

(12) Every student is likely to cheat on the exam.

(12) contains the QP *every student* and the modal-like element *likely*. It is ambiguous between the reading where *every student* has wide scope and the one where *likely* has wide scope. Given the assumption that *likely* is not subject to QR and thus constitutes a one-membered chain, the LF representation for (12) is the following:

(13) [$_{IP}$ every student$_i$ [$_{IP}$ x_i is likely$_j$ [$_{IP}$ t_i to cheat on the exam]]]

(13) observes the MBR, the variable x_i being bound by the QP. In this representation, *every student* c-commands and can take scope over *likely*. At the same time, *likely* can

take wide scope since it c-commands the trace t_i within the embedded clause, a member of the chain headed by *every student*.

Ernst suggests that "c-command" in the SP should be c-command given in (5a) (see Ernst (1991)). We will follow him in this respect and strictly distinguish m-command from c-command.

2.1.3. Chains

As (11) indicates, chains play a crucial role in the SP. Before we proceed, we would like to make explicit the notion of chain whose precise definition is not provided in Aoun and Li (1989) or Ernst (1991). We will assume that a chain is defined in terms of antecedent government (cf. Chomsky (1986b), Rizzi (1990:92)):

(14) $(\alpha_1,\ldots,\alpha_n)$ is a chain if n=1 or if, for $1 \leq i < n$, α_i antecedent-governs α_{i+1}.

What (14) means is that an unmoved element counts as a one-membered chain and that a moved element and its trace or traces form a chain only if antecedent government holds between them. The definition of antecedent government is provided below (Rizzi (1990:6)):

(15) Antecedent Government: X antecedent-governs Y iff
(i) X and Y are coindexed
(ii) X c-commands Y
(iii) no barrier intervenes
(iv) Relativized Minimality is respected

We will adopt Cinque's (1990:42) definition of barrier:[8]

(16) Every maximal projection that fails to be directly selected by a category nondistinct from [+V] is a barrier for government.

Descriptively, C, I, V, and A are nondistinct from [+V]. The part "to be directly selected" in (16) can be paraphrased as "to be a complement of" in the X-bar-theoretic sense. Relativized Minimality (henceforth RM) proposed by Rizzi (1990:27) is given in (17):

(17) Relativized Minimality: X α-governs Y only if there is no Z such that
(i) Z is a base-generated position
(ii) Z is α-GT compatible with Y
(iii) Z c-commands Y and does not c-command X.

Intuitively, RM says that government of Y by X is blocked by intervening Z only if Z is a potential governor of the same kind as X for Y. The value of "α" in (17) ranges over "head", "A-antecedent", "A'-antecedent", and "head-antecedent". A descriptive statement of GT (government theory) Compatibility for antecedent government is as follows (we ignore head government):

(18) Z is compatible with Y,
a. a member of an A-chain, iff Z is an A specifier c-commanding Y.
b. a member of an A'-chain, iff Z is an A' specifier c-commanding Y.
c. a member of an X^0-chain, iff Z is an X^0 category c-commanding Y.

An argument for defining chains in terms of antecedent government comes from examples like the following (taken from Rizzi (1990:99-100)):

(19) Tell me what you think that everyone should give to Bill.

(20) a. ??Tell me what you heard rumors that everyone wanted to give to Bill.
b. ?Tell me what you wonder why everyone gave to Bill.

(20a,b) are somewhat degraded since they violate the Subjacency Condition (cf. Chomsky (1977) among others). But what concerns us here is that there is a scope contrast between (19) and (20).[9] (19) is ambiguous, whereas (20a,b) are unambiguous, only the *wh*-operator taking scope over the standard QP. This contrast can readily be accounted for by the SP which utilizes the notion of chain given in (14). The LF representation for (19) would be the following:[10]

(21) tell me $[_{CP}$ what$_i$ you think that $[_{IP}$ everyone$_j$ $[_{IP}$ x_j should t_j give x_i to Bill]]]

(21) does not violate the MBR (see footnote 3). (19) is ambiguous because in (21) *what* c-commands *everyone* and the latter in turn c-commands a member of the chain containing the former, i.e., the variable x_i. Next, consider the LF representations for

(20a,b), which comply with the MBR:

(22) a. tell me $[_{CP}$ what$_i$ you heard rumors $[_{CP}$ that $[_{IP}$ everyone$_j$ $[_{IP}$ x_j t_j wanted to give
x_i to Bill]]]]
b. tell me $[_{CP}$ what$_i$ you wonder $[_{CP}$ why $[_{IP}$ everyone$_j$ $[_{IP}$ x_j t_j gave x_i to Bill]]]]

In (22) the link between *what* and its variable is "broken" since an antecedent government relation between the two does not hold; in (22a) there is an intervening barrier, i.e., the CP selected by the noun *rumor*, a [-V] category, and in (22b) there is an intervening A'-specifier, i.e., *why*, which induces a RM effect. The reason why (20a,b) are unambiguous is evident from (22a,b); *what* c-commands *everyone*, but the latter does not c-command a member of the chain headed by the former. Therefore, examples like (19) and (20) support the claim that the definition of chain should refer to antecedent government.

Rizzi (1990), observing examples like (19) and (20), suggests that LF reconstruction seems to be possible only when antecedent government holds between an operator and its variable. Saito (1990) shows, however, that *wh*-movement cannot be "undone" since it establishes operator-variable relations. Therefore, we cannot resort to reconstruction in (19) and (20). Cinque (1990), on the other hand, argues that the contrast in question is simply due to the referential quality of the "long-distance" moved element in (20) and that the operator with this referential quality is resistant to scope interaction.[11] But notice that (19) and (20) are equally "out of the blue" or out of pragmatic context. Thus, we claim that the above purely syntactic account is superior to Cinque's account.

2.1.4. The Ban on Transitive Application

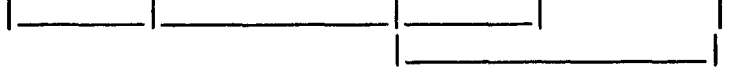
There is one more point to make about the nature of the SP which neither Aoun and Li (1989) nor Ernst (1991) are aware of.[12] It is that the SP does not tolerate

transitive applications. The issue of transitivity becomes sharp when we examine sentences involving more than two operators. Consider the raising construction in (23) taken from Chomsky (1981:177) who attributes this observation to May (1977) (">" should be read as "takes scope over"):

(23) Some senator is likely to speak at every rally.
(i) $\exists$ > likely > $\forall$
(ii) likely > $\exists$ > $\forall$
(iii) likely > $\forall$ > $\exists$

(23) is three-ways ambiguous, as indicated above. Its LF representation would be as follows:[13]

(24) $[_{IP}$ some senator$_i$ $[_{IP}$ x_i is likely$_j$ $[_{IP}$ every rally$_k$ $[_{IP}$ t_i to speak at $x_k]]]]$



(24) does not violate the MBR. If the SP applied in a transitive manner, we would predict from (24) that (23) has the six logically possible readings. In particular, we would expect that there should be interpretations on which *every rally* takes scope over *likely*; this is because *some senator* can have scope over *likely*, and *every rally* can take scope over *some senator*. Note that the absence of such interpretations cannot be attributed to some sort of semantic restriction; as (12) shows, universal quantifiers can take scope over *likely*. Therefore, examples like (23) suggest that the task of the SP is to fix the relative scope of a given pair of operators. Let us call this constraint on the SP "the Ban on Transitive Application (BTA)"[14].

2.2. Scrambling

In this section, we discuss scrambling in Japanese in relation to scope interpretation. First, the scope of a standard QP and a "*wh*-phrase" is examined from the perspective on scope outlined in the previous section. It is argued that further relativization of Rizzi's (1990) RM enables us to subsume the MBR of Aoun and Li

(1989) under the ECP. It is also argued that the relevant scope facts in Japanese can be accounted for by the SP together with LF reconstruction of scrambling (Saito (1990)).

Secondly, the scope of standard QPs in simplex, multiple, and "long distance" scrambling is considered. It is suggested that given the LF clause-boundedness condition on QPs (cf. May (1977), Hornstein (1984)) and the auxiliary assumption that reconstruction of scrambling (at least in Japanese) is an "all-at-once" operation, the scope of QPs in Japanese is amenable to the SP.

Throughout, we will assume with Saito (1985) and Hoji (1985) that scrambling in Japanese is an adjunction operation (but see Mahajan (1990) for a different view of scrambling).

2.2.1. Scope of a QP and a Wh-QP

Let us first consider the scope relation between a QP and a "*wh*-phrase". Observe the Japanese examples in (25), which are taken from Hoji (1986):[15]

(25) a. ?*Daremo -ga dare-o syootaisita no?
everyone-NOM who-ACC invited Q
'Everyone invited who?'

b. Dare-o_i daremo -ga t_i syootaisita no?
who-ACC everyone-NOM invited Q
'Who$_i$, everyone invited t_i?'

c. Dare-ga daremo -o syootaisita no?
who-NOM everyone-ACC invited Q
'Who invited everyone?'

d. Daremo -o_i dare-ga t_i syootaisita no?
everyone-ACC who-NOM invited Q
'Everyone$_i$, who invited t_i?'

Hoji's (1985, 1986) generalizations about these kinds of examples are given in (26):

(26) a. *QP-ga [$_{VP}$ WH-o V] Q
-NOM -ACC

b. WH$_i$-o QP-ga [$_{VP}$ t_i V] Q (WH taking wide scope)

c. WH-ga [$_{VP}$ QP-o V] Q (WH taking wide scope)

d. QP$_i$-o WH-ga [$_{VP}$ t_i V] Q (WH taking wide scope)

The configuration in (26a) is ungrammatical. Those in (26b-d) are grammatical but unambiguous, the "*wh*-phrase" taking wide scope over the standard QP. The D-structure configuration in which a QP c-commands a "*wh*-phrase" leads to ungrammaticality unless the "*wh*-phrase" scrambles over the QP.

Here we will make the two assumptions about Japanese in (27):

(27) (i) "*wh*-phrases" are syntactically QPs and are subject to QR (Kim (1989))[16]
(ii) "*wh*-phrases" must be licensed by the feature [+Q] (= +Quantificational) in COMP under "government" at LF (cf. Nishigauchi (1990))[17]

From now on, following Kim (1989), we will distinguish "*wh*-phrases" in Japanese from those in languages like English by calling the former *wh*-QPs. The word "*wh*-element" will be used as a cover term for *wh*-phrases and *wh*-QPs.

Let us first consider the possible (but ill-formed) schematic LF representations for (26a) in (28):

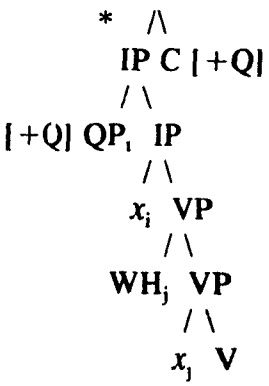
(28) a. *$[_{IP}$ WH-o_j $[_{IP}$ QP-ga_i $[_{IP}$ x_i $[_{VP}$ x_j V]]]] Q
b. *$[_{IP}$ QP-ga_i $[_{IP}$ WH-o_j $[_{IP}$ x_i $[_{VP}$ x_j V]]]] Q
c. *$[_{IP}$ QP-ga_i $[_{IP}$ x_i $[_{VP}$ WH-o_j $[_{VP}$ x_j V]]]] Q

(28a,b) violate the MBR of Aoun and Li (1989).[18] In (28a) QP-*ga* does not bind the variable x_j though the former qualifies as the most local potential antecedent for the latter. In (28b) the variable x_i is not bound by its most local potential antecedent WH-*o*. However, nothing seems to rule out (28c), given the assumptions made so far. Thus, constraining LF representations by the MBR is in fact insufficient. What we would need is a more general constraint which prohibits not only (28a,b) but also (28c).

It is certainly desirable to derive the MBR from some principle of grammar. As a matter of fact, Aoun and Li (1989) suggest two possibilities; one is to derive the MBR from RM proposed by Rizzi (1990), the other is to integrate the MBR into the Binding Theory. In what follows, we will argue for the first option.

Let us look at why (28c) is ill-formed. Its structure is given in (29):

(29)



We would like to suggest that in structures like (29), "government" of the *wh*-QP by the [+Q] in COMP is blocked by the intervening [+Q] the standard QP contains.[19] This is a situation where a minimality requirement is imposed on "government" by a licensing feature. Following Chomsky (1986a), let us assume the Principle of Full Interpretation:

(30) Principle of Full Interpretation (FI)
Every element at PF and LF must receive an appropriate interpretation or must be licensed.

Applied to the case at hand, FI requires that *wh*-QPs be licensed or "governed" by [+Q] in COMP at LF ((27ii)). (28c) is then ultimately excluded as a violation of FI since the *wh*-QP is unlicensed due to the presence of the QP with the feature [+Q].

Note that scope relations among QPs are (partially) determined by antecedent government (see (11) and (14)). Then, it would be natural to think that antecedent government also plays a role in locality in QR. It would be theoretically desirable that the same principle, RM in particular, constrains scope phenomena in general. Note also that the ill-formedness of (28c) cannot be explained by an approach which utilizes the notion of binding. As mentioned above, the MBR cannot rule out (28c). But RM together with FI can.

Therefore, we advocate the idea to derive the MBR from RM, ultimately from the ECP. We will adopt the "conjunctive" ECP in (31) (see (15) for antecedent government):[20]

(31) Empty Category Principle: A nonpronominal empty category must be
(i) properly head-governed
(ii) antecedent-governed.

For present purposes, the clause (i) of (31) is orthogonal (see Cinque (1990), Rizzi (1990)). We will assume without discussion that argument traces and X^0 traces can satisfy the clause (ii) of the ECP at any level of representation (S-structure or LF) (cf. Chomsky (1989)).[21] We will assume further, following Lasnik and Saito (1984, 1992), that adjunct traces can satisfy the clause (ii) of the ECP (can be positively Γ-marked in their terminology) only at LF.

In order to subsume the MBR under RM, the latter must be made sensitive to adjoined positions created by QR as well as base-generated positions. Thus, we propose to eliminate (17i), and RM will be as in (32):[22]

(32) Relativized Minimality: X α-governs Y only if there is no Z such that
(i) Z is α-GT compatible with X
(ii) Z c-commands Y and does not c-command X.

We also depart from Rizzi (1990) in changing Y in the clause (17ii) to X, as in (32i), for a theory-internal reason that will become clear below (see Section 2.4.). This is tantamount to saying that GT-compatibility is the relation between a governor and a potential governor and need not to refer to governees. GT-compatibility can loosely be characterized in the following way (adapted from Rizzi (1990)):[23]

(33) Government Theory Compatibility:
An element Z is α-GT compatible with X only if the conditions for the appropriate subcase of government are met (the relevant property of the governor) while the substantive condition on Z as a governor need not be (e.g. actual coindexation for antecedent government).

(33) differs from Rizzi's original statement in that it makes no reference to syntactic positions of governors (and, as we suggested above, governees). Instead, it makes use of the property of governors. For instance, the relevant property of operators, we tentatively suggest, is operatorhood.

Given (32), (33), and the assumption that a *wh*-QP and a standard QP are GT-compatible with each other (see (27i)), the ill-formedness of (26a) can be accounted for: variables must be antecedent-governed by the closest potential governor (cases like (28a,b)); "government" by a licensing feature is blocked by an intervening feature of the same kind (cases like (28b,c)).[24]

Now, let us turn to (26b-d). It may seem that in (26b) the S-structure trace is not antecedent-governed because of the intervening QP. But that is not the case. Notice that neither the QP nor the scrambled *wh*-QP has yet acquired operatorhood at S-structure: Saito (1990) shows that scrambling in Japanese does not establish operator-variable relations. Therefore, S-structures like (26b) are in effect "invisible" to RM. Given Saito's (1990) claim, it is reasonable to assume that scrambled QP's as well as non-scrambled ones are subject to QR. (34) is a well-formed LF representation for (26b):[25]

(34) $[_{IP}$ WH$_i$-o $[_{IP}$ x_i $[_{IP}$ QP-ga$_j$ $[_{IP}$ x_j $[_{VP}$ t_i V]]]]] Q

|_____| |_____|

We have already seen cases ((20a,b)) where chain formation fails at S-structure. A natural consequence of our concept of chain (see (14)) is that a chain can get "broken" also at LF. We suggest in fact that in (34) the link between the variable x_i and the trace t_i gets broken at LF since antecedent government does not hold between the two due to the presence of the intervening variable x_j. This is an instance where syntactic variables enter into RM. There is nothing wrong with the trace t_i because it has already satisfied the ECP at S-structure. According to the SP, (34) represents the interpretation on which the *wh*-QP takes wide scope, a desirable result.

The LF representation for (26c) would be the following:

(35) $[_{IP}$ WH-ga$_i$ $[_{IP}$ x_i $[_{VP}$ QP-o$_j$ $[_{VP}$ x_j V]]]] Q

|_____| |_____|

(35) in which the *wh*-QP has scope over the QP correctly represents the interpretation of (26c). LF representations other than (35) are excluded by the ECP.

In (26d) the S-structure trace, like that in (26a), satisfies the ECP. The (ill-formed) LF representations for (26d) are as follows:

(36) a. *$[_{IP}$ WH-ga$_i$ $[_{IP}$ QP$_j$-o $[_{IP}$ x_j $[_{IP}$ x_i $[_{VP}$ t_j V]]]]] Q
b. *$[_{IP}$ QP$_i$-o $[_{IP}$ WH-ga$_j$ $[_{IP}$ x_i $[_{IP}$ x_j $[_{VP}$ t_i V]]]]] Q
c. *$[_{IP}$ QP$_i$-o $[_{IP}$ x_i $[_{IP}$ WH-ga$_j$ $[_{IP}$ x_j $[_{VP}$ t_i V]]]]] Q

(36a,b) are cases of an ECP violation; in (36a) antecedent government of the variable x_i by the *wh*-QP is disrupted by the QP, whereas in (36b) x_i cannot be antecedent-governed by the QP due to the presence of the *wh*-QP. (36c) is not allowed because the QP bars "government" of the *wh*-QP by the [+Q] in COMP.

Then, why is (26d) grammatical at all? Let us assume with Saito (1990) that scrambling in Japanese can be undone at LF.[26] Given this assumption, (26d) has the alternative LF representation in (35); the scrambled QP first moves back to its D-structure position and then adjoins to VP by QR. (35) represents the correct interpretation of (26d). Notice that when reconstruction takes place in (26b), the resultant LF representations are all ill-formed, as in (28).[27]

2.2.2. Scope of QPs

2.2.2.1. Simplex Cases

We saw above that scrambled sentences in Japanese like (37) are ambiguous:

(37) (=7b) Daremo -o$_i$ dareka -ga t_i aisiteiru (koto)
everyone-ACC someone-NOM love fact
'Everyone$_i$, someone loves t_i.'

The question is: Can the analysis developed in the preceding subsection handle cases like (37)? The answer is yes. The LF representation for unreconstructed (37) would be the following:

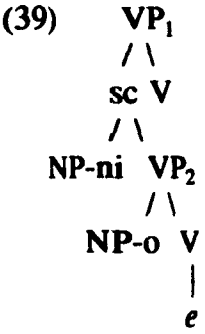
(38) $[_{IP}$ daremo-o$_i$ $[_{IP}$ x_i $[_{IP}$ dareka$_j$-ga $[_{IP}$ x_j $[_{VP}$ t_i aisiteiru]]]]]
|_____| |_____|

(38), in which there is no link between the variable x_i and the trace t_i, represents the

interpretation on which *daremo* takes scope over *dareka*. When the scrambled QP is put back to its original position, the LF representation for (37) will be identical to (9) in which *dareka* takes wide scope. Therefore, (37) is ambiguous.[28]

2.2.2.2. Multiple Scrambling

Next, let us consider multiple scrambling involving QPs. We will assume following Hoji (1985) that in Japanese the indirect object asymmetrically c-commands the direct object at D-structure and that the Japanese dative construction has the pseudo-Larsonian VP-structure in (39) proposed by Aoun and Li (1989) for the double object construction (*e* denotes an empty verb):[29]

(39)


The relevant examples are given in (40)-(42):

(40) a. [$_{IP}$ Junko-ga [$_{VP1}$ dareka -ni [$_{VP2}$ daremo -o *e*] shookaisita]] (koto)
-NOM someone-DAT everyone-ACC introduced fact
'Junko introduced everyone to someone.'
b. Dareka-ni$_i$ daremo-o$_j$ [$_{IP}$ Junko-ga [$_{VP1}$ t_i [$_{VP2}$ t_j *e*] shookaisita]] (koto)
'To someone$_i$, everyone$_j$, Junko introduced t_j t_i.'
c. Daremo-o$_j$ dareka-ni$_i$ [$_{IP}$ Junko-ga [$_{VP1}$ t_i [$_{VP2}$ t_j *e*] shookaisita]] (koto)
'Everyone$_j$, to someone$_i$, Junko introduced t_j t_i.'

(41) a. [$_{IP}$ Dareka -ga [$_{VP1}$ daremo -ni [$_{VP2}$ Kazumi-o *e*] shookaisita]] (koto)
someone-NOM everyone-DAT -ACC introduced fact
'Someone introduced Kazumi to everyone.'
b. Daremo-ni$_i$ Kazumi-o$_j$ [$_{IP}$ dareka-ga [$_{VP1}$ t_i [$_{VP2}$ t_j *e*] shookaisita]] (koto)
'To everyone$_i$, Kazumi$_j$, someone introduced t_j t_i.'
c. Kazumi-o$_j$ daremo-ni$_i$ [$_{IP}$ dareka-ga [$_{VP1}$ t_i [$_{VP2}$ t_j *e*] shookaisita]] (koto)
'Kazumi$_j$, to everyone$_i$, someone introduced t_j t_i.'

(42) a. $[_{IP}$ Dareka -ga $[_{VP1}$ Kiwako-ni $[_{VP2}$ daremo -o $e]$ shookaisita]] (koto)
someone-NOM -DAT everyone-ACC introduced fact
'Someone introduced everyone to Kiwako.'
b. Kiwako-ni$_i$ daremo-o$_j$ $[_{IP}$ dareka-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ shookaisita]] (koto)
'To Kiwako$_i$, everyone$_j$, someone introduced t_j t_i.'
c. Daremo-o$_j$ Kiwako-ni$_i$ $[_{IP}$ dareka-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ shookaisita]] (koto)
'Everyone$_j$, to Kiwako$_i$, someone introduced t_j t_i.'

The (b) examples and the (c) examples are derived from the unambiguous (a) examples. The scrambled sentences are represented schematically in (43) with their respective interpretations:

(43) Multiple Scrambling

a. QP-ni$_i$ QP-o$_j$ $[_{IP}$ NP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ V]] (-DAT -ACC -NOM)	(unambiguous)
b. QP-o$_j$ QP-ni$_i$ $[_{IP}$ NP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ V]]	(ambiguous)
c. QP-ni$_i$ NP-o$_j$ $[_{IP}$ QP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ V]]	(ambiguous)
d. NP-o$_j$ QP-ni$_i$ $[_{IP}$ QP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ V]]	(ambiguous)
e. NP-ni$_i$ QP-o$_j$ $[_{IP}$ QP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ V]]	(ambiguous)
f. QP-o$_j$ NP-ni$_i$ $[_{IP}$ QP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j $e]$ V]]	(ambiguous)

The prediction the previous discussion makes is that a scrambled sentence is ambiguous when it involves the flip of the D-structure c-command relation between QPs. As we can see in (43), this prediction is indeed borne out, though a technical question arises as to the nature of reconstruction. Suppose that in (43a) we can choose to reconstruct only QP-*ni* at LF. Then, we would expect that (43a) should be ambiguous, having the following two LF representations:

(44) a. QP-ni$_i$ x_i QP-o$_j$ x_j $[_{IP}$ NP-ga $[_{VP1}$ t_i $[_{VP2}$ t_j V]]]
|__| |__|_____|
b. QP-o$_j$ x_j $[_{IP}$ NP-ga $[_{VP1}$ QP-ni$_i$ x_i $[_{VP2}$ t_j V]]]
|__| |__|

This is, however, not the case. To account for the unambiguity of (43a), we hypothesize that the following holds at least for Japanese-type scrambling:[30]

(45) Reconstruction of scrambling is an "all-at-once" operation.

In other words, there are only two options; either no reconstruction or reconstruction of

every scrambled element. Given (45), (44b) is not permitted, and the scope interpretations in multiple scrambling in (43) can be explained by the present analysis.

2.2.2.3. Long Distance Scrambling

Let us turn now to "long distance" scrambling involving QPs. We will assume following Saito (1985) that scrambling can take place successive-cyclically. Consider the examples (46)-(48):

(46) a. Hisayuki-ga [$_{CP}$ daremo -ga dareka -o aisiteiru to] omotteiru (koto)
-NOM everyone-NOM someone-ACC love COMP think fact
'Hisayuki thinks that everyone loves someone.'
b. Dareka-o$_i$ Hisayuki-ga [$_{CP}$ daremo-ga t_i aisiteiru to] omotteiru (koto)
'Someone$_i$, Hisayuki thinks that everyone loves t_i.'

(47) a. Daremo -ga [$_{CP}$ dareka -ga Kenichi-o aisiteiru to] omotteiru (koto)
everyone-NOM someone-NOM -ACC love COMP think fact
'Everyone thinks that someone loves Kenichi.'
b. Kenichi-o$_i$ daremo-ga [$_{CP}$ dareka-ga t_i aisiteiru to] omotteiru (koto)
'Kenichi$_i$, everyone thinks that someone loves t_i.'

(48) a. Daremo -ga [$_{CP}$ Tomohisa-ga dareka -o aisiteiru to] omotteiru (koto)
everyone-NOM -NOM someone-ACC love COMP think fact
'Everyone thinks that Tomohisa loves someone.'
b. Dareka-o$_i$ daremo-ga [$_{CP}$ Nobuhiro-ga t_i aisiteiru to] omotteiru (koto)
'Someone$_i$, everyone thinks that Tomohisa loves t_i.'

In the above examples, the (b) sentences are derived from the unambiguous (a) sentences by preposing the object of the embedded clause to the sentence initial position. Their schematic S-structure representations are provided in (49):

(49) Long Distance Scrambling
a. QP-o$_i$ [$_{IP}$ NP-ga [$_{CP}$ [$_{IP}$ QP-ga [$_{VP}$ t_i V]]] V] (ambiguous)
-ACC -NOM -NOM
b. NP-o$_i$ [$_{IP}$ QP-ga [$_{CP}$ [$_{IP}$ QP-ga [$_{VP}$ t_i V]]] V] (unambiguous)
c. QP-o$_i$ [$_{IP}$ QP-ga [$_{CP}$ [$_{IP}$ NP-ga [$_{VP}$ t_i V]]] V] (unambiguous)

(49a) is ambiguous. On the other hand, (49b,c) are unambiguous, the subject QP of the matrix clause taking wide scope. The unambiguity of (49c) indicates that there is a restriction on the scope interaction of QPs; QPs can yield ambiguity only if they are

clause mates. The scope interpretations in (49) are predicted by the account developed above together with the general clause-boundedness condition on LF representations of QPs.[31] The unambiguity of (49b) needs no explanation since no QP has scrambled over another QP. The LF representations for (49a) would be as follows:

(50) a. $[_{IP}$ NP-ga $[_{CP}$ $[_{IP}$ QP-o$_i$ $[_{IP}$ x_i $[_{IP}$ QP-ga$_j$ $[_{IP}$ x_j $[_{VP}$ t_i V]]]]]] V]
b. $[_{IP}$ NP-ga $[_{CP}$ $[_{IP}$ QP-ga$_j$ $[_{IP}$ x_j $[_{VP}$ QP-o$_i$ $[_{VP}$ x_i V]]]]] V]

Whatever principle is responsible for LF clause-boundedness of QPs, it forces the scrambled QP to move back to a position within the embedded clause. In (50a) the scrambled QP has been reconstructed to the intermediate adjoined position. In (50b), on the other hand, it has been put back to the base-generated position. In the former representation, QP-*o* takes scope over QP-*ga*, and in the latter, QP-*ga* takes scope over QP-*o*. Hence the ambiguity of (49a).

The two possible LF representations for (49c) would be the following:

(51) a. $[_{IP}$ QP-ga$_j$ $[_{IP}$ x_j $[_{CP}$ $[_{IP}$ QP-o$_i$ $[_{IP}$ x_i $[_{IP}$ NP-ga $[_{VP}$ t_i V]]]]] V]]
b. $[_{IP}$ QP-ga$_j$ $[_{IP}$ x_j $[_{CP}$ $[_{IP}$ QP-o$_i$ $[_{IP}$ NP-ga $[_{VP}$ x_i V]]]] V]]

No matter whether the scrambled QP is reconstructed to the IP-adjoined position ((51a)) or to the original position ((51b)), it will be c-commanded by the matrix subject QP. Therefore, (49c) allows only the interpretation on which QP-*ga* has wide scope.

2.3. NP-Movement

This section is concerned with a second kind of Move α, i.e., NP-movement. In particular, we examine raising, passives, unaccusatives, psych constructions, and dative constructions. It is shown that various kinds of data involving NP-movement fall under the present account.

2.3.1. Raising

We have already seen above that the ambiguity in raising cases like (6) and (12), whose LF representations are recapitulated below, can be accounted for by the SP:

(52) (=(8)) $[_{IP}$ someone$_i$ $[_{IP}$ x_i $[_{VP}$ everyone$_j$ $[_{VP}$ t_i loves $x_j]]]]$

(53) (=(13)) $[_{IP}$ every student$_i$ $[_{IP}$ x_i is likely$_j$ $[_{IP}$ t_i to cheat on the exam$]]]$

In (52) and (53), the link between the variable x_i and the trace t_i holds since antecedent government of the latter by the former is blocked neither by the QP *everyone* nor by the modal-like element *likely*, given the characterization of GT-compatibility in (33). It would be feasible that a syntactic variable on the one hand and the QP and modals on the other do not share the relevant property for the purpose of GT-compatibility.

2.3.2. Passives

Let us turn to passives, typical NP-movement constructions. The relevant examples are provided in (54) and (55) (the latter is from Japanese):

(54) Everyone was seen by someone.

(55) Daremo -ga$_i$ dareka -ni t_i nagur-are-ta.
everyone-NOM someone by hit-PASS-PAST
'Everyone was hit by someone.'

Both of these sentences are ambiguous.[32] The ambiguity of the English (54) is not surprising since its active counterpart is ambiguous. (55) is more informative since active sentences containing QPs are generally unambiguous in Japanese. Consider its LF representation below:

(56) $[_{IP}$ daremo-ga$_i$ $[_{IP}$ x_i $[_{VP}$ dareka$_j$ $[_{VP}$ $[_{PP}$ x_j-ni$]$ t_i nagurareta$]]]]$

In (56) the link between the variable x_i and the trace t_i is not disrupted because the

variable of the agentive QP x_j does not intervene between them; x_j is inside the PP (postpositional phrase) and does not c-command t_i (see (32ii)). (56) successfully represents ambiguity since the higher QP c-commands the lower QP and the latter c-commands t_i, the tail of the chain containing the former. The ambiguity of (54) can be accounted for in the same manner.

Now, consider the Japanese passive sentences in (57) which involve a *wh*-QP and a QP:[33]

(57) a. *[John ya Mary]-ga$_i$ [$_{VP}$ dare-ni t_i sikar-are-ta] no?
and -NOM who by scold-PASS-PAST Q
'John, Mary and so on were scolded by whom?'
b. Dare-ni$_j$ [John ya Mary]-ga$_i$ [$_{VP}$ t_j t_i sikar-are-ta] no?
who by and -NOM scold-PASS-PAST Q
'[By whom]$_i$, John, Mary and so on were scolded t_i?'
c. Dare-ga$_i$ [$_{VP}$ [John ya Mary]-ni t_i sikar-are-ta] no?
who-NOM and by scold-PASS-PAST Q
'Who was scolded by John, Mary and so on?'
d. [John ya Mary]-ni$_j$ dare-ga$_i$ [$_{VP}$ t_j t_i sikar-are-ta] no?
and by who-NOM scold-PASS-PAST Q
'[By John, Mary and so on]$_i$, who was scolded t_i?'

The S-structure schematic configurations of these examples are given below:

(58) a. *QP-ga$_i$ [$_{VP}$ WH-ni t_i V] Q
-NOM by
b. WH-ni$_j$ QP-ga$_i$ [$_{VP}$ t_j t_i V] Q (WH taking wide scope)
c. WH-ga$_i$ [$_{VP}$ QP-ni t_i V] Q (ambiguous)
d. QP-ni$_j$ WH-ga$_i$ [$_{VP}$ t_j t_i V] Q (ambiguous)

(58a) is ungrammatical. (58b), derived from (58a) by scrambling the *wh*-QP, is rendered grammatical but is unambiguous, the *wh*-QP taking wide scope. (58c) is ambiguous. (58d) derived from (58c) is also ambiguous.

The scope interpretations observed in (58) are exactly what is predicted by the SP, RM, and the assumption that scrambling can be undone at LF. (58a) is ruled out in the same way as (26a). Its LF representations are provided in (59):

(59) a. *$[_{IP}$ WH_j $[_{IP}$ QP-ga_i $[_{IP}$ x_i $[_{VP}$ $[_{PP}$ x_j-ni] t_i V]]]] Q
b. *$[_{IP}$ QP-ga_i $[_{IP}$ WH_j $[_{IP}$ x_i $[_{VP}$ $[_{PP}$ x_j-ni] t_i V]]]] Q
c. *$[_{IP}$ QP-ga_i $[_{IP}$ x_i $[_{VP}$ WH_j $[_{VP}$ $[_{PP}$ x_j-ni] t_i V]]]] Q

(59a,b) violate the ECP; in (59a) antecedent government by the *wh*-QP of its variable is blocked by the QP, and in (59b) antecedent government by the QP of its variable is blocked by the *wh*-QP. (59c) violates FI since the *wh*-QP cannot be "governed" by the [+Q] COMP due to the QP with [+Q]. The LF representations for (58b-d) would be (60-62) respectively:

(60) $[_{IP}$ WH_k $[_{IP}$ $[_{PP}$ x_k-ni]$_j$ $[_{IP}$ QP-ga_i $[_{IP}$ x_i $[_{VP}$ t_j t_i V]]]]] Q
|_____| |_____|_____|

(61) $[_{IP}$ WH-ga_i $[_{IP}$ x_i $[_{VP}$ QP_j $[_{VP}$ $[_{PP}$ x_j-ni] t_i V]]]] Q
| | |_____| |
|_____|_____|

(62) *$[_{IP}$ QP_k $[_{IP}$ $[_{PP}$ x_k-ni]$_j$ $[_{IP}$ WH-ga_i $[_{IP}$ x_i $[_{VP}$ t_j t_i V]]]]] Q

In (60) the *wh*-QP takes wide scope. When (58b) undergoes reconstruction, it will have the illegitimate LF representations given in (59). Thus, it is unambiguous. (58c) is ambiguous since in (61) the *wh*-QP c-commands the QP and the latter in turn c-commands the trace t_i, a member of the chain headed by the former. Although (62), where the *wh*-QP is not "governed" by the [+Q] in COMP, is ill-formed, (58d) will have the LF representation in (61) when reconstruction of the scrambled QP occurs. Thus, it is ambiguous.

2.3.3. Unaccusatives

Consider the unaccusative constructions in (63) from Japanese and (64) from Korean (the former is cited from Hoji et al. (1989), the latter from Kim (1991)):

(63) Dareka -ga_i subeteno heya -ni t_i hair-ta.
someone-NOM every room into enter-PAST
'Someone entered every room.'

(64) Sey conglyu-uy koki-ka motun yenmos-ey iss-ta.
three kinds -GEN fish-NOM all ponds -LOC be-IND
'There are three kinds of fish in all the ponds.'

It has been argued that unaccusative constructions involve NP-movement. As expected, (63) and (64) are interpreted ambiguously. (65) would be the LF representation for (63):

(65) $[_{IP}$ dareka-ga$_i$ $[_{IP}$ x_i $[_{VP}$ subeteno heya$_j$ $[_{VP}$ $[_{PP}$ x_j-ni] t_i hair-ta]]]]

As shown in this representation, the SP straightforwardly accounts for the ambiguity observed in unaccusatives.

2.3.4. Psych Constructions

Belletti and Rizzi (1988) propose an analysis of psych predicates in which the theme argument undergoes NP-movement for Case reasons. If their analysis is on the right track, we would predict that the scope ambiguity exhibited by NP-movement cases also shows up in psych constructions.[34] The following examples, (66) from Japanese and (67) from Korean, demonstrate that the prediction is borne out (Kim and Larson (1989:686)):

(66) Dareka -ga dare-ni-mo hitsuyoo-da.
someone-NOM everyone-DAT need
'Everyone needs someone.'

(67) Mwuenka-ka nwukwu-eykey-na hwuhoyslep-ta.
something-NOM everyone-DAT be regrettable-IND
'Everyone regrets something.'

(66) would receive the following LF representation:

(68) $[_{IP}$ dareka-ga$_i$ $[_{IP}$ x_i $[_{VP}$ daremo$_j$ $[_{VP}$ $[_{PP}$ x_j-ni] t_i hitsuyoo-da]]]]

It is by now clear that (68) correctly represents the two interpretations.

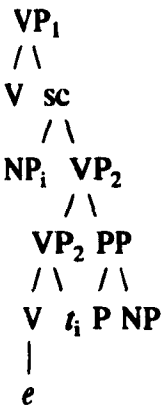
2.3.5. Dative Constructions

Aoun and Li (1989:167) note that the following dative constructions are ambiguous, contrasting with their unambiguous double object counterparts ((70) is from Chinese):[35]

(69) Mary gave some book to everyone.

(70) Wo song sanben shu gei meigeren.
I gave three book to everyone
'I gave three books to everyone.'

They claim, contra Larson (1988), that it is the dative construction, not the double object construction, that involves NP-movement. According to their analysis, the schematic S-structure of the dative construction looks like the following (again, *e* denotes an empty verb):

(71)


If (71) is correct, the ambiguity of (69) and (70) is explainable under the SP. The LF representation for (69) would be as follows:

(72) Mary gave $[_{VP1}$ some book$_i$ $[_{SC}$ x_i $[_{VP2}$ everyone$_j$ $[_{VP2}$ $[_{VP2}$ *e* $t_i]$ $[_{PP}$ to $x_j]]]]]$

The overlap of the chains in (72) successfully represents the ambiguity of (69).

2.4. *Wh*-Movement

In this section, we are concerned with the scope interaction between a *wh*-phrase

(as opposed to a *wh*-QP) and a standard QP. Simplex cases, passives, psych constructions, and double object constructions in English are examined in turn. Observing that the SP cannot cope with these cases, we modify the SP in such a way that it incorporates the notion of GT-compatibility into its formulation.

2.4.1. Wh-Arguments

May (1985) recognizes a contrast in scope interaction in pairs like (73):

(73) a. What did everyone bring?
b. Who brought everything?

(73a), with a QP in subject position and a *wh*-trace in object position, is ambiguous, whereas (73b), with a QP in object position and a *wh*-trace in subject position, is unambiguous.[36]

We assume with Saito (1990) that *wh*-movement cannot be undone at LF because it establishes operator-variable relations unlike scrambling in Japanese. The LF representations of (73a,b) would be (74a,b) respectively:[37]

(74) a. $[_{CP}$ what$_i$ did $[_{IP}$ everyone$_j$ $[_{IP}$ x_j $[_{VP}$ t_j bring $x_i]]]]$

b. $[_{CP}$ who$_i$ $[_{IP}$ x_i $[_{VP}$ everything$_j$ $[_{VP}$ t_i brought $x_j]]]]$

Let us make the reasonable assumption that a standard QP is not GT-compatible with a "true" *wh*-phrase. Then, antecedent government by *wh*-phrases like *what* and *who* as opposed to *wh*-QPs in Japanese will not be blocked by standard QPs like *everyone* and *everything*. The ambiguity of (74a) is expected by the SP; the IP-adjoined QP c-commands the variable x_i, a member of the chain headed by the *wh*-phrase. The unambiguity of (74b), however, is not expected and calls for further modification of the SP in (11).

It seems that the restriction is tighter on the interaction between a *wh*-phrase and

a standard QP than on the interaction between standard QPs (including *wh*-QPs). In particular, in order for a QP to take scope over a *wh*-phrase, all the members of the chain headed by it must be properly contained in the chain headed by the *wh*-phrase, as in (74a). An intersection of chains, as in (74b), does not suffice to yield ambiguity.

It was suggested above that RM plays a significant role in scope assignment; it directly imposes a locality condition on QR and indirectly contributes to the determination of scope via the notion of chain in the SP. Then, it is not unreasonable to think that the relevance of RM to scope assignment may be reflected in the formulation of the SP itself, though there is no *a priori* reason to believe that this should be the case. To be more specific, it is possible that the GT-(in)compatibility between two operators decides the kind of restriction on the scope interaction between them. Pursuing this possibility, we reformulate the SP as in (75), where "A $\subseteq$ B" means that A is GT-compatible with B, while "A $\not\subseteq$ B" means that A is not GT-compatible with B:

(75) The Scope Principle (SP)
An operator A has scope over an operator B in case
(i) A c-commands a member of the chain containing B (A $\subseteq$ B) or
(ii) all members of the chain containing A c-command a member of the chain containing B (A $\not\subseteq$ B).

Since a standard QP is not GT-compatible with a *wh*-phrase, its relative scope with respect to a *wh*-phrase is subject to (75ii). The scope contrast in (73) follows directly from the SP in (75).

What about the status of *wh*-phrases with respect to standard QPs in GT-compatibility? Obviously, the *wh*-phrases in (74) must be subject to (75i). Otherwise, we would incorrectly expect (74a) to be unambiguous with only the wide scope reading of the QP. Thus, let us tentatively suppose that *wh*-phrases are GT-compatible with standard QPs. This assumption may make intuitive sense in light of the fact that *wh*-phrases are often referred to as "quasi-quantifiers" (cf. Chomsky (1975) and others). What we are suggesting is that GT-(in)compatibility is not necessarily a symmetrical relation. In this

case, a *wh*-phrase is GT-compatible with a standard QP, but not vice versa.

The reason that GT-compatibility has been changed into the relation between governors (compare (17ii) and (32i)) lies in the SP in (75). Under Rizzi's GT-compatibility which concerns the relation of a potential governor to a governee, it is impossible to incorporate the notion of GT-compatibility into the SP. This is because operators can be antecedent-governors but can never be antecedent-governees at LF.

2.4.2. Wh-Adjuncts

Let us turn to cases where the *wh*-phrase is an adjunct. Consider the examples in (76) (taken from May (1985)):

(76) a. When did everyone see Max?
b. When did Max see everyone?

Both of these sentences are ambiguous. We will assume following McConnell-Ginet (1982), Larson (1988), and Stroik (1990) that V-modifying adverbs such as *when* are sister to V at D-structure.[38] Under this assumption, the LF representations for (76a,b) would be (77a,b) respectively:

(77) a. $[_{CP}$ when$_i$ did $[_{IP}$ everyone$_j$ $[_{IP}$ x_j $[_{VP1}$ t_j see $[_{VP2}$ Max $[_{V'}$ t_V $x_i]]]]]]$

b. $[_{CP}$ when$_i$ did $[_{IP}$ Max $[_{VP1}$ everyone$_j$ $[_{VP1}$ see $[_{VP2}$ x_j $[_{V'}$ t_V $x_i]]]]]]$

Given (77a,b), the SP predicts that (76a,b) should both be ambiguous, which is indeed the case. This is because *when* c-commands *everyone*, and all the members of the chain headed by *everyone* c-command the variable of *when*.

What about the case of *why*, which is often regarded as an IP-adverb? Consider the following pair from Collins (1991:38):

(78) a. Why did everybody hate John?
b. Why did John hate everybody?

Collins (1991) notes that (78a) is ambiguous while (78b) is not. Let us assume with Collins (1991) that *why* can be generated in a VP-adjoined position. The LF representations for (78a,b) would be the following (assuming the Larson-type layered VP (cf. Travis (1991)) and adjunction of *why* to the lower VP):

(79) a. $[_{CP}$ why$_i$ did $[_{IP}$ everybody$_j$ $[_{IP}$ x_j $[_{VP1}$ t_j hate $[_{VP2}$ $[_{VP2}$ John$]$ $x_i]]]]]]$

b. $[_{CP}$ why$_i$ did $[_{IP}$ everybody$_j$ $[_{IP}$ John $[_{VP1}$ hate $[_{VP2}$ $[_{VP2}$ $x_j]$ $x_i]]]]]$

The SP can take care of the scope contrast in (78), given the above representations. (78a) is ambiguous since in (79a) the *wh*-phrase c-commands the QP, and all the members of the chain containing the QP c-command the variable of the *wh*-phrase. In (78b) the QP cannot take wide scope since in (79b) the variable of the *wh*-phrase is not c-commanded by the variable of the QP.

2.4.3. Passives

Now, consider the following passive examples involving *wh*-operators (cited from Kim (1991)):

(80) a. What will be brought in by every guest?
b. By whom will everything be purchased?

These sentences are both reported to be ambiguous. We will tentatively assume that the *by*-phrase can appear in a VP-adjoined position or a sister position to V at D-structure. When the *by*-phrase is in an adjoined position, it is assumed to be licensed by predication, and when it is sister to V, it is assumed to be licensed by (adverbial) θ-role assignment, as in the case of adjuncts such as *when* (cf. Larson (1988), Stroik (1990)). If this assumption is correct, the following would be possible LF representations for (80a,b):

(81) a. $[_{CP}$ what$_i$ $[_{IP}$ x_i will $[_{VP}$ $[_{PP}$ by every guest$]_j$ $[_{VP}$ $[_{VP}$ be brought in $t_i]$ $x_j]]]]$

b. $[_{CP}$ $[_{PP}$ by whom$]_i$ will $[_{IP}$ everything$_j$ $[_{IP}$ x_j $[_{VP}$ be t_j purchased $x_i]]]]$

Given (81a,b), the ambiguity of (80a,b) is expected by the SP. We should not be fooled by the lines in (81a). In (81a), despite its appearance, all the members of the chain containing the pied-piped QP c-command the NP-trace of *what*. In (81b), all the members of the chain headed by the QP c-command the variable of *by whom*.

2.4.4. Psych Constructions

The SP in (75) can be extended to the pairs involving psych verbs in (82) taken from Kim and Larson (1989:682):[39]

(82) a. Who does everyone excite?
b. Who excites everyone?

The examples in (82) contrast with those in (73) in that it is the (b) sentence that is ambiguous; the (a) sentence is unambiguous.[40] Given Belletti and Rizzi's (1988) analysis of psych predicates mentioned above, the LF representations for (82a,b) would be as in (83a,b) respectively:

(83) a. $[_{CP}$ who$_i$ does $[_{IP}$ everyone$_j$ $[_{IP}$ x_j $[_{VP}$ excite $[_{V'}$ $t_j]$ $x_i]]]]$

b. $[_{CP}$ who$_i$ $[_{IP}$ x_i $[_{VP}$ everyone$_j$ $[_{VP}$ excites $[_{V'}$ $t_i]$ $x_j]]]]$

Here again, we should not be visually deceived by the chains drawn with the lines. It is important to bear in mind that the variable of the experiencer argument (x_i in (83a) and x_j in (83b)) is hierarchically higher than the NP-trace of the theme argument (t_j in (83a) and t_i in (83b)). (83b) is ambiguous since all the members of the chain containing *everyone* c-command t_i, whereas (83a) is unambiguous since x_i, the variable of *who*, is

not c-commanded by t_j, the tail of the chain containing *everything*.

Tacit in the representations in (83) is the assumption that a variable of a standard QP and a variable of a *wh*-operator are not GT-compatible with each other. Suppose that variables were GT-compatible with each other regardless of the nature of their operators. Then, we would expect the reverse of the actual judgements for (82a,b). This is because in (83a) antecedent government of t_j by x_j would be blocked by x_i, which would allow all the members of the chain (*everyone*, x_j) to c-command x_i, the tail of the chain headed by *who*, and in (83b) antecedent government of t_i by x_i would be blocked by x_j, which would leave no overlap of the chains.

2.4.5. Dative Constructions

Finally, let us examine dative constructions from the viewpoint of the SP in (75). Relevant examples are given below ((84a) is taken from Kim (1991), who attributes it to Robert May, and (84b) is adapted from Kim (1991)):

(84) a. What did John give to everyone.
b. To whom did John give everything?

The (a) sentence is ambiguous, while the (b) sentence is unambiguous. Assuming the dative structure in (71), their respective LF representations would be the following:

(85) a. $[_{CP}$ what$_i$ did $[_{IP}$ $[_{PP}$ to everyone$]_j$ $[_{IP}$ John $[_{VP1}$ give $[_{SC}$ x_i $[_{VP2}$ $[_{VP2}$ t_V $t_i]$ $x_j]]]]]]$

b. $[_{CP}$ $[_{PP}$ to whom$]_i$ did $[_{IP}$ everything$_j$ $[_{IP}$ Alex $[_{VP1}$ give $[_{SC}$ x_j $[_{VP2}$ $[_{VP2}$ t_V $t_j]$ $x_i]$

With (85a,b), the SP can explain the interpretative contrast in (84a,b). (84a) is ambiguous since in (85a) all the members of the chain headed by *to everyone* c-command t_i, the trace of *what*. (84b), on the other hand, is unambiguous since in (85b) the trace of *everything* t_j fails to c-command x_i, the variable of *to whom*.

2.5. Summary

In this chapter, we have considered the relation between Move α and the scope of operators. Based on scope facts regarding Japanese scrambling, we have claimed that the MBR follows from a modified version of RM and thus ultimately from the ECP. The preceding examination of scope interaction between a *wh*-phrase and a standard QP has prompted us to refine the SP in such a way that it refers to GT-compatibility in its formulation. It has been shown that the SP can account for a wide range of data involving scrambling, NP-movement, and *wh*-movement.

The central claim of this chapter is that RM (GT-compatibility, to be exact) plays a key role in fixing scope of operators. In particular, there are three ways in which GT-compatibility contributes to scope assignment. First, it determines the kind of restriction on interaction of given two operators through the SP. Secondly, it imposes a locality condition on (LF) movement of operators through the ECP. Thirdly, it decides whether or not a given link in a chain holds, which is crucial information for the purpose of the SP.

FOOTNOTES TO CHAPTER 2

1. The Scope Principle of May (1985) will not be considered here. Various kinds of problems with it have already been pointed out by a number of authors (Aoun and Li (1989), Ernst (1991), Kim (1991), to name a few). For instance, it cannot account for the cross-linguistic scope difference between (6) and (7a) below.

2. α binds ß if α c-commands ß and α is coindexed with ß.

3. This definition is meant to make the MBR apply to variables bound by *wh*-operators. For example:

(i) what$_i$ did everyone$_j$ x_j buy x_i

Aoun and Li assume that variables coindexed with *wh*-operators are R-expressions. Under this assumption, in (i) *everyone* is not a potential antecedent for x_i since the coindexation of the latter with the former will result in a violation of Condition C of the Binding Theory, i.e., x_i will be A-bound by x_j.

As for variables coindexed with standard QPs, Aoun and Li assume that they are not subject to Condition C.

4. May (1985) also suggests the possibility of NP-adjunction by QR. We will assume that QR can freely adjoin QPs to any maximal projection.

5. Aoun and Li suggest that the lack of Subject Raising in Chinese may be traced back to the degenerate nature of INFL or the ability of SPEC of IP and VP to be assigned Case and a Θ-role. We do not commit ourselves to their suggestion in this regard. See Kim (1991) for criticism. It is beyond the scope of this thesis to explore the question of what the source of the parametric variation in question is.

6. In the representations or structures to follow, irrelevant details will be omitted throughout.

7. Saito (1990) shows that an A'-position cannot be equated with an operator position. Partly following Taraldsen (1986), let us adopt the definition of a variable in (i):

(i) X is a variable iff X is locally bound by an operator at LF.

8. We will assume that adjunction to non-barriers does not create barriers.

9. The same kind of contrast is also observed in Italian (see Cinque (1990:11-14)).

10. The Principle of Full Interpretation, which requires that no superfluous elements be present at LF and PF, forces intermediate *wh*-traces, if any, to be deleted at LF (see Chomsky (1986a, 1989)).

11. This (marginal) referential quality of *what* is supposed to allow (20a,b) to avoid violating the "conjunctive" ECP. See footnote 20.

12. We are indebted to Lisa Travis (p.c.) for bringing our attention to this matter.

13. QR of the standard QP over operators like *likely* is prohibited by the ECP (see Chapter 4).

14. A strict interpretation of the SP would of course imply the BTA. It is not clear at the moment from what principle of grammar the BTA can be derived. It might prove to be a piece of evidence for the autonomy of linguistic knowledge; in our logical thinking, deduction with the use of transitivity is a quite common activity.

15. For discussion of the ungrammaticality of (25a) and similar examples, see Hoji (1986).

16. In Japanese, standard QPs are derived from "*wh*-phrases" in a systematic way. Existential QPs and universal QPs are obtained by attaching *ka* and *mo* to "*wh*-phrases" respectively (the latter can also be used as negative polarity items). For example:

(i) a. dare 'who' --- **dareka** 'someone', **daremo** 'everyone'
b. nani 'what' --- **nanika** 'something', **nanimo** 'everything'
c. doko 'where' --- **dokoka** 'somewhere', **dokomo** 'everywhere'

Kim (1989) claims that "*wh*-phrases" in Korean are standard QPs, too. As long as his data are concerned, Korean patterns fully with Japanese in terms of scope interpretation.

17. We depart from Nishigauchi (1990) in assuming (i) for "government" in (27ii):

(i) Feature Government: X feature-governs Y iff
(i) X m-commands Y
(ii) Relativized Minimality is respected

We will assume further that *wh*-QPs can take scope over the c-command domain of the [+Q] COMP which licenses them (see Chapter 3).

It is worth mentioning that in Japanese, it is [+Q] COMP that determines quantificational force of "*wh*-phrases". Consider the following:

(ii) a. Dare-ga kuru no?
who-NOM come Q
'Who will come?'
b. Dare-ga kite mo, boku-wa kinisi-nai.
who-NOM come Q I -TOP care-NEG
'No matter who comes, I don't care.'

If a "*wh*-phrase" is "governed" by *no* or *ka*, as in (iia), it acquires existential force. If it is "governed" by *mo*, as in (iib), it acquires universal force. As seems reasonable, we identify the Q-morphemes *ka* and *mo* with the morphemes *ka* and *mo* in QPs respectively (see the preceding footnote).

18. Assuming that variables of *wh*-QPs are not subject to Condition C (see footnote 3 and (27i)).

19. See footnotes 16 and 17.

20. Traces of head movement are exempt from (31i), and those of *wh*-elements that are D-linked (discourse-linked) in the sense of Pesetsky (1987) are exempt from (31ii). The latter can meet the "identification" requirement through binding. See Cinque (1990, ch.1).

21. Lasnik and Saito (1984:263) present the "principle" in (i) which is intended to explain among other things *that*-trace effects:

(i) Only an argument receives a Γ-feature at S-structure.

See Rizzi (1990) for an alternative account of *that*-trace effects within the conjunctive ECP.

22. One may well ask what motivated Rizzi to include (17i) in his RM. We will argue later in Chapter 4 that there is in fact no motivation for limiting RM inducers to elements in base-generated positions.

23. We will elaborate on the notion of GT-compatibility in Chapter 4 where we argue that it should be defined in terms of a set of features. For the time being, we will leave it as vague as (33).

24. Thus, (28b) doubly violates the ECP and FI. The same remark applies to (36b) and (59b) below. One could get rid of this unwanted redundancy by assuming that the ECP is simply part of FI (see Fanselow (1991) for a proposal along this line).

25. Since QR can adjoin QPs to any maximal projection (footnote 4), QP-*ga* is allowed to adjoin to the IP immediately dominating it.

26. This operation leaves no trace behind, as required by FI.

27. Hoji (1985:248) presents the following condition to explain data such as those in (7) and (25):

(i) at LF *QP_i QP_j t_j t_i where each member c-commands the member to its right

This condition, however, is unsatisfactory not only conceptually but also empirically. Within his framework based on May (1977) and Lasnik and Saito (1984), multiple scrambling such as (40b) and long distance scrambling such as (46b) and (48b) are problematic for (i) (see Hoji (1985:299, fn.25)).

28. Instances of VP-internal scrambling receive the same explanation.

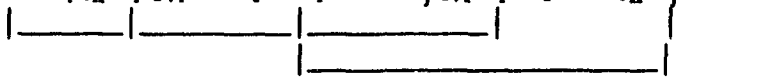
(i) a. NP-ga $[_{VP}$ QP-o$_i$ $[_{VP}$ QP-ni t_i V]] (ambiguous)
-NOM -ACC -DAT
b. QP-ga $[_{VP}$ NP-o$_i$ $[_{VP}$ QP-ni t_i V]] (unambiguous)
c. QP-ga $[_{VP}$ QP-o$_i$ $[_{VP}$ NP-ni t_i V]] (unambiguous)

29. We will assume that VP_2 is not a barrier since scrambling of the direct object is fully grammatical.

30. Alternatively, one might suggest that in (43a,c, and e), what has been scrambled is the small clause (Japanese allows CP scrambling).

31. Strictly speaking, the precise characterization of the condition in question cannot be based on clause-boundedness. The ambiguity of (i) shows that the LF representation in (ii) should be available and thus in certain cases QR is not clause-bound (Aoun and Hornstein (1985:624)):

(i) Someone expects every Republican to be elected.
(ii) $[_{IP}$ someone$_i$ $[_{IP}$ x_i $[_{VP}$ every Republican$_j$ $[_{VP}$ t_i expects $[_{IP}$ x_j to be elected]]]]]



See Hornstein (1984) and Aoun and Hornstein (1985) for a binding approach to the relevant phenomena. We suspect that consideration of tense may prove crucial.

32. Hoji et al. (1989, fn.12) remark that "the Japanese passive does not seem to induce scope ambiguity between the preposed subject and the agentive NP, unlike Chinese". Thus for them, (55) should be unambiguous. But we find it ambiguous. They discuss the following passive sentence:

(i) Dareka -ga$_i$ subeteno hito -ni t_i syookais-rare-ta.
someone-NOM every person-DAT introduce-PASS-PAST
'Someone was introduced to everyone.'

We agree that (i) is ambiguous and that if we take *subeteno hito* 'everyone' as an agentive NP (*-ni* can be used either as the dative marker or 'by'), (i) is not ambiguous, *dareka* 'someone' taking scope over *subeteno hito*. We suggest that the unavailability of the wide scope interpretation of the agentive NP can be attributed to the fact that agentive NPs in Japanese strongly favor group readings (for reasons unknown to us). Observe the ambiguous (ii):

(ii) Dareka -ga$_i$ subeteno hito -ni sorezore t_i syookais-rare-ta.
someone-NOM every person by each introduce-PASS-PAST
'Someone was introduced by everyone each.'

In (ii) *sorezore* 'each' is added to guarantee the non-group reading on *subeteno hito* (this technique is due to Hoji (1985)).

33. Considering the property of the agentive phrase mentioned in the preceding note, we avoid using *daremo* 'everyone'. We assume that the conjunctive morpheme *ya* 'and so on', like *ka* and *mo*, has the feature [+Q].

34. Notice, however, that English psych constructions like (i) would be expected to be unambiguous under the present account ((ii) is the LF representation for (i)):

(i) Something is worrying everyone.
(ii) $[_{IP}$ something$_i$ $[_{IP}$ x_i is $[_{VP}$ everyone$_j$ $[_{VP}$ $[_{V'}$ worrying $t_i]$ $x_j]]]]$

This is because antecedent government of the trace t_i by the variable x_i is blocked by the variable x_j at LF. Contrary to the prediction, (i) is judged to be ambiguous. We have no explanation for (i) at the moment.

35. Compare (69) and (70) with (ia) and (ib) respectively:

(i) a. Mary gave someone every book.
b. Wo song sange ren meiben shu.
I gave three man every book
'I gave three men every book.'

The SP coupled with the double object structure in (39) accounts for the lack of ambiguity of (ia,b).

36. Spanish exhibits the same scope contrast as the one in (73). See Jaeggli (1987).

37. Recall that we are assuming Cinque's (1990) barrier system, not Chomsky's (1986b). Hence, VP-adjunction of *wh*-operators is unnecessary.

38. Stroik (1990:658) observes the following scope contrast:

(i) a. Who did John see everywhere (he went)?
b. Where did John (last) see everyone?

(ia) is unambiguous, whereas (ib) is ambiguous. The contrast falls under the SP in (75).

39. In Nakamura (to appear, fn.24), we mistakenly remarked that our account there based on Kim (1991) could cover psych constructions like (82a).

40. The same "reversed" judgements about scope as those in (82) obtain in Spanish psych-constructions as well. See Jaeggli (1987).

CHAPTER 3

FEATURE GOVERNMENT

3.0. Introduction

This chapter deals with linguistic entities which are dependent on licensing features. Specifically, in addressing the question in (1) raised at the outset of this thesis, we examine *wh*-elements, negative polarity items, and adverbs:

(1) What is the mechanism governing licensing of (certain) operators?

In Section 3.1., the *Wh*-Criterion and the Neg-Criterion proposed as an answer to (1) in the case of *wh*-elements and negative elements are introduced. In Section 3.2., licensing of *wh*-elements, negative polarity items, and adverbs is discussed. It is suggested that the *Wh*-Criterion and the Neg-Criterion are too tight to capture the relevant cross-linguistic variations in licensing of *wh*-elements and negative polarity items. As an alternative to these criteria, we present what we call the Feature-Dependent Item Criterion which exploits the new concept of feature government. Feature government is simply defined in terms of m-command and Relativized Minimality. We also present the Feature Government Parameter which determines the degree of locality in feature government required in a given language for a given item. It is argued that distributional and interpretative characteristics of adverbs (in English) can also be accounted for by the Feature-Dependent Item Criterion and the Relativized Minimality.

Thus, our answer to (1) is the Feature-Dependent Item Criterion, which makes crucial use of Relativized Minimality. Relativized Minimality plays an important role not only in fixing the relative scope of operators (see Chapter 2) but also in licensing of operators.

3.1. The *Wh*-Criterion and The Neg-Criterion

In this section, we will briefly consider the *Wh*-Criterion of Rizzi (1991) and the Neg-Criterion of Haegeman (1991) which are claimed to be the well-formedness conditions on *wh*-elements and negative elements respectively.

3.1.1. The Wh-Criterion

Wh-movement has attracted considerable attention in the course of development of linguistic theory (cf. Chomsky (1977) and many others). Putting aside the issue of locality,[1] the most notable questions surrounding syntactic *wh*-movement are: (i) What triggers it? and (ii) Why do some languages have it, while some do not?

Concerning the second question, the prevailing view since Huang (1981, 1982) has been that the presence or the absence of overt *wh*-movement can be ascribed to a parameter which is sensitive to the level of representation. In other words, *wh*-movement applies at S-structure in languages like English, while it applies at LF in languages like Chinese in which all *wh*-elements appear in-situ at S-structure (cf. Lasnik and Saito (1984)).

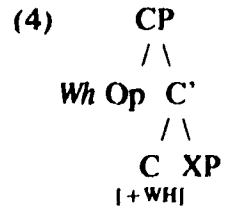
What about the first question? Along the lines of Huang (1981, 1982), Rizzi (1991), based on May (1985), formulates the following general well-formedness condition on *wh*-structures, which he takes to be universal:

(2) The *Wh*-Criterion
   a. A *Wh*-operator must be in a Spec-head configuration with an $X^0_{[+WH]}$.
   b. An $X^0_{[+WH]}$ must be in a Spec-head configuration with a *Wh*-operator.

The definition of a *wh*-operator for the purpose of (2) is (3):

(3) *wh*-operator = a *wh*-phrase in a scope position.

where a scope position is a left-peripheral A'-position (either a Spec or an adjoined position). (2) requires that in typical cases, the following configuration hold in any language at LF (linear order irrelevant):[2]

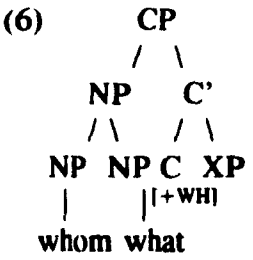
(4)


Thus, according to Rizzi (1991), it is the *Wh*-Criterion that triggers *wh*-movement.

The *Wh*-Criterion together with (3) accounts for the familiar contrast in (5) under the assumption that it applies as early as at S-structure in English:

(5) a. *Tom gave what to whom?
b. What did Tom give to whom?

(5a) is ruled out since neither of the clauses in (2) is fulfilled at S-structure. (5b) is well-formed since *what* and the [+WH] COMP are in a Spec-head relation at S-structure, satisfying the *Wh*-Criterion, and *whom* in-situ, which does not qualify as a *wh*-operator at S-structure under (3), trivially fulfills the *Wh*-Criterion. On Rizzi's account, the relevant portion of the LF structure for (5b) would be as follows:

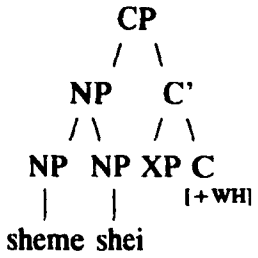
(6)


The *wh*-in-situ *whom* moves to SPEC of CP at LF to satisfy the clause (2a).

In languages like Chinese, the *Wh*-Criterion is supposed to be satisfied only at LF. Thus, the Chinese sentence in (7) from Huang (1981) is assumed to have the LF structure in (8):

(7) shei mai-le sheme?
who bought what
'Who bought what?'

(8)



3.1.2. The Neg-Criterion

Haegeman (1991), extending Rizzi's (1991) analysis of *wh*-movement to negative polarity items (henceforth NPIs), argues for the Neg-Criterion in (9):[3]

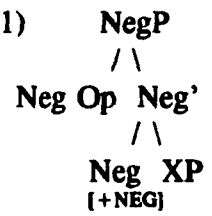
(9) The Neg-Criterion
a. A Neg-operator must be in a Spec-head configuration with an $X^0_{[+NEG]}$.
b. An $X^0_{[+NEG]}$ must be in a Spec-head configuration with a Neg-operator.

By analogy with (3), a neg-operator is characterized as follows:

(10) neg-operator = a neg-phrase in a scope position.

(9) requires that the following configuration hold at LF (again linear order irrelevant):

(11)

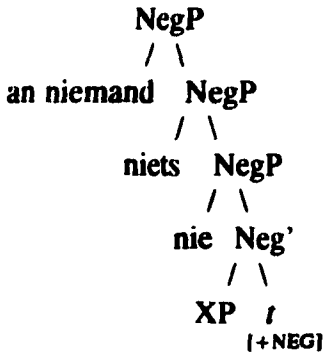


As in the case of the above-mentioned "Move *wh* parameter", the relevant parameter for the Neg-Criterion is supposed to determine at which level of representation the criterion must be satisfied; in some languages, (9) must be satisfied as early as at S-structure, while in some languages, it can be satisfied as late as at LF. Let us consider the following example from West Flemish (taken from Haegeman (1991)):

(12) da Valère an niemand niets nie gezeid en-eet
that to no one nothing not said has
'that Valère did not tell anyone anything.'

West Flemish has a clitic-like negative element *en* which attaches to finite verbs. This

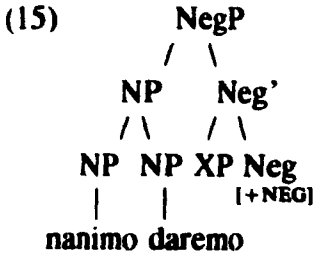
element cannot express sentential negation on its own, and the presence of other negative elements such as *nie* 'not', *niemand* 'no one', and *niets* 'nothing' is required. West Flemish, as in standard French, exhibits negative concord. (12) is a case of negative concord where *niemand* and *niets* have been scrambled from their base-generated positions to the left of *nie* and are construed as NPIs. Under Haegeman's analysis, (12) has the following configuration at S-structure:

(13)


where the head of NegP *en* has moved to Agr. The movement of NPIs in examples like (12) receives a straightforward account if we assume that (i) the Neg head, when it moves, leaves the negative feature behind, (ii) the relation of *an niemand/niets* and the [+NEG] is a variation of Spec-head agreement, and (iii) the Neg-Criterion applies at S-structure in West Flemish.[4]

The Neg-Criterion requires that in-situ NPIs raise at LF to enter into a Spec-head relation with a negative X^0 category (or its trace). Then, the example in (14) from Japanese would have (15) as its partial LF structure (assuming the existence of NegP):[5]

(14) Daremo nanimo iwa-nakat-ta.
anyone anything say-NEG-PAST
'No one said anything.'

(15)


3.2. Licensing of Feature-Dependent Items

This section discusses *wh*-elements, NPIs, and adverbs. It has been noted in the literature that these elements are somehow dependent on their licensers. For instance, it is clear from examples like the Japanese (16a,b) that *wh*-elements must be licensed by the relevant feature in COMP:

(16) a. Yoshinori-ga dare-o sikarimasita ka?
-NOM who-ACC scolded Q
'Yoshinori scolded who?'
b. *Yoshinori-ga dare-o sikarimasita?
-NOM who-ACC scolded
'Yoshinori scolded who?'

In interrogative sentences in Japanese, a [+Q] COMP is spelled out morphologically as either *ka* or *no*. The lack of a Q-morpheme leads to ungrammaticality, as in (16b).[6]

(17), taken from Linebarger (1987:326), shows that NPIs like *any* must co-occur with a negative element:

(17) a. George didn't eat any breakfast today.
b. *George ate any breakfast today.

(17b), where there is no negative element, is ungrammatical.

But notice that the presence of an appropriate licenser is merely a necessary condition, not a sufficient condition on licensing of *wh*-elements or NPIs. Consider (18a,b) from Japanese:

(18) a. Dare-ga Mikiko-ni [Satoshi-ga deisuisita to] hanasimasita ka?
who-NOM -DAT -NOM got plastered COMP told Q
'Who told Mikiko that Satoshi got plastered?'
b. *Dare-ga Mikiko-ni [Satoshi-ga deisuisita ka] hanasimasita?
who-NOM -DAT -NOM got plastered Q told
'Who told Mikiko whether Satoshi got plastered?'

(18b) is ungrammatical despite the presence of *ka* in the embedded clause. On the intuitive level, the licenser *ka* must be "high" enough in the tree to license a *wh*-QP. The same holds in the case of licensing of NPIs in English, as is illustrated by the following

pair:

(19) a. Richard did not see anyone.
b. *Anyone did not see Richard.

Licensing of adverbs is not as transparent as that of *wh*-elements or NPIs in that there appears to be no overt licenser in this case. But we know from licensing of *wh*-phrases in English that the superficial absence of a licenser does not necessarily mean the absence of a licensing mechanism. In this light, consider (20a,b) cited from Jackendoff (1972:50):

(20) a. Stanley completely ate his Wheaties.
b. *Completely Stanley ate his Wheaties.

The parallel of (20) with (19) is straightforward. It is reasonable to assume that in (20b), the licenser of the adverb *completely*, though invisible, is not "high" enough, as in the case of (19b).

In what follows, we will consider how *wh*-elements, NPIs, and adverbs are syntactically licensed. Let us use "Feature-Dependent Item (FDI)" as a cover term for these entities.

3.2.1. Licensing of Wh-Elements

In this subsection, we will focus on *wh*-elements. First, it is pointed out that the *Wh*-Criterion cannot be a universal condition, given the arguments of Kim (1989) and Aoun and Li (to appear) among others. As its alternative, we put forth the FDI-Criterion, which is based on feature government. Feature government is a combination of m-command and Relativized Minimality. The FDI-Criterion among other things explains certain ungrammatical cases involving a QP and a *wh*-QP and apparent Subjacency effects in Japanese.

Secondly, considering *wh*-movement in Slavic languages, we present the Feature Government Parameter which concerns the degree of locality in feature government, i.e.,

Spec-head agreement or government. Then, following Pesetsky (1987) and Aoun and Li (to appear), we suggest that *wh*-movement in languages like English is triggered by a morphological requirement on [+WH] COMP. A condition is invoked to cope with languages such as Italian and Irish which permit only one *wh*-phrase per clause.

3.2.1.1. Feature Government

In the treatment of *wh*-QPs in Japanese in Chapter 2, we assumed with Kim (1989) that they are syntactically QPs in nature and are subject to QR at LF. If this assumption is on the right track, it immediately casts doubt on the alleged universality of the *Wh*-Criterion in (2). In particular, a question arises as to how *wh*-QPs in Japanese are licensed in sentences like the following:

(21) Yasuhiro-wa [$_{CP}$ dare-ga nani-o hatsumeisita to] itta no?
-TOP who-NOM what-ACC invented COMP said Q
'Who did Yasuhiro say invented what?'

Given the general clause-boundedness of QR (cf. May (1977), Aoun et al. (1981), Hornstein (1984)), the LF representation for (21) would be as follows:

(22) Yasuhiro-wa [$_{CP}$ [$_{IP}$ dare-ga$_i$ [$_{IP}$ x_i [$_{VP}$ nani-o$_j$ [$_{VP}$ x_j hatsumeisita to]]]]] itta no

It is clear from (22) that even at LF, Japanese *wh*-QPs are not in Spec-head relations with a [+WH] COMP (or a [+Q] COMP) which is marked by a Q-morpheme.

Recently, Aoun and Li (to appear) have argued that *wh*'s-in-situ in English and Chinese need not move to SPEC of CP and stay in-situ at LF. If their claim is correct, it undermines the *Wh*-Criterion in a major way. Aoun and Li base their argument on the interaction of *only* and a *wh*-in-situ. Following Tancredi (1990), they assume the following generalization:

(23) Principle of Lexical Association (PLA)
An operator like *only* must be associated with a lexical constituent in its c-command domain.

Crucially, (23) applies at LF (see Aoun and Li (to appear) for justification). Observe the

contrast between (24) and (25):

(24) a. *Mary, he only likes.
b. *Who does he only like?

(25) Who only likes what?

(24a,b) show that the postverbal object cannot escape the c-command domain of *only*, as predicted by the PLA. (25) constitutes a piece of evidence that *wh*'s-in-situ do not raise to SPEC of CP at LF; *what* in (25) can be properly associated with *only*. Notice, however, that examples like (25) do not rule out the possibility of *wh*'s-in-situ undergoing QR. Mahajan (1990), along the lines of Fiengo et al. (1988), proposes that *wh*'s-in-situ, even those in English, are subject to QR. We depart from Mahajan (1990) in allowing *wh*'s-in-situ to be adjoined not only to IP but other maximal projections such as VP. We will assume that *wh*'s-in-situ have two options at LF; they can remain where they are generated or they can raise by the application of QR.[7]

Therefore, there is good reason to believe that, contrary to the *Wh*-Criterion, *wh*-elements do not necessarily enter into a Spec-head agreement relation with a [+WH] COMP. Then, how are *wh*-elements licensed? It appears that we need to somehow loosen the strong locality requirement imposed by the *Wh*-Criterion.

We would like to suggest that the *Wh*-Criterion represents the most strict instance of a much laxer condition on licensing of *wh*-elements. In particular, in light of the null hypothesis that all FDIs obey the same licensing mechanism, we propose (26) which is meant to cover not only *wh*-elements but also other FDIs such as NPIs and adverbs:

(26) The FDI-Criterion
A feature-dependent item (FDI) must be feature-governed by its licensing feature at LF.

(26) is assumed to be universal. We propose the definition of feature government in (27):

(27) Feature Government: X feature-governs Y iff
(i) X m-commands Y
(ii) Relativized Minimality is respected

(27) differs from the definition of antecedent government in three respects; (i) it does not require coindexation of X and Y,[8] (ii) it uses m-command rather than c-command, and (iii) it is insensitive to barriers. The definitions of m-command and RM are recapitulated below for convenience:

(28) α m-commands ß if α and ß do not dominate each other and the first maximal node dominating α also dominates ß.

(29) Relativized Minimality: X α-governs Y only if there is no Z such that
(i) Z is α-GT compatible with X
(ii) Z c-commands Y and does not c-command X.

It should be obvious that the Spec-head agreement is simply a subcase of feature government.

The FDI-Criterion virtually takes the forms in (30) when the FDI in question is a *wh*-element (in the languages under consideration here):

(30) a. *Wh*-phrases must be feature-governed by [+WH] in COMP at LF.
b. *Wh*-QPs must be feature-governed by [+Q] in COMP at LF.

The FDI-Criterion straightforwardly accounts for the contrast in examples like (16) and (18). The (a) examples are fine since the *wh*-QP is feature-governed by the matrix [+Q] in COMP at LF. In (16b) there is no licensing feature in the first place. (18b) has no well-formed LF representation; the *wh*-QP, if it stays in-situ, will not be feature-governed by the [+Q] in the embedded COMP, and if it lowers to be feature-governed by the [+Q], its trace will inevitably violate the ECP.

Recall from Chapter 2 that (27ii) has already been motivated by Japanese examples like (31) involving a *wh*-QP and a QP:

(31) a. ?*Daremo -ga dare-o syootaisita no?
everyone-NOM who-ACC invited Q
'Everyone invited who?'
b. Dare-ga daremo -o syootaisita no?
who-NOM everyone-ACC invited Q
'Who invited everyone?'

The schematic LF representations for (31a,b) are provided below:

(32) a. *$[_{IP}$ WH-o$_j$ $[_{IP}$ QP-ga$_{[+Q]i}$ $[_{IP}$ x_i $[_{VP}$ x_j V]]]] $Q_{[+Q]}$
b. *$[_{IP}$ QP-ga$_{[+Q]i}$ $[_{IP}$ WH-o$_j$ $[_{IP}$ x_i $[_{VP}$ x_j V]]]] $Q_{[+Q]}$
c. *$[_{IP}$ QP-ga$_{[+Q]i}$ $[_{IP}$ x_i $[_{VP}$ WH-o$_j$ $[_{VP}$ x_j V]]]] $Q_{[+Q]}$

(33) $[_{IP}$ WH-ga$_i$ $[_{IP}$ x_i $[_{VP}$ QP-o$_{[+Q]j}$ $[_{VP}$ x_j V]]]] $Q_{[+Q]}$

(32a,b) are ruled out by the ECP in conjunction with RM. The ill-formedness of (32b,c) is attributed to the lack of feature government of the *wh*-QP by the [+Q] in COMP; the [+Q] that the standard QP contains blocks the required feature government. Thus, a violation of FI ensues. (33) is well-formed since it satisfies the ECP and FI; the *wh*-QP is feature-governed by the [+Q] in COMP.

A question we must ask about examples like (21) and (25) is: How can the *wh*-elements take the scope over the matrix clause? Here we follow Baker (1970a), Riemsdijk and Williams (1981), Pesetsky (1987), Aoun and Li (to appear) among others and suggest that *wh*-elements get coindexed with a [+WH]/[+Q] COMP under feature government. This is a way to implement the intuition that a Q-morpheme, overt or covert, functions as a scope marker for *wh*-elements. We will make the natural assumption that NPIs and adverbs employ the same coindexation mechanism. Under this coindexation, FDIs are interpreted as if they were in the positions of their licensers. The LF representations for (21) and (25) then would be as follows:

(34) $[_{CP}$ Yasuhiro-wa $[_{CP}$ $[_{IP}$ dare-ga$_i$ $[_{IP}$ x_i $[_{VP}$ nani-o$_j$ $[_{VP}$ x_j hatsumeisita to]]]] itta no$_{i,j}$]

(35) $[_{CP}$ who$_i$ C$_{i,j}$ $[_{IP}$ only likes what$_j$]]

This approach to scope assignment of *wh*-elements predicts that there should be minimality effects in cases involving more than one COMP bearing the feature [+WH]/[+Q]. Consider the following minimal pair of Japanese sentences:

(36) a. Osamu-wa Akiko-ni $[_{CP}$ dare-ga nani-o sita to] hanasimasita ka?
-TOP -DAT who-NOM what-ACC did COMP told Q
'Who did Osamu tell Akiko did what?'
b. Osamu-wa Akiko-ni $[_{CP}$ dare-ga nani-o sita ka] hanasimasita ka?
-TOP -DAT who-NOM what-ACC did Q told Q
'Does Osamu tell Akiko who did what?'

(36b) differs from (36a) in that it contains *ka* in the embedded COMP instead of *to*. In (36a) the *wh*-QPs takes scope over the matrix clause, as expected. Thus, (36a) is interpreted as a multiple *wh* question. In (36b), on the other hand, the *wh*-QPs can only be construed as having scope over the embedded clause. (36b) is a yes-no question.[9] Observing examples like (36b), Nishigauchi (1990) claims that Japanese exhibits *wh*-island effects of Subjacency at LF. On the present account, the apparent Subjacency effect in (36b) is simply due to RM; the *wh*-QPs can be feature-governed by the [+Q] in the embedded COMP but not by the [+Q] in the matrix COMP at LF since feature government by the latter is blocked by the former.

Turning to cases involving more than one [+WH] COMP, consider the following sentence:

(37) Who remembers where John bought what?

Baker (1970a) notes that examples like (37) can be understood in one of two ways. On one interpretation, only *who* takes the matrix scope and *where* and *what* are associated with the embedded clause. On the other interpretation, *what* is paired with *who*, yielding a multiple *wh* question. Under the assumptions made so far, one may reasonably expect the second interpretation to be unavailable because feature government by the matrix [+WH] of the *wh*-in-situ should be disrupted by the intervening [+WH] in the embedded COMP. Do examples like (37) constitute counterexamples to the above account? We will defer answering this question until Chapter 4.

An immediate welcome consequence of the above analysis is that it explains the well-known difference (in English and other languages) between *wh*'s-in-situ and syntactically moved *wh*-phrases, i.e., the former but not the latter are free from various sorts of island violations. The following examples taken from Fiengo et al. (1988) illustrate the point:

(38) a. *Who do you like books that criticize?
b. Who likes books that criticize who?

(39) a. *Who do you think that pictures of are on sale?
b. Who thinks that pictures of who are on sale?

(40) a. *Who did you get jealous because I spoke to?
b. Who got jealous because I spoke to who?

(38)-(40) pertain to the Complex NP Constraint (CNPC), the Subject Condition, and the Adjunct Condition respectively. The lack of an ECP violation in the (b) examples follows automatically from the assumption that *wh*'s-in-situ in English do not have to obey the *Wh*-Criterion. In other words, the in-situ *wh*-phrases in the (b) examples do not cross the barriers (the relative clause in (38), the subject NP in (39), and the adjunct clause in (40)) at LF. The (b) examples show in turn that feature government can "penetrate" into barriers, supporting the formulation of feature government in (27).[10]

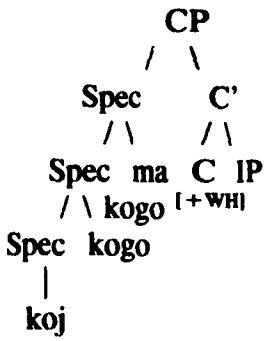
3.2.1.2. The Feature Government Parameter

In the preceding subsection, we motivated the notion of feature government. It has been noted in the literature that languages do not neatly split into two types with regard to *wh*-movement, those which have overt *wh*-movement to SPEC of CP and those which have *wh*'s-in-situ. A natural question that arises is: Does feature government have an advantage over other approaches in dealing with cross-linguistic variations in *wh*-movement?

Let us start with Bulgarian-type languages. Rudin (1988) argues convincingly that in Bulgarian and Rumanian, all *wh*-elements in a given sentence obligatorily move to SPEC of CP whose head is marked with [+WH] at S-structure. Thus, example (41) from Bulgarian, where *wh*-phrases are multiply preposed, has the partial S-structure representation in (42):

(41) Koj kogo ma kogo e pokazal?
who whom to whom has pointed out
'Who pointed out whom to whom?'

(42)



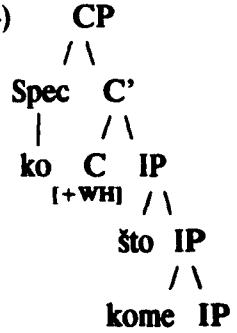
Bulgarian and Rumanian belong to a group of languages which imposes the most tight restriction on licensing of *wh*-elements, i.e., Spec-head agreement. These are languages which satisfy the *Wh*-Criterion at S-structure.

Next, let us turn to Polish-type languages. Rudin (1988) distinguishes Serbo-Croatian, Polish, and Czech from Bulgarian-type languages in terms of multiple *wh*-fronting. Consider the example in (43) from Serbo-Croatian:

(43) Ko je što kome dao?
who has what to whom given
'Who gave what to whom?'

Rudin claims that (43) has the partial S-structure representation in (44) despite its surface similarity with (41):

(44)



In (44) only one *wh*-phrase *ko* 'who' is in SPEC of CP and the rest of the *wh*-phrases are adjoined to IP.[11] We will assume with Adams (1984) and Rudin (1988) that there is a condition of the following sort ((45) is taken from Rudin (1988:490)):

(45) Condition on SpecCP Adjunction (CSA)
*[$_{\text{SpecCP}}$ α SpecCP] (nothing may be adjoined to SpecCP)

Furthermore, we will assume following Rudin (1988) that Polish-type languages obey (45), while Bulgarian-type languages do not.[12]

What would be the parameter that differentiates Bulgarian-type languages from Polish-type languages? An obvious difference between (42) and (44) is the degree of locality; in (42) Spec-head agreement is required, whereas in (44) what seems to be required is less tight relation, i.e., government. Let us adopt the following definition of government:[13]

(46) α governs ß iff α m-commands ß and there is no Γ, Γ a maximal projection, such that Γ excludes α and includes ß.

Exclusion and inclusion are defined as follows ((47) is from Chomsky (1986b:9)):

(47) α excludes ß if no segment of α dominates ß.

(48) α includes ß if every segment of α dominates ß.

Given the definition of government in (46), *što* 'what' and *kome* 'to whom' (as well as *ko* 'who') in (44) are governed by the [+WH] COMP. The relevant parameter, which we call the feature government parameter, may be formulated as follows:

(49) The Feature Government Parameter
   a. An FDI must be in a Spec-head agreement relation with its licensing feature at S-structure.
   b. An FDI must be governed by its licensing feature at S-structure.

Bulgarian-type languages choose the setting (49a) and Polish-type languages the setting (49b) when the FDI in question is a *wh*-element.[14] Notice that both Spec-head agreement and government are subcases of feature government. Thus, our approach which employs feature government can naturally express the cross-linguistic variation in *wh*-movement in terms of the degree of locality. Notice also that Spec-head agreement and government are core notions in the theoretical framework adopted here. It would be plausible that these notions are utilized in certain parameters.

One may well ask if the representations in (42) and (44) violate any locality principle, in particular, the ECP. Given the assumptions made so far, in (42) RM applies

at S-structure since the *wh*-phrases are in SPEC of CP, i.e., an operator position, and acquire operatorhood at S-structure (see Chapter 2). Technically, the representation in (42) does not violate the ECP. The traces of the *wh*-phrases satisfy the ECP at S-structure (and LF) since the [+WH] COMP, which gets coindexed with the *wh*-phrases as a result of Spec-head agreement or feature government, antecedent-governs them.

How about (44)? One of the *wh*-phrases *ko* 'who' is in SPEC of CP and thus bears operatorhood. Thus, (44) is assumed to be visible to RM. In this case, *što* 'what' and *kome* 'to whom' clearly intervene between *ko* 'who' and its trace. As a consequence, antecedent-government by *ko* of its trace is blocked by *što* and *kome*. Then, why is (43) allowed? We will return to this question when we consider superiority in Chapter 4.

The parameter in (49) actually permits a third kind of setting (or non-setting), i.e., no S-structure requirement. We can assume that a group of *wh*-in-situ languages such as Chinese, Japanese, and Korean have such a setting.[15] A question arises as to which setting is chosen by languages like English.

One may speculate that English has the setting (49a) but at the same time obeys the condition in (45). This speculation, however, is not on the right track within the present analysis. Recall that the parameter in (49) is concerned with S-structure representations. Once the value (49a) is chosen, all *wh*-phrases in a sentence must end up being in SPEC of CP at S-structure. Thus, we must conclude that English does not specify a value for the parameter. If this is the case, the question is: What drives *wh*-movement in English at all?

Following the suggestions made by Pesetsky (1987) and Aoun and Li (to appear), we will assume that what forces *wh*-movement in languages like English is a morphological requirement on [+WH] COMP. As opposed to [+Q] COMP in a language such as Japanese which is filled by an independent Q-morpheme, [+WH] COMP lacks an overt morpheme. Therefore, it would be natural to assume that [+WH] COMP must be identified or, in Pesetsky's (1987) term, cliticized to by a *wh*-phrase. In other words,

[+WH] COMP is inert on its own and must "get activated" by the appropriate feature contained in a *wh*-phrase. This amounts to saying that it is [+WH] COMP that must be licensed by a *wh*-phrase first; only after that will it be able to license *wh*-phrases. Since this process supposedly involves a licensing feature in a *wh*-phrase, it is reasonable to think that feature government is responsible for it. One way to execute this idea would be to present the following:

(50) [+WH] COMP Identification
A [+WH] COMP must be feature-governed by a *wh*-element.

(50) ensures that a *wh*-element appears in SPEC of CP whose head is marked with [+WH]. In English (50) applies as early as at S-structure. English does not allow SPEC of CP to be filled by more than one *wh*-phrase since it conforms to (45).

One may wonder if satisfaction of (50) can be postponed until LF in some languages. If it can, we would expect that such languages have LF movement of *wh*'s-in-situ to SPEC of CP and that they exhibit ECP effects in constructions where barriers are involved or RM is relevant. In this light, consider the following examples from Iraqi Arabic (taken from Wahba (1991)):

(51) a. *Mona nasat $[_{CP}$ li-meno$_i$ $[_{IP}$ tinti sheno $t_i]]$?
forgot to-whom to-give what
'What did Mona forget to whom to give?'
b. *sh-nasat Mona $[_{CP}$ li-meno$_i$ $[_{IP}$ tinti sheno $t_i]]$?
QP forgot to-whom to-give what
'What did Mona forget to whom to give?'

(51a,b) are both ungrammatical under the intended interpretation (though (51a) is grammatical under the interpretation where both of the *wh*-elements take scope over the embedded clause). The sole difference between (51a) and (51b) is that in the latter example, there is a question particle *sh-* (glossed as QP) which defines the scope of *wh*'s-in-situ.[16] The LF representation for (51a) would be the following in which *sheno* 'what' has been forced by (50) to move to the matrix SPEC of CP:

(52) *[$_{CP}$ sheno$_j$ [$_{IP}$ Mona nasat [$_{CP}$ li-meno$_i$ [$_{IP}$ tinti x_j x_i]]]]

This representation is excluded by RM or the ECP; antecedent government of the variable x_j by *sheno* 'what' is blocked by the potential governor *li-meno* 'to whom'.

How about (51b)? FI requires that there be no superfluous elements at LF. Thus, the question particle in Iraqi Arabic must be deleted and "replaced" by a *wh*-element at LF since it has no semantic content. Then LF movement of the *wh*-in-situ into SPEC of CP in the matrix clause is obligatory in (51b). As a consequence, the LF representation for (51b) is identical to the ill-formed (52) except the subject-verb order in the matrix clause.

We suggest that (50) can be fulfilled either at S-structure or at LF.[17] In examples like (51a,b) from Iraqi Arabic, (50) must be satisfied at LF.

Note that Q-morphemes in languages like Japanese cannot be deleted at LF since they bear semantic significance; it is the Q-morpheme that provides *wh*-QPs with quantificational force.[18] Thus, there is no need for a *wh*-QP to move to SPEC of CP at LF. *Wh*-phrases moved to SPEC of CP in languages like English cannot be deleted since they obviously have semantic content.

Before leaving this section, let us mention another cross-linguistic difference regarding licensing of *wh*-elements. In languages like Italian and Irish, only one instance of a *wh*-phrase is allowed per clause, i.e., multiple interrogations with a *wh*-in-situ are ungrammatical, as the Italian (53) from Adams (1984:2) and the Irish (54) from McCloskey (1979:71) illustrate:

(53) *Mi domando chi ha incontrato chi.
myself ask-1s who has met who(m)
'I wonder who met who.'

(54) *Cé aL bhí ag caint le cé?
who COMP was at talking to who
'Who was talking to who(m)?'

The intuitive idea we would like to pursue is that in these languages the feature [+WH]

in COMP is "used up" once it licenses one *wh*-phrase. The following condition, which we take to be essentially lexical, may be invoked:

(55) Condition on FDI Licensing Feature
A licensing feature can license a single instance of FDI.

As far as *wh*-elements are concerned, Italian and Irish obey (55), while English, Japanese, etc. do not.

3.2.2. Licensing of Negative Polarity Items

Let us now turn to licensing of NPIs. If NPIs are subject to the same licensing mechanism as other FDIs, we would predict that their licensing is dictated by the FDI-Criterion. In particular, we would expect that (i) the Neg-Criterion is too strong a requirement, (ii) minimality effects can be observed, (iii) licensing of NPIs is in principle insensitive to barriers, (iv) the Feature Government Parameter is relevant. We suggest that these expectations are in fact borne out. Thus, the FDI-Criterion and feature government gain further support.

We will assume with Progovac (1988) that "negation" is the only licenser of NPIs (see below). We will assume further that licensers of NPIs contain the feature [+NEG]. Given these assumptions, the FDI-Criterion requires the following:

(56) A NPI must be feature-governed by its licensing feature [+NEG] at LF.

Let us consider licensing of NPIs in English from the viewpoint of (56). Observe the subject-object asymmetry in (57) taken from Linebarger (1987:328):

(57) a. John didn't invite any students.
b. *Any students weren't invited by John.

Here a few remarks on *not* and auxiliary verbs in English are in order. We will assume following Ernst (1992) that *not* is an adverb which can appear in SPEC of VP headed by an auxiliary verb or can be adjoined to VP (cf. Baker (1991)). We will also assume with Ernst (1992) that *not* does not raise at LF and that all finite auxiliaries, including modals,

raise over negation to INFL at S-structure.[19] We will presume, as is common, that contraction is a PF rule and thus has no influence on LF. The LF representations for (57) would be as follows:[20]

(58) a. $[_{IP}$ John did $[_{VP}$ not$_{[+NEG]}$ $[_{VP}$ invite any students$]]]$
b. *$[_{IP}$ any students were $[_{VP}$ not$_{[+NEG]}$ $[_{VP}$ invited by John$]]]$

What is noteworthy is that in (58a) the NPI cannot be in a Spec-head relation with the adverb *not* at LF; *not* is supposed to be either in SPEC of VP or in a VP-adjoined position. This shows that English does not obey the Neg-Criterion even at LF and that we have to reject the Neg-Criterion as a universal condition.

(56) straightforwardly accounts for the contrast in (57). In (58a) the NPI is m-commanded and feature-governed by [+NEG], as is required. On the other hand, in (58b) the NPI is not feature-governed by [+NEG]; the latter does not m-command the former. The examples in (19) receive the same explanation.

In English, NPIs can be licensed not only by a clausemate negation, as in (57a), but also by expressions like an antecedent of conditional, as in (59) from Linebarger (1987:328):

(59) If you steal any food, they'll arrest you.

In addition, a NPI in English can be licensed by a superordinate negation, as shown in (60) (cited from Progovac (1988:74)):

(60) a. Mary did not say that anyone left.
b. Mary did not say that Peter had seen anyone.

A question arises as to what licenses the NPI in sentences without overt negation. Progovac (1988) claims that there is a negative null operator in SPEC of CP in cases like (59). She relates the ungrammaticality of (62) to that of (61):

(61) ?*Who must have killed Yuri?

(62) *If John must know the answer, he is lucky.

She assumes following McDowell (1987) that epistemic modals move to SPEC of CP at

LF. If this is correct, (61) would be excluded by the condition in (45) (assuming of course that (45) applies also at LF in English); SPEC of CP is already occupied by a *wh*-phrase at S-structure. (62) would be ruled out in the same way under the assumption that a negative operator is situated in SPEC of CP at S-structure.

But there is a problem with this account. As Progovac (1988:185) herself notices, examples like the following remain unexplained:

(63) a. How did anyone fix that car?
b. Which of you has ever been to Italy?

Progovac's account wrongly expects these examples to be ungrammatical just like (61) and (62). This is because a null negative operator must co-occur with the *wh*-phrase in SPEC of CP in (63), violating the condition in (45).

Contra Progovac, we will assume with Laka (1991) that it is COMP that qualifies as a licenser of NPIs; COMP containing [+NEG] functions as a licenser in sentences like (59). Thus, we will posit a negative COMP in the environments where Progovac postulates a null negative operator in SPEC of CP. Under this assumption, (63a,b) are fine since they do not violate (45). As for the ungrammatical examples (61) and (62), we will present a RM account of them in Chapter 4.

The well-formedness of (59) and (60a,b) is consistent with (56). Their LF representations would be the following:

(64) $[_{CP}$ if$_{[+NEG]}$ $[_{IP}$ you steal any food]], they'll arrest you

(65) a. $[_{IP}$ Mary did $[_{VP}$ not$_{[+NEG]}$ $[_{VP}$ say $[_{CP}$ that anyone left]]]]
b. $[_{IP}$ Mary did $[_{VP}$ not$_{[+NEG]}$ $[_{VP}$ say $[_{CP}$ that Peter had seen anyone]]]]

In these representations, [+NEG] feature-governs the NPIs. In examples (63a,b), the NPI is licensed by the m-commanding COMP in the same way as that in (59); COMP in interrogative sentences contains [+NEG] (see Progovac (1988)).

If the present approach to NPIs is on the right track, we would expect that RM effects should be observed in cases where there are more than one instance of the feature

[+NEG]. A relevant example would be the following (from Baker (1970b)):

(66) Nobody didn't see anyone.

(66) involves two potential NPI licensers, i.e., *nobody* and *not*. Our analysis predicts that in (66) the NPI is licensed only by [+NEG] of *not*; it cannot be feature-governed by *nobody* due to the presence of intervening [+NEG] causing a RM effect. (67a,b) represent the interpretation on which the NPI is associated with *not* and the one on which it is associated with *nobody* respectively:

(67) a. NOT $\exists$ x, x a person, NOT $\exists$ y, y a person, [x saw y]
'Everyone saw someone.'
b. NOT $\exists$ x, x a person, $\exists$ y, y a person, NOT [x saw y]
'Everyone saw everyone.'

In actuality, (66) is interpreted as meaning (67a) but not (67b). Therefore, it is *not* that licenses the NPI in (66), and the prediction at hand is borne out by examples like (66).[21]

The above formulation of feature government predicts that the feature [+NEG] can in principle "penetrate" into barriers. The following examples confirm the prediction ((68) is from Ladusaw (1980) and (69) from Linebarger (1987:337)):[22]

(68) Waldo didn't report the possibility that anybody might leave.
(69) He didn't move because anyone pushed him.

(68) is a case of CNPC and (69) is a case of Adjunct Condition. The appositive clause and the adjunct clause form a barrier in (68) and (69) respectively. These kinds of examples support the concept of feature government.

A natural question that arises in connection with the Feature Government Parameter is: Is the parameter relevant to NPIs? If it is, we would expect the existence of languages in which NPIs must be in a Spec-head configuration with [+NEG] at S-structure and those in which NPIs must be governed by [+NEG] at S-structure. The first class of languages do seem to exist. We have already seen in (12) that West Flemish

belongs to this class (Hungarian appears to belong to this class (see Haegeman (1991))).

What about the second class? Unfortunately, we do not know of any language that arguably belongs to the second class at the moment. But if there exits such a language, it would lend further support for positing the Feature Government Parameter.

It is important to notice the setting of the Feature Government Parameter for one kind of FDIs can be different from that for another kind of FDIs within the same language. Let us take West Flemish as an example. As we saw above, NPIs in West Flemish are subject to (49a). *Wh*-elements in this language, however, are not, as the following examples show (Haegeman (1991)):

(70) a. Van wien is Valère ketent *t*?
of whom is pleased
'Who is Valère pleased with?'
b. Wien is-ter ketent van wien?
who is there pleased of whom
'Who is pleased with whom?'

(70a) shows that West Flemish has overt *wh*-movement. (70b) shows that not all *wh*-elements have to enter into a Spec-head relation with a [+WH] COMP at S-structure, i.e., *wh*'s-in-situ are allowed. Thus, when it comes to *wh*-phrases, West Flemish has neither the setting (49a) nor (49b). One can claim that overt *wh*-movement in this language, like English, is triggered by the COMP Identification requirement in (50).

A similar remark can be made about the condition in (55). It was mentioned above that a [+WH] COMP in Italian can license only one *wh*-element. Apparently, the condition (55) does not extend to licensing of NPIs in Italian. Observe the following example (from Rizzi (1982:175, fn.13)):

(71) Non pretendo che nessuno dica niente.
NEG (I) pretend that nobody say nothing
'I do not pretend that anybody say anything.'

(71) indicates that a [+NEG] feature in Italian is capable of licensing multiple instances of NPIs. Therefore, [+NEG] in Italian does not obey (55). It is possible that some

features are subject to (55) while some are not even within the same language.[23]

We saw above that in English NPIs can be licensed by (i) clausemate negation, (ii) negative COMP, and (iii) superordinate negation. It has been frequently noticed that there are considerable variations in licensing of NPIs across languages and even in the same language. If each of the three kinds of licensers is taken to be a deciding factor for classifying NPIs, the following typology of NPIs is obtained:

(72) NPIs which are licensed
- a. only by clausemate negation
- b. only by negative COMP
- c. only by superordinate negation
- d. by clausemate negation and negative COMP
- e. by clausemate negation and superordinate negation
- f. by negative COMP and superordinate negation
- g. by clausemate negation, negative COMP, and superordinate negation

According to Progovac (1988), of the above seven logically possible types of NPIs, five (those except (b) and (c)) are actually attested (see Progovac (1988) for details and a binding approach).[24] We will not go into these variations, but simply note that the restrictions in question must be lexical in nature.

3.2.3. Licensing of Adverbs

In this section, we examine how adverbs are licensed, drawing heavily on the insights of Travis (1988). The discussion is limited to English adverbs, largely because their behavior is perhaps the most well-documented (cf. Jackendoff (1977), McConnell-Ginet (1982), Ernst (1984)), but we hope that the analysis presented below carries over to adverbs in other languages as well.

Given Travis' (1988) claim that adverbs are licensed by a set of features, it is quite natural to think that their distributions and interpretations (see below) are restricted by the same licensing mechanism that other FDIs are subject to, i.e., the FDI-Criterion. It is argued that this is indeed the case. In the course of discussion, we slightly modify

the definition of m-command. We also claim that the "intervention" in RM should be relativized according to the type of government in question. It is suggested that sequencing of adverbs can be explained by RM.

Travis (1988) discusses how licensing of adverbs is carried out and relates it to their behavior which is different from other syntactic categories. She emphasizes that a proper analysis of licensing of adverbs should serve to explain the idiosyncratic properties of adverbs, i.e., "transportability", interpretation, and sequencing. Let us consider these properties in turn.

First, the term "transportability" represents the characteristic of some adverbs which exhibit fairly free distribution. The following examples illustrate this characteristic (Travis (1988:282)):

(73) a. Cleverly/Clumsily John dropped his cup of coffee.
b. John cleverly/clumsily dropped his cup of coffee.
c. John dropped his cup of coffee cleverly/clumsily.

Secondly, some adverbs receive a different interpretation depending on the positions they occupy. Relevant examples are given below (taken from Travis (1988:285)):

(74) a. The police carelessly will arrest Fred.
b. Fred carelessly will be arrested by the police.
c. The police arrested Fred carelessly.
d. Fred was arrested carelessly by the police.

In (74a,b) the adverb *carelessly* appears in the pre-INFL position and is associated with the subject of the sentence. In (74c,d), on the other hand, it appears in the post-VP position and refers to the agent in the sentence. To put it differently, in the latter examples, it is the police that were being careless.

Thirdly, only certain sequences of different types of adverbs are permitted, as we can see in the following examples ((a) and (c) taken from Jackendoff (1972:89), (b) and (d) from Travis (1988:286)):

(75) a. Probably Max carefully was climbing the walls of the garden.
b. Max probably was carefully climbing the walls of the garden.
c. *Carefully Max probably was climbing the walls of the garden.
b. *Max carefully was probably climbing the walls of the garden.

Before going into Travis' (1988) proposals, let us review her typology of English adverbs provided below (adapted from Travis (1988)):[25]

(76) a. Type Ia: Initial, Aux (cleverly, clumsily, ...(subject-sensitive))
b. Type Ib: VP-initial, VP-final (cleverly, clumsily, ...(agent-sensitive))
c. Type IIa: Initial, Aux (quickly, slowly, ...(event-modifying))
d. Type IIb: VP-initial, VP-final (quickly, slowly, ...(process-modifying))
e. Type III: Initial, Aux (evidently, probably, unbelievably, ...)
f. Type IV: VP-initial, VP-final (completely, easily, totally, ...)

Notice that Type I and Type II adverbs are those whose interpretations vary according to their positions. Examples of Type I adverbs are already given in (74). With regard to Type II, Travis recognizes the subtle difference in meaning in the following examples (Travis (1988:292)):

(77) a. Quickly John will be arrested by the police.
b. John quickly will be arrested by the police.
c. John will be quickly arrested by the police.
d. John will be arrested quickly by the police.

In (77a,b) what the adverb *quickly* modifies is the event of the arrest; the arrest will take place right away. In (77c,d), on the other hand, the adverb modifies the process of the arrest; the manner of the arrest will be quick.

Travis (1988) proposes that adverbs are licensed by syntactic features of the X^0 categories INFL and V. Specifically, she suggests that Type I-IV adverbs are licensed by the following features (Travis (1988:299)):

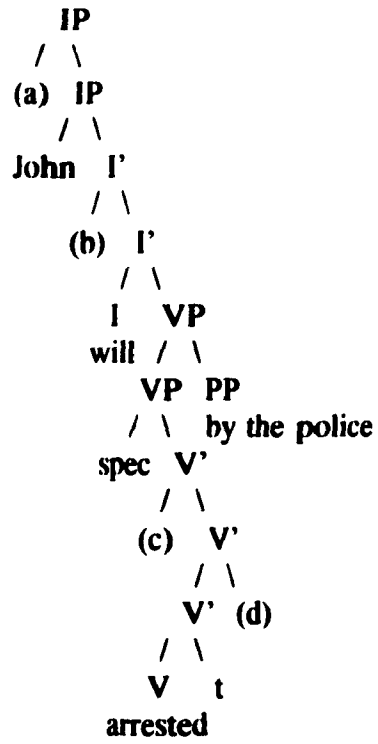
(78)

INFL	VERB
Type Ia: AGR	Type Ib: Agent
Type IIa: Event	Type IIb: Manner
Type III: Event	Type IV: Manner

Travis accounts for the above-mentioned properties of adverbs in terms of this licensing mechanism. Her answer to the first question is that it is feature percolation that is

responsible for transportability of adverbs. Feature percolation allows adverbs to be licensed anywhere within the maximal projection of the licensing head. As an illustration, consider the following tree for the examples in (77) (we ignore *be*):

(79)



The adverb *quickly* belongs to Type II. Thus, it can be licensed either by INFL or V. As we can see in (79), all the four adverb positions are located within the maximal projection of INFL or V. Under Travis' analysis, this is why (certain) adverbs can enjoy free distribution.

With regard to the second question, Travis claims that the interpretation of a given adverb is determined by the feature of its licensing head. Let us consider again the examples in (77) and the structure in (79). In (77a,b) the adverb is licensed by the feature Event in INFL and thus modifies the event. In (77c,d), on the other hand, the adverb is licensed by the feature Manner in V and thus modifies the process.

To account for the restricted relative sequencing of adverbs observed in such examples as (75), Travis makes the following three assumptions (Travis (1988:301)):

(80) a. scope is assigned by feature percolation
b. percolating features may not cross paths (but a head may contain more than one index)
c. speaker-oriented adverbs (CP scope)
S-adverbs (IP scope)
subject-oriented (INFL)
manner/agent (V)

Given (80), examples like those in (75) can be explained. By (80c), the speaker-oriented adverb *probably* must take scope over CP, and the subject-oriented adverb *carefully* must take INFL scope. In (75a,b) *probably* is hierarchically higher than *carefully*. Hence, the two paths of feature percolation created by these adverbs will not intersect, observing the condition (80b). In (75c,d), however, the reversed hierarchical relation between the two adverbs inevitably creates an overlap of the percolation lines, which results in a violation of (80b). The grammaticality judgements in (75) correctly follow.

It seems that Travis' approach to adverbs is essentially on the right track. But we would like to suggest another way of looking at the adverb facts, taking advantage of some of her insights. We will follow Travis in assuming that (certain) adverbs are licensed by INFL and/or V and that their interpretations depend on the licensing head. For concreteness, let us assign the feature [+ADV] to INFL and V. This feature represents the ability to license adverbs. We will not adopt the fine-tuned feature system in (78), but we will keep the intuition captured by it intact. If, for example, a Type II adverb happens to be licensed by [+ADV] in INFL, it becomes event-modifying. If it is licensed by [+ADV] in V, it becomes process-modifying. We will also assume following Travis (1988, p.c.) that adverbs can be either bare heads or maximal projections. The X^0 categorial status of adverbs allows them to be incorporated into another X^0 category.[26] It will be assumed that adverbs can only be incorporated into functional categories.[27] Thus, adverbs can adjoin to INFL but not to V.

But we depart from Travis in that we will not use percolation of licensing features. Instead, we suggest that the mechanism for licensing adverbs is feature

government, which has already been motivated in the above discussion of *wh*-elements and NPIs. It would be conceptually preferable to account for licensing of adverbs without recourse to an adverb-specific device. The null hypothesis in the present context is that feature government is responsible for every instance of licensing of a FDI.

Now, let us examine the relevant examples in light of feature government. The FDI-Criterion, applied to adverbs, requires the following:[28]

(81) An adverb must be feature-governed by its licensing feature [+ADV] at LF.

We will assume that in general adverbs do not raise at LF (cf. Travis (1988), Ernst (1991)).[29] Hence, (81) in effect must be met as early as at S-structure in most cases.

Let us start with Type I and Type II adverbs which behave similarly in terms of transportability and interpretation. Consider (77) and (79) regarding a Type II adverb once again. The adverb positions (b) and (c,d) are m-commanded and thus feature-governed by INFL and V respectively, satisfying (81). Under the definition of m-command we have been assuming (see (28) in 3.2.1.1.), however, INFL does not m-command the position (a). To accommodate this case into the FDI-Criterion, we revise the definition of m-command slightly as in (82) (see (47) and (48) for exclusion and inclusion respectively):[30]

(82) α m-commands ß if α and ß do not dominate each other and there is no Γ, Γ a maximal projection, such that Γ includes α and excludes ß.

Given this definition, the position (a) in (79) is m-commanded and feature-governed by INFL. In (77a,b) the adverb *quickly* is licensed by INFL and construed as event-modifying. In (77c,d) it is licensed by V and construed as process-modifying.

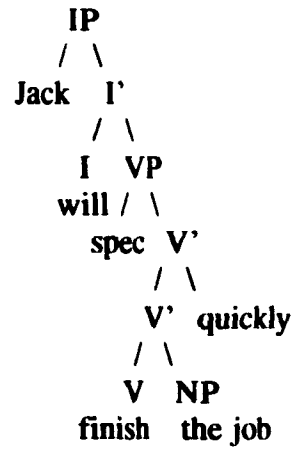
A technical problem arises when we consider sentences like the following:

(83) Jack will finish the job quickly.

In (83) the adverb *quickly* can only be interpreted as process-modifying. The question is: Why can it not be interpreted as event-modifying? The following would be the structure

for (83):

(84)



Under the definition of RM in (29) which refers to c-command, V does not intervene between INFL and the position of *quickly*. The adverb position is m-commanded and feature-governed by INFL. As a result, it is wrongly expected that *quickly* in (83) can be construed as event-modifying. What we need is to guarantee that V "protects" these positions from being m-commanded by INFL.

We would like to suggest that "intervention" in RM, i.e., the clause (ii), should be relativized depending on the types of government. This is a natural extension of the idea we have been pursuing that RM has to be further relativized. Since feature-government involves m-command rather than c-command, intervention for its purpose should be defined in terms of m-command as well. Accordingly, RM can be reformulated as in (85):

(85) Relativized Minimality: X α-governs Y only if there is no $\bar{Z}$ such that
(i) Z is α-GT compatible with X
(ii) Z ß-commands Y and does not ß-command X.

where ß-command is the type of command utilized in α-government.[31] Now, relativization of minimality initiated by Rizzi (1990) is complete. The definition allows us to keep c-command as the relevant command that defines intervention for the purpose of antecedent government. Given (85), V qualifies as "a closer potential licenser" for the adverb in (84).

With (82) and (85), the transportability and interpretations of Type II adverbs follow directly from feature government. Notice that the adoption of (82) and (85) does not necessitate any change in the above account of *wh*-elements, NPIs, and adverbs.

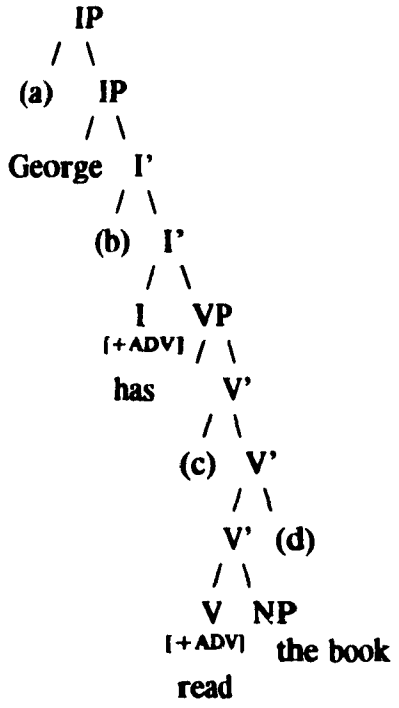
As one can verify easily, Type I adverbs can be accounted for in the same way as type II.

Next, let us turn to Type III and Type IV adverbs. Consider first Type IV adverbs which are licensed exclusively by V. The following examples are developed based on Jackendoff (1972):

(86) a. *Completely George has read the book.
b. *George completely has read the book.
c. George has completely read the book.
d. George has read the book completely.

These examples can be readily explained by the requirement in (81). The structure for (86) would be the following:

(87)



(86a,b) are ruled out as a violation of (81). The adverbs in (86a,b) are not m-commanded by V in the first place. On the other hand, the adverbs in (86c,d) are feature-governed

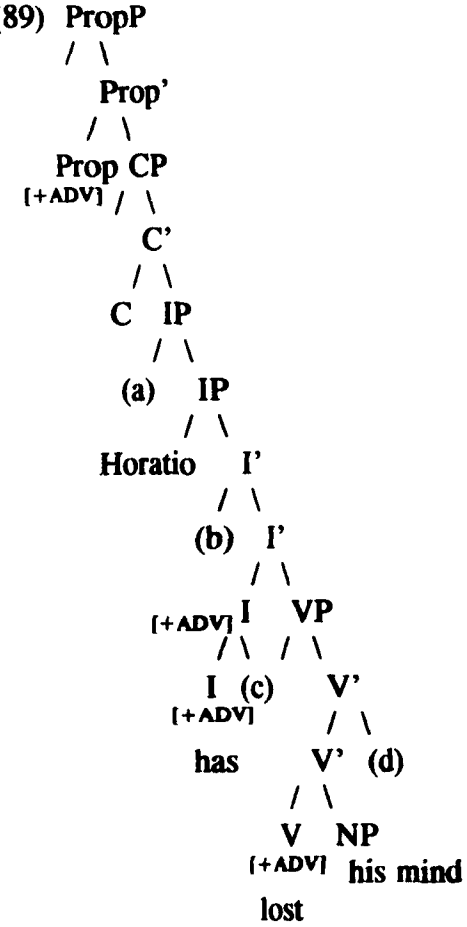
by V. INFL does not intervene between V and the positions (c) and (d).

Now, consider the following examples of Type III adverbs developed again based on Jackendoff (1972):

(88) a. Evidently Horatio has lost his mind.
b. Horatio evidently has lost his mind.
c. Horatio has evidently lost his mind.
d. *Horatio has lost his mind evidently. (without comma reading)[32]

Under Travis' analysis, Type III adverbs are licensed by INFL but not by V. Here we depart from Travis and hypothesize that Type III speaker-oriented adverbs are licensed by the feature [+ADV] of the (empty) head Prop(osition) which we assume is sister to CP.[33] The S-structure configuration for the examples in (88) would be as follows:

(89)



Given the above-mentioned hypothesis, the judgements in (88) can be explained by (81) and the ECP. Suppose that the adverb *evidently* remains in the S-structure positions (a)-

(d) at LF. Then, it cannot be licensed by [+ADV] in Prop except in the case where it is in the incorporated position (c). If the adverb is in the positions (a) and (b), [+ADV] in INFL m-commands it and intervenes between it and [+ADV] in Prop. As a result, the adverb cannot be feature-governed by [+ADV] in Prop, violating (81). If the adverb is in the position (d), it is "protected" from [+ADV] in Prop by [+ADV] in V, being left unlicensed. The adverb in (c) is feature-governed by [+ADV] in Prop if we assume, as seems reasonable, that in (89) the feature [+ADV] dominates and thus does not m-command the position (c); [+ADV] in INFL by definition does not intervene between [+ADV] in Prop and the adverb.[34]

Since the adverb *evidently* cannot be feature-governed by [+ADV] in Prop in the positions, (a), (b), and (d), it must raise at LF in order to satisfy (81). For the sake of concreteness, let us suppose that speaker-oriented adverbs can move to SPEC of PropP at LF. Then, the respective LF representations for the examples in (88) would be the following (the representation for (88c) is provided though LF raising of the adverb is not obligatory in this case):

(90) a. $[_{PropP}$ evidently$_i$ Prop$_{[+ADV]i}$ $[_{CP}$ $[_{IP}$ t_i $[_{IP}$ Horatio has$_{[+ADV]}$ lost his mind]]]]
b. $[_{PropP}$ evidently$_i$ Prop$_{[+ADV]i}$ $[_{CP}$ $[_{IP}$ Horatio t_i has$_{[+ADV]}$ lost his mind]]]
c. $[_{PropP}$ evidently$_i$ Prop$_{[+ADV]i}$ $[_{CP}$ $[_{IP}$ Horatio $[_I$ has$_{[+ADV]}$ $t_i]$ lost his mind]]]
d. *$[_{PropP}$ evidently$_i$ Prop$_{[+ADV]i}$ $[_{CP}$ $[_{IP}$ Horatio has$_{[+ADV]}$ lost his mind t_i]]]

Since we are assuming that a licenser gets coindexed with its licensee under feature government, in (90) Prop is coindexed with the adverb and qualifies as the antecedent governor for the trace. In (90a,b, and c) Prop with [+ADV] antecedent-governs the trace as required by the ECP since INFL with [+ADV] does not intervene between the two (recall that for antecedent government, "intervention" in RM is defined in terms of c-command). We are assuming of course that in (90c) [+ADV] in INFL dominates and does not c-command the trace. (90d) is ruled out by the ECP together with RM since INFL with [+ADV] intervenes between Prop with [+ADV] and the trace, disrupting antecedent government by Prop.

In brief, LF raising of the adverb saves (88a,b) from violating (81) (88c), even with the adverb in its S-structure incorporated position at LF, is well-formed. There is no well-formed LF representation for (88d); if the adverb stays in its S-structure position, it cannot be feature-governed by [+ADV] in Prop, violating (81); if it raises at LF, the trace cannot be antecedent-governed, violating the ECP. This account explains why Type III adverbs cannot appear within VP.

The assumption that Type III adverbs undergo raising at LF also helps us to account for the problem of sequencing exemplified in (75). The LF representations for (75) in which the adverb *probably* remains where it is generated would be the following:[35]

(91) a. *$[_{PropP}$ $Prop_{[+ADV]}$ $[_{IP}$ probably $[_{IP}$ Max $carefully_{[+ADV]}$ $was_{[+ADV]}$ $[_{VP}$ climbing the walls of the garden]]]]
b. $[_{PropP}$ $Prop_{[+ADV]}$ $[_{IP}$ Max $[_{I}$ probably $was_{[+ADV]}]$ $[_{VP}$ carefully climbing the walls the garden]]]
c. *$[_{PropP}$ $Prop_{[+ADV]}$ $[_{IP}$ $carefully_{[+ADV]}$ $[_{IP}$ Max $[_{I}$ probably $was_{[+ADV]}]$ $[_{VP}$ climbing the walls of the garden]]]]
d. *$[_{PropP}$ $Prop_{[+ADV]}$ $[_{IP}$ Max $carefully_{[+ADV]}$ $[_{I}$ $was_{[+ADV]}$ probably] $[_{VP}$ climbing the walls of the garden]]]

(91a, c, and d) are ill-formed, whereas (91b) is well-formed. In (91a) *probably* is m-commanded by [+ADV] in INFL, which blocks feature government of the adverb by [+ADV] in Prop. Thus, the adverb is unlicensed. In (91b), where the incorporated structure of INFL is assumed, *probably* is not m-commanded by [+ADV] in INFL by definition and is feature-governed by [+ADV] in Prop, satisfying (81). It may seem that in (91c,d) *probably* is feature-governed by [+ADV] in Prop since, as in (91b), the adverb is incorporated into INFL and [+ADV] in INFL does not induce a RM effect. To account for (91c,d) and other examples, we present the following:

(92) Feature Government Algorithm
A licenser and its licensee come to share an index and (a) feature(s) as a result of feature government.

We have already mentioned that a licenser and its licensee get coindexed under feature

government. Now we will also assume that there is copying of features between the licenser and the licensee at the same time. Given (92), in (91c,d) *carefully* acquires [+ADV] at LF by virtue of being licensed by [+ADV] in INFL. As a consequence, feature government of *probably* by [+ADV] in Prop is blocked by [+ADV] in *carefully*. (91c,d) are excluded as a violation of (81).

Turning next to raising of *probably*, consider the following LF representations:

(93) a. $[_{PropP}$ probably$_i$ Prop$_{[+ADV]i}$ $[_{IP}$ t_i $[_{IP}$ Max carefully$_{[+ADV]}$ was$_{[+ADV]}$ $[_{VP}$ climbing the walls of the garden]]]]

b. $[_{PropP}$ probably$_i$ Prop$_{[+ADV]i}$ $[_{IP}$ Max t_i was$_{[+ADV]}$ $[_{VP}$ carefully climbing the walls of the garden]]]

c. *$[_{PropP}$ probably$_i$ Prop$_{[+ADV]i}$ $[_{IP}$ carefully$_{[+ADV]}$ $[_{IP}$ Max t_i was$_{[+ADV]}$ $[_{VP}$ climbing the walls of the garden]]

d. *$[_{PropP}$ probably$_i$ Prop$_{[+ADV]i}$ $[_{IP}$ Max carefully$_{[+ADV]}$ $[_{I}$ was$_{[+ADV]}$ t_i] $[_{VP}$ climbing the walls of the garden]]

The question here is whether these representations observe the ECP. (93a) is well-formed since the trace is antecedent-governed by Prop; there is no feature [+ADV] intervening between the two that will block the antecedent government. (93a) also satisfies (81) because *probably* is now feature-governed by [+ADV] in Prop. In (93b) the trace is antecedent-governed by Prop; whether or not *probably* is incorporated into INFL before raising, [+ADV] in INFL does not intervene between the adverb and its trace. (93c) violates the ECP regardless of the original position of *probably* due to the presence of *carefully*, which gets [+ADV] from INFL and blocks antecedent government of the trace by Prop with [+ADV]. (93d) is ruled out as an ECP violation in the same way as (93c). (75c,d) are ungrammatical since they have no legitimate LF representations.

In the case of licensing of adverbs, which is very local, barriers are irrelevant. It seems that the Feature Government Parameter is also irrelevant.[36]

3.3. Summary

In this chapter, we have examined licensing of three kinds of feature-dependent

operators, i.e., *wh*-elements, NPIs, and adverbs. We have proposed the FDI-Criterion, which requires that a FDI be feature-governed by its licenser at LF. Feature government has been claimed to be defined in terms of m-command and RM. It has been argued that the FDI-Criterion is superior to the *Wh*-Criterion and the Neg-Criterion both empirically and theoretically. The FDI-Criterion can account for (i) minimality effects, (ii) the irrelevance of barriers, (iii) the cross-linguistic variations (in conjunction with the Feature Government Parameter), and (iv) licensing of adverbs. The above discussion of adverbs has led us to revise the definitions of m-command and RM.

It is worth pointing out that RM via the definition of feature government plays an integral part in licensing of FDIs. We have already seen in Chapter 2 that RM contributes to fixing scope of operators in a significant way. Thus, we conclude that RM, more precisely GT-compatibility, governs the interpretative and distributional behavior of operators to a large extent.

FOOTNOTES TO CHAPTER 3

1. For recent discussions, see among others Lasnik and Saito (1984, 1992), Chomsky (1986b), Cinque (1990), Rizzi (1990). The issue of superiority will be dealt with in Chapter 4.

2. Rizzi (1991) proposes that the *Wh*-Criterion must be construed as requiring that the chain of the X^0 position in question have the feature [+WH]. Given this extension and coindexation under agreement, the *Wh*-Criterion is satisfied in (i) at S-structure, as desired (Rizzi claims that the main inflection can be specified as [+WH]):

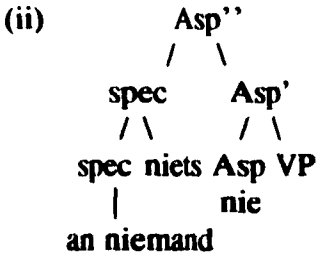
(i) $[_{CP}$ who$_i$ C$_i$ $[_{IP}$ t_i I$_i$ love-s$_{[+WH]}$ Sylvia]]

3. Here we are simplifying matters. The Neg-Criterion is intended to cover not only NPIs but also other kinds of negative elements (see Haegeman (1991)). We will restrict our attention to NPIs.

4. What prompted Haegeman (1991) to assume the adjunction structure in (13) is the fact that NPIs do not have to be adjacent to *nie*, as shown in (i):

(i) da Valère an niemand dienen boek nie gegeven en-oat
that to no one that book not given had
'that Valère had not given anyone that book.'

Along the lines of Johnson (1991), Travis (1992) suggests the possibility that the double object in (i) forms a DP and moves as a constituent. Travis also presents an alternative account of West Flemish NPIs which avoids the problems Haegeman faces (e.g. VP-fronting facts). She analyzes *nie* as an adverb incorporated into Asp within VP (see Travis (1991)). The relevant portion of S-structure representation for (12) would be as follows:

(ii)


We will not choose between the two analyses. What concerns us here is that NPIs must enter into a Spec-head agreement relation with a negative head at S-structure in West Flemish (and presumably in Hungarian (see Haegeman (1991)).

5. As (14) shows, NPIs are licensed in the subject position of the matrix clause in Japanese. This is not the case in English (without inversion of negative elements). See Section 3.2.2. below.

6. But see footnote 5 to Chapter 1.

7. See Aoun and Li (to appear) for arguments that various phenomena that used to be explained in terms of LF *Wh*-movement (e.g. scope, the ECP effects, Weak Crossover effects) can be handled even without postulating such movement.

8. But as a result of feature government, X and Y get coindexed (see below).

9. Nishigauchi (1990:33-36) notes that in examples like (36b), some speakers allow the matrix reading of *dare* 'who' if it is pronounced with an extra heavy stress. This option is not open to *nani* 'what' even if it receives a heavy stress. Thus, the word order seems to play a role. Nishigauchi concludes correctly, in our opinion, that the wide scope reading of the *wh*-QP is attributed to extra focus and is pragmatic in nature.

10. If the assumption that *wh*'s-in-situ do not have to move to SPEC of CP is correct (but see (51)), it is not clear how the well-known argument-adjunct asymmetries can be explained syntactically. One might explore a semantic or pragmatic account. See, for instance, Liejiong (1990) for a semantic account of Huang's (1981, 1982) Chinese data. We will leave this issue open for further work.

11. S-structure adjunction of *wh*-phrases is optional (i.e., *wh*'s-in-situ are allowed in a nonecho question) in Serbo-Croatian but it is obligatory in Polish and Czech (see Rudin (1988)).

12. (45) also explains among other things the presence of *wh*-island effects in Polish-type languages and the absence of such effects in Bulgarian-type languages (see Rudin (1988)).

13. The definition of government given in Chomsky (1986:9) is the following:

(i) α governs ß iff α m-commands ß and there is no Γ, Γ a barrier for ß, such that Γ excludes α.

Since Cinque's (1991) barriers system is adopted in this thesis, we translated (i) into the present context, as in (46).

14. Polish sentences like (i) are problematic for the present account (Rudin (1988)):

(i) Maria myśli, że co Janek kupił?
thinks that what bought
'What does Maria think that Janek bought?'

But notice that Polish normally does not allow extraction out of finite clauses. We will leave this issue open.

15. One may suggest along the lines of Kim (1989, 1991) that nonmovement of *wh*-QPs directly follows from their QP status. But behind this suggestion is the assumption that QPs universally remain in-situ at S-structure. There are, however, languages like Hungarian which obligatorily move QPs at S-structure (see Kiss (1986)).

16. The full form of the question particle is *sheno* 'what'. The contracted form *sh-* is used when followed by verbs. See Wahba (1991).

17. From the viewpoint of acquisition, the "default" value must be to meet (50) at LF (cf. Chomsky (1989)). Children learning languages like English can switch to the "nondefault" value by encountering positive evidence which involves overt *wh*-movement.

18. See footnote 17 to Chapter 2.

19. See Baker (1991) for a proposal similar to Ernst's (1992).

20. It has been argued that polarity *any* is an existential quantifier (see Carlson (1980) and Linebarger (1980)). Then, it would be plausible that NPIs containing *any* can undergo QR. In the representations to follow in this chapter, however, we place such NPIs in-situ.

21. Linebarger's (1987) Immediate Scope Constraint (ISC), which requires that no "logical element" intervene between a NPI and its licenser, implies that this should be the case if the feature |+NEG| is taken to be a "logical element", as seems reasonable. The ISC will be discussed in Chapter 4.

22. There seem to be semantic or pragmatic restrictions on NPI licensing. Compare (68) with following example from Ross (1967):

(i) *Waldo did not report the possibility that anyone had left.

Discussion of such restrictions is beyond the scope of this thesis. See Linebarger (1987) for a proposal regarding a pragmatic constraint.

23. We are not aware of any language in which the feature |+NEG| obeys (55).

24. One of the problems with Progovac's (1988) approach is that NPIs are assumed to raise into SPEC of CP when they are to be licensed by superordinate negation. Thus, in examples like (60a,b) *anyone* must move to SPEC of CP. Since NPIs like *anyone* are arguably existential quantifiers (see footnote 20), this assumption runs counter to the general clause-boundedness of QR of QPs.

25. There are two types of adverbs other than those listed in (75):

(i) a. Type V: VP-final (hard, well, more, ...)
b. Type VI: Aux (truly, virtually, merely, ...)

These types will not be discussed. We suspect that Type V adverbs may be those that must be sister to V at D-structure (cf. Larson (1988), Stroik (1990)) and that Type VI adverbs may be heads that cannot be adjoined to maximal projections at D-structure.

26. For thorough discussion of "syntactic" incorporation, see Baker (1988).

27. The more precise generalization seems to be that only arguments (such as N and P) can incorporate into lexical categories, as Lisa Travis (p.c.) suggests.

28. To be precise, only adverbs that must be licensed by a feature are subject to (81). XP adverbs such as *wh*-adjuncts and PP adverbs, which we assume to be licensed by either predication or adverbial Θ-role assignment (see Larson (1985) and Stroik (1990) for the latter), do not obey (81).

29. But we do assume that Type III adverbs raise at LF (see below).

30. The precise formulation of (82) was suggested to us by Lisa Travis (p.c.). As Travis points out, (82) makes government the relation of mutual m-command (see (46)).

31. We will not discuss implications of (85) for other types of government (see Rizzi (1990:112, fn.4)).

32. Presumably, the comma reading permits the adverb to be adjoined to IP (cf. Ernst (1984)).

33. Rochette (to appear) argues that Type III adverbs select a proposition, which is syntactically realized as a CP. But examples like (ib) appear to be problematic for this analysis (Rochette (to appear)):

(i) a. Mary believes that John will probably win the prize.
b. *Mary wishes that John probably win the prize.

It is not clear why the adverb is not allowed in (ib) since under the standard assumption, the embedded clauses in (i) are both full CPs. (ib) also shows that the presence of COMP is not a sufficient condition on licensing of speaker-oriented adverbs. Since the feature [+ADV] seems to be an inherent feature, it would be feasible to assume that COMP does not contain [+ADV]. For these reasons, we assume that Prop has [+ADV]. (ia) is fine because *believe* subcategorizes for PropP, while (ib) is ungrammatical because *wish* subcategorizes for CP. Empirical justification for positing PropP awaits further work.

34. Note that the impossibility of the event-modifying reading of *quickly* in (77c,d) does not pose a problem since adverbs cannot incorporate into V and [+ADV] in V protects *quickly* from being feature-governed by [+ADV] in INFL in (77c,d).

35. Note that the adverb *carefully* cannot be incorporated into INFL. If it is, INFL with [+ADV] will dominate it and, as a result, will not m-command it. This will lead to a violation of (81). If *probably* is outside of the m-command domain of INFL in (91a), the representation is well-formed. If it is not incorporated into INFL in (91b), the representation is ruled out for the same reason as (91a). Also, (91c,d) with the unincorporated structure of INFL are excluded in the same way.

36. The Condition on FDI Licensing Feature given in (55), however, may be at work in licensing of English adverbs. Jackendoff (1972:88) notes that in English, there cannot be more than one "subject-oriented" adverbs. For example (Jackendoff 1972:90):

(i) *Carefully, Max quickly was climbing the walls of the garden.

It may be that the feature [+ADV] in INFL in English is constrained by the condition.

CHAPTER 4

RELATIVIZED MINIMALITY

4.0. Introduction

In the preceding chapters, it was argued that the scope assignment of operators and the licensing of feature-dependent items are constrained by Relativized Minimality. But so far, its proper formulation, the notion of Government Theory Compatibility in particular, has been kept rather vague.

This chapter is thus devoted to the issue of how Government Theory Compatibility should be characterized. The above discussions have already suggested a direction to take in order to solve this problem; the main thrust of this chapter is the claim that Government Theory Compatibility is better defined in terms of a set of lexical features.

In Section 4.1., inner island phenomena, which prompted the proviso in Rizzi's (1990) Relativized Minimality that elements in an adjoined position do not induce Relativized Minimality effects, are briefly considered. It is pointed out that the proviso is ill-motivated and unnecessary. The modified version of Relativized Minimality proposed above can handle inner islands and thus gains further support. In Section 4.2., we seek an appropriate notion of Government Theory Compatibility. A formulation within our feature system is presented tentatively in the hope that it will prove useful for discovering the exact formulation which awaits further work. In Section 4.3., the Scope Principle advanced in Chapter 2 is reexamined. It is shown that data other than those considered in Chapter 2 can be accounted for by the Scope Principle.

4.1. A Problem of Rizzi (1990)

Let us reconsider Rizzi's (1990) original formulation of RM introduced in Section 2.1.3. It is repeated below for convenience:

(1) Relativized Minimality: X α-governs Y only if there is no Z such that
(i) Z is a base-generated position
(ii) Z is α-GT compatible with Y
(iii) Z c-commands Y and does not c-command X.

Recall that GT-compatibility for (1) is defined purely in syntactic terms. For example, an element Z is GT-compatible with an element Y in an A'-chain only if it is in an A'-specifier position.

The consideration of the scope of a QP and a *wh*-QP in Japanese in Section 2.2.1 led us to propose to eliminate the clause (i), as in (2) where α-government uses ß-command in its definition (see Section 3.2.3.):[1]

(2) Relativized Minimality: X α-governs Y only if there is no Z such that
(i) Z is α-GT compatible with X
(ii) Z ß-commands Y and does not ß-command X.

The elimination of (1i) has enabled us to subsume under the ECP the Minimal Binding Requirement of Aoun and Li (1989) which states that variables must be bound by the most local potential antecedent (A'-binder) (see Section 2.2.1.). It appears that (2) has empirical and theoretical advantages over (1).

The questions that we have put off until now are: What motivated Rizzi (1990) to include the clause (1i) in RM? Is there any price to pay for adopting (2) instead of (1)? Before going into these questions, let us remind ourselves of the definitions of the ECP and antecedent government adopted in this thesis:

(3) Empty Category Principle: A nonpronominal empty category must be
(i) properly head-governed
(ii) antecedent-governed.

(4) Antecedent Government: X antecedent-governs Y iff
(i) X and Y are coindexed
(ii) X c-commands Y
(iii) no barrier intervenes
(iv) Relativized Minimality is respected

Rizzi's (1990) argument for the clause (i) in (1) is based on so-called "inner island" effects (see Ross (1984)). A relevant pair of examples is the following:

(5) a. It is for this reason that no one believes that Bill was fired.
b. It is for this reason that everyone believes that Bill was fired.

There is a contrast in interpretation between (5a) and (5b). (5b) is ambiguous, allowing both the higher and the lower construals of the adverbial phrase. It can mean either "This is the reason for the fact that everyone believes that Bill was fired", or "This is the reason such that everyone believes that Bill was fired for it". In (5a), on the other hand, the lower construal is impossible, though the higher one is possible.

Rizzi assumes that "affective" operators (i.e., operators that enter into negative polarity cf. Klima (1964)) like *no one* move to SPEC of CP at LF, whereas "nonaffective" operators like *everyone* are subject to QR. Under this assumption, the LF representations for (5a,b) with the lower construals are as follows:

(6) a. *It is [for this reason]$_i$ [$_{CP}$ no one$_j$ that [$_{IP}$ x_j believes [that Bill was fired t_i]]]
b. It is [for this reason]$_i$ [$_{CP}$ that [$_{IP}$ everyone$_j$ [$_{IP}$ x_j believes [that Bill was fired t_i]]]]

Notice that the focused element is an adjunct. Therefore, its trace must satisfy the ECP at LF (see Lasnik and Saito (1984, 1992)). According to RM in (1), antecedent government will never be blocked by an element in an adjoined position. Thus, (6b) observes the ECP. (6a) is ruled out by the ECP; the offending A'-specifier *no one* blocks antecedent government by the focused element which heads an A'-chain. Hence, the contrast in (6).

We would like to point out, however, that RM in (1) wrongly rules out even the higher construal of (5a), as is illustrated in (7):[2]

(7) *It is [for this reason]$_i$ [$_{CP}$ no one$_j$ that [$_{IP}$ t_i [$_{IP}$ t_j believes [that Bill was fired]]]]

In (7) the antecedent government relation between the focused element and its trace is disrupted by the A'-specifier *no one*, exactly as in the case of (6a).

The question then is: How can we rule out the lower construal, while ruling in the higher one in (5a)? Suppose that QPs, affective or nonaffective, are subject to QR. Let us assume that negative elements have the feature [+NO] meaning "negative

operator" in addition to the licensing feature |+NEG| and that adverbial elements, including those which are feature-dependent and those which are not (see footnote 28 to Chapter 3), have the feature |+AO| meaning "adverbial operator" [3] Let us further assume that negative element with |+NO| are GT-compatible with adverbial elements with |+AO|.[4] Given these assumptions, the two LF representations for (5a) would be as follows:

(8) a. *It is [for this reason]$_{[+AO]i}$ [$_{CP}$ that [$_{IP}$ no one$_{[+NO]j}$ [$_{IP}$ x_j believes [that Bill was fired t_i]]]]
b. It is [for this reason]$_{[+AO]i}$ [$_{CP}$ that [$_{IP}$ t_i [$_{IP}$ no one$_{[+NO]j}$ [$_{IP}$ x_j believes [that Bill was fired]]]]]

RM in (2) can account for the unambiguity of (5a). (8a) violates the ECP because antecedent government of the trace t_i by the adverbial focused element is disrupted by intervening *no one*. (8b) is well-formed since *no one* does not intervene between the focused element and its trace.[5] Antecedent government by an adverbial element will not be blocked by a nonnegative operator like *everyone* which lacks the feature |+NO|. This is why (5b) is ambiguous.

To sum up, there is no motivation for restricting RM inducers to elements in base-generated positions, as far as inner islands are concerned. In fact, the elimination of the clause (i) in (1) is necessary to explain such construal contrast as the one in (5). Examples like those in (5) support the version of RM given in (2) after all.

4.2. Feature-Based Relativized Minimality

Having established that RM in (2) is superior to Rizzi's (1990) RM, the remaining task is to determine what would be a better way to characterize the notion of GT-compatibility. Without a proper notion of GT-compatibility, RM simply would not work. A promising way to go about this task has already suggested itself in the course of the preceding discussions, i.e., it seems appropriate to define GT-compatibility in terms of lexical features of governors.

The purpose of this section is to present a feature-based formulation of GT-compatibility. Though this formulation is highly theory-internal and depends entirely on what kind of feature system is employed, our hope is that it will serve as a stepping stone for a still better concept of GT-compatibility.

4.2.1. Minimality in Antecedent Government

This subsection is concerned with minimality effects in antecedent government (see (4) for the definition of antecedent government). First, we review the relevant configurations examined in the preceding discussions. Secondly, superiority effects are considered in the light of RM. It is suggested that superiority violations can be explained by the ECP. Observing the lack of superiority at LF, we present a condition, which accounts for the problem posed by Baker (1970a). We incorporate this condition into the definition of RM. To handle Bulgarian-type languages, in which superiority effects are totally absent, another condition is put forth. Thirdly, given LF raising of epistemic modals, the restriction on them with respect to sequencing and the unsolved examples from the previous chapter receive a RM account.

4.2.1.1. Review

So far, we have used licensing features such as [+WH], [+Q], [+NEG], and [+ADV] and operator features such as [+NO] and [+AO]. Here we would like to introduce two new operator features. It will be assumed that standard QPs are assigned the feature [+SO] meaning "standard operator" and that *wh*-phrases are assigned the feature [+QO] meaning "quasi-operator" (cf. Chomsky (1975) and others).[6]

Consider first antecedent government involving the four operator features, i.e., [+SO], [+QO], [+AO], and [+NO]. The configuration we are interested in is the following sort:

(9) ... X ... Z ... Y ...

where an element c-commands the one to its right, Y is the trace of X, and there is no barrier between them. By replacing X and Z with the four different features, we obtain the sixteen logically possible configurations. Of these configurations, we will examine the four in (10) which are directly relevant to the current discussion:

(10) a. * ... |+SO|$_i$ ... |+SO| ... t_i ...
b. (*) ... |+QO|$_i$ ... |+QO| ... t_i ...
c. * ... |+AO|$_i$ ... |+AO| ... t_i ...
d. * ... |+AO|$_i$ ... |+NO| ... t_i ...

As far as we can tell, the configurations other than those in (10) are either nonexistent or consistent with the present analysis.[7]

Putting (10b) aside for the moment (we will return to it in the next subsection), let us start with (10a and c). The relevant examples we have already examined above are repeated below:

(11) Dareka -ga daremo -o aisiteiru (koto)
someone-NOM everyone-ACC love fact
'Someone loves everyone.'

(12) *Max carefully was probably climbing the walls of the garden.

Recall that the Japanese (11) is unambiguous with the wide scope reading of *dareka* 'someone'. It was suggested above that the following LF representations for (11) and (12) are excluded by the antecedent government requirement of the ECP:[8]

(13) *[$_{IP}$ dareka$_{[+SO]i}$-ga [$_{IP}$ daremo$_{[+SO]j}$-o [$_{IP}$ x_i [$_{VP}$ x_j aisiteiru]]]]

(14) *[$_{PropP}$ probably$_{[+AO]i}$ Prop$_{[+AO]i}$ [$_{IP}$ Max carefully$_{[+AO]}$ was t_i climbing the walls of the garden]]

The SP advanced in Section 1.4.1. is recapitulated in (15) for convenience:

(15) The Scope Principle (SP)
An operator A has scope over an operator B in case
(i) A c-commands a member of the chain containing B (A $\subseteq$ B) or
(ii) all members of the chain containing A c-command a member of the chain containing B (A $\not\subseteq$ B).

where "A $\subseteq$ B" means that A is GT-compatible with B, whereas "A $\not\subseteq$ B" means that A is not GT-compatible with B. Since (13) involves QPs, it is subject to the first clause of the SP. If it were well-formed, (11) would be expected to be ambiguous since *dareka* c-commands *daremo* and the latter c-commands the variable of the former, contrary to the fact. (14), where the speaker-oriented adverb *probably* has raised to be licensed by Prop, is ill-formed. If *probably* does not raise at LF, it will be left unlicensed (*carefully* blocks the required licensing), causing a violation of FI. Thus, (12) is simply ungrammatical.

It has already been noted that the configuration (10d), exemplified above in (8a), results in ungrammaticality. An additional piece of evidence for the general ill-formedness of (10d) is provided by (16a) (from Jackendoff (1972:84)):

(16) a. *Did Frank probably beat all his opponents?
b. Did Frank easily beat all his opponents?

Since (16a,b) are questions capable of licensing NPIs, the COMP contains the feature [+NO] (as well as [+NEG]). The LF representations for (16a,b) are (17a,b) respectively:[9]

(17) a. *$[_{PropP}$ probably$_{[+AO]i}$ Prop$_{[+AO]i}$ $[_{CP}$ did$_{[+NO]}$ $[_{IP}$ Frank t_i $[_{VP}$ beat all his opponents]]]]
b. $[_{CP}$ did$_{[+NO]}$ $[_{IP}$ Frank easily$_{[+AO]}$ beat all his opponents]]

Since in (16a) feature-government of *probably* by [+ADV] in Prop is disrupted by [+NEG] in COMP (we will come back to this point below), the LF raising of the adverb is obligatory. As shown in (17a), however, such raising is impossible due to the negative COMP, resulting in the ill-formed configuration given in (10d). (17b) is fine since nothing drives the V-licensed adverb *easily* to move at LF.

Notice that an account of (16a) resorting to some semantic restriction on the co-occurrence of the adverb and the question seems unsatisfactory. The well-formedness of the following tag-question challenges such a semantic account (Jackendoff (1972:85)):[10]

(18) Frank probably beat all his opponents, didn't he?

Thus, we suggest that the present syntactic analysis is preferable over a possible semantic analysis.

Turning next to antecedent government by variables, we introduce two variable features. One is [+SV] "standard variable", the other [+QV] "quasi-variable". Let us suppose that the former feature is assigned to variables of standard quantifiers and that the latter one is assigned to variables of *wh*-phrases. The generalizations regarding antecedent government in question seem to be the following:[11]

(19) a. * ... x[+SV]$_i$ ... x[+SV] ... t_i ...
b. ... x[+SV]$_i$ ... x[+QV] ... t_i ...
c. ... x[+QV]$_i$ ... x[+SV] ... t_i ...

Consider the configuration (19a). A relevant Japanese example is repeated below from Chapter 2:

(20) Dare-o$_i$ daremo -ga t_i syootaisita no?
who-ACC everyone-NOM invited Q
'Who$_i$, everyone invited t_i?'

This scrambled sentence is interpreted unambiguously, the *wh*-QP taking wide scope with respect to the QP. Its LF representation would be (21):

(21) [$_{IP}$ dare$_i$-o [$_{IP}$ $x_{[+SV]i}$ [$_{IP}$ daremo-ga$_j$ [$_{IP}$ $x_{[+SV]j}$ [$_{VP}$ t_i syootaisita]]]]] no
|_____| |_____|

In (21) the link between the variable x_i and the trace t_i is broken by the intervening variable x_j since an antecedent government relation does not hold between them (recall that a chain is defined in terms of antecedent government). The *wh*-QP and the QP, being GT-compatible with each other, are interpreted according to the first clause of the SP. The SP together with (21) makes a correct prediction about the interpretation of (20). If the broken link were present in (21) or if (19a) were well-formed, (20) would be wrongly expected to be ambiguous. When scrambling is reconstructed, the resulting LF representations are all illegitimate (see Section 2.1.1.).

Now consider the following psych constructions in connection with (19b, c) (again

repeated from Chapter 2):

(22) a. Who does everyone excite?
b. Who excites everyone?

As noted above, (22a) is unambiguous with the *wh*-phrase having wide scope, whereas (22b) is ambiguous. Continuing to assume Belletti and Rizzi's (1988) analysis of psych constructions, the LF representations for (22a,b) would be as follows:

(23) a. $[_{CP}$ who$_i$ does $[_{IP}$ everyone$_j$ $[_{IP}$ $x_{[+SV]i}$ $[_{VP}$ excite $[_{V'}$ $t_j]]$ $x_{[+QV]i}]]]]$

b. $[_{CP}$ who$_i$ $[_{IP}$ $x_{[+QV]i}$ $[_{VP}$ everyone$_j$ $[_{VP}$ excites $[_{V'}$ $t_i]]$ $x_{[+SV]i}]]]]$

Notice that at D-structure the experiencer argument (*who* in (22a) and *everyone* in (22b)) c-commands the theme argument (*everyone* in (22a) and *who* in (22b)). Since standard QPs are not GT-compatible with *wh*-phrases, those in (23) are subject to the second clause of the SP. In (23a) the link between the QP variable x_j and the trace t_j is not disrupted by the presence of the *wh*-variable x_i. Given (23a), the SP can account for the unambiguity of (22a). If there were not a link between x_j and t_j, (22a) would be incorrectly predicted to be ambiguous. In (23b) the *wh*-variable x_i antecedent-governs the trace t_i despite the intervening QP-variable x_j. With (23b), the ambiguity of (22b) is expected by the SP. If there were not a link between x_i and t_i, we would predict that (22b) should be unambiguous, contrary to the fact. In short, the configurations in (19b,c) are well-formed.

One might wonder if we need to distinguish between operators from their variables vis-à-vis features? Simple examples like (24) repeated from Chapter 1 force us to distinguish a standard QP from its variable. The LF representation for (24) is provided in (25):

(24) Someone loves everyone.

(25) $[_{IP}$ someone$_{[+SO]i}$ $[_{IP}$ $x_{[+SV]i}$ $[_{VP}$ everyone$_{[+SO]j}$ $[_{VP}$ t_i loves $x_{[+SV]j}]]]]$

If variables left by standard QPs have the feature [+SO], it would be incorrectly expected that (25) is not available. In particular, (25) would involve the ill-formed configuration given in (10a), and the link between the variable x_i and the trace t_i would be broken by *everyone*. Then (24), which is subject to the first clause of the SP, would be wrongly expected to be unambiguous since there would be no overlap of the chains in its LF representation.

As for the distinction between *wh*-phrases and their variables, consider the following Bulgarian sentence repeated from Chapter 3 and its S-structure and LF representation:[12]

(26) Koj kogo ma kogo e pokazal?
who whom to whom has pointed out
'Who pointed out whom to whom?'

(27) $[_{CP}$ koj$_i$, kogo$_j$, ma kogo$_k$ $e_{[+QO]i,j,k}$ $[_{IP}$ pokazal x_i x_j $x_k]]$

Suppose that a *wh*-variable has the feature [+QO]. Then (27) would result in the configuration in (10b); the subject variable x_i would block antecedent government of the variables x_j and x_k by COMP (but see the next subsection). It would thus violate the ECP, and (26) would be expected to be ungrammatical, which is not the case. Therefore, it seems necessary to differentiate *wh*-phrases from their variables.

Finally, let us consider the following illegitimate configuration where the licensing feature [+ADV] is involved:

(28) * ... $[+ADV]_i$ ... [+ADV] ... t_i ...

We have already seen that examples like (29) are ungrammatical. The LF representation for (29) after the raising of *evidently* would be (30):

(29) *Horatio has lost his mind evidently. (without comma reading)

(30) *$[_{PropP}$ evidently$_i$ Prop$_{[+ADV]i}$ $[_{IP}$ Horatio has$_{[+ADV]}$ $[_{VP}$ lost$_{[+ADV]}$ his mind $t_i]]]$

In (29) the LF raising of *evidently* is triggered by the requirement that the adverb must be feature-governed by |+ADV| in Prop. If the adverb remains in-situ, |+ADV| in V will count as the closer potential licenser for it, blocking feature government by |+ADV| in Prop (30) must be excluded, otherwise (29) would be incorrectly predicted to be grammatical. It was suggested above that in (30) antecedent government by Prop with |+ADV| of the trace of the adverb is blocked by |+ADV| in INFL. In short, (28) is ruled out by the ECP.

Notice that (29) does not contain the configuration (10c) with the operator feature |+AO|. Since examples like (12) can be excluded as involving (28) (see footnote 8), a question arises as to whether we can dispense with |+AO|. Examples such as the following show that it is necessary to have both |+ADV| and |+AO|:

(31) When did everyone see Max?

(32) $[_{CP}\ when_{[+AO]i}\ did_{[+AO]}\ [_{IP}\ everyone_j\ [_{IP}\ x_j\ I_{[+ADV]}\ [_{VP}\ t_j$ see Max $x_i]]]]$

(32) is the LF representation for (31). Since *when* is an adjunct, its variable must satisfy the ECP at LF. If |+AO| can be identified with |+ADV|, (32) would be expected to be ungrammatical, contrary to the fact. In particular, the *wh*-variable would not be antecedent-governed by COMP with |+AO| due to INFL with |+ADV|, violating the ECP. In addition, the ambiguity of (31) indicates that there is a chain between COMP and the variable x_i at LF. Otherwise, the SP would wrongly predict the unambiguity of (31). Since chains are defined on the basis of antecedent government, it must be the case that in (32) COMP antecedent-governs x_i. This also points to the conclusion that |+AO| is different from |+ADV|.

4.2.1.2. Superiority

Going back to the configuration (10b) now, observe the classic paradigm of

superiority in (33) (cf. Chomsky (1973) and others):

(33) a. Who admires what?
b. *What does who admire?

(34a,b) would be the S-structure and LF representations for (33a,b):[13]

(34) a. $[_{CP}$ who$_{[+QO]i}$ C$_{[+QO]i,j}$ $[_{IP}$ x_i $[_{VP}$ t_i admires what$_{[+QO]j}]]]$
b. *$[_{CP}$ what$_{[+QO]i}$ does$_{[+QO]i,j}$ $[_{IP}$ who$_{[+QO]j}$ $[_{VP}$ t_j admire $x_i]]]$

In (34a) the two traces are antecedent-governed, satisfying the ECP. In (34b) x_i, the variable of *what*, cannot be antecedent-governed by COMP with [+QO] due to the presence of *who* with [+QO]. Thus, in examples like (33b), the configuration (10b) is ungrammatical. The same is true of "pure" superiority cases like (35b) (from Lasnik and Saito (1984:241, fn.10)):

(35) a. Who did you tell to read what?
b. ?*What did you tell who to read?

The S-structure and LF representations for these examples would be the following:

(36) a. $[_{CP}$ who$_{[+QO]i}$ did$_{[+QO]i,j}$ $[_{IP}$ you $[_{VP}$ tell x_i to read what$_{[+QO]j}]]]$
b. *$[_{CP}$ what$_{[+QO]i}$ did$_{[+QO]i,j}$ $[_{IP}$ you $[_{VP}$ tell who$_{[+QO]j}$ to read $x_i]]]$

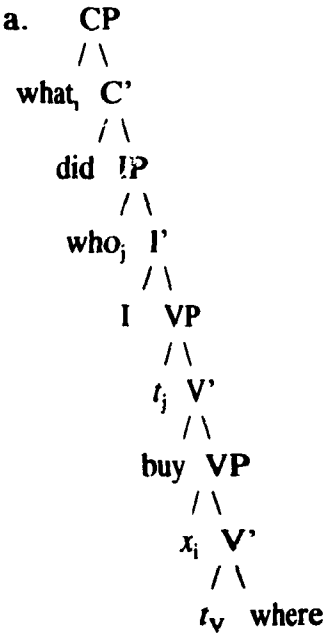
In (36a) the variable x_i is antecedent-governed by COMP. In (36b) *who* with [+QO] blocks antecedent government of the variable x_i by COMP with [+QO]. The ungrammaticality of (35b) does not follow from the Lasnik and Saito-type (1984) theory in which the disjunctive ECP and LF *wh*-movement are assumed since both of the *wh*-traces will be "lexically" governed at LF. In the present approach, superiority violations including (35b) can be accounted for in terms of the conjunctive ECP. This is an advantage over Lasnik and Saito (1992) who conclude that the Superiority Condition must be maintained independently of the ECP.[14]

It has been noted in the literature (cf. Kayne (1983), May (1985)) that the addition of *wh*-phrases considerably ameliorates superiority violations. Thus, we have the contrast in (37) (Lasnik and Saito (1984:241, fn.10)):

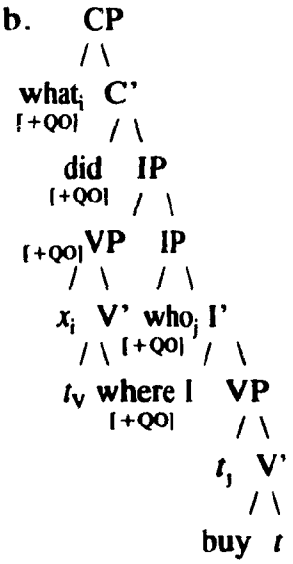
(37) a. *What did who buy at the store?
b. ?What did who buy where?

(37a) is ruled out in the same way as (33b) by the ECP. The question is: How can (37b) avoid an ECP violation which supposedly results in total ungrammaticality? In the present analysis under which the conjunctive ECP is assumed, the amelioration in grammaticality in (37b) must be attributed to some LF operation which makes antecedent government by *what* of its variable possible. In other words, the LF representation for (37b) like (38b) must be available ((38a) is the S-structure representation):

(38) a.



b.



In (38b) the lower VP has been adjoined to IP.[15] As a result, *what* antecedent-governs its variable in SPEC of the adjoined VP. The representation (38b) can be made available if we assume that the feature [+QO] of *where* percolates up to VP and that maximal projections marked with this feature can undergo QR.

There is one more assumption we need to explain the (relative) well-formedness of (38b), i.e., that antecedent government by the preposed VP of its trace is not blocked by *who*. Since we are advocating the approach under which RM is defined in terms of features, the antecedent government in question would not hold; both the preposed VP

and *who* share the feature [+QO]. To cope with this problem, some condition along the lines of (39) would be necessary:

(39) Visibility Condition on Relativized Minimality:
The configuration ... X ... Z ... Y ... is visible for Relativized Minimality only if the α-governor X is in its canonical position.

Let us say that for operators, canonical positions are those in which they acquire operatorhood in the unmarked case (Section 2.2.1.). Thus, the canonical position for *wh*-phrases is SPEC of CP. The canonical position for QPs is an adjoined position. In cases where the α-governor is a feature-governor, (39) applies vacuously.

Given (39), the preposed VP in (38b), being in a non-canonical position, antecedent-governs its trace despite the intervention of *who*. As a result, the configuration (10b) is allowed in cases like (38a). This is why the asterisk is put in parentheses in (10b).

A condition such as (39) can also be used to solve the problem posed by Baker (1970a). The relevant example is repeated below from Chapter 3:

(40) Who remembers where John bought what?

Recall that this sentence is ambiguous between the reading on which *what* is paired with *where* and the one on which it is paired with *who*. The unanswered question was: Why is the latter reading possible? Given (39), the following is a legitimate LF representation for (40):[16]

(41) $[_{CP}$ who$_i$ $C_{i,j}$ $[_{IP}$ x_i $[_{VP}$ what$_j$ $[_{VP}$ t_i remembers $[_{CP}$ where John bought $x_j]]]]]$

In the above representation, *what* has been raised and adjoined to the matrix VP. Nothing prevents this movement; regardless of the presence of *where*, *what* can antecedent-govern its variable x_j by virtue of (39) since it is in a non-canonical position; there is no barrier between *what* and its trace; movement of *wh*-phrases is not clause-bound. In (41) *what* is licensed by the matrix COMP rather than the embedded COMP, yielding the interpretation on which it is paired with *who*. The ambiguity of examples like (40) lends

further support to (39).

Note that the fact that the Japanese counterpart of (40) is unambiguous, the *wh*-QP in the embedded clause being unable to take the matrix scope, follows directly from the QP status of *wh*-QPs; raising of *wh*-QPs is clause-bound.

Returning now to the problem posed by multiple *wh*-questions in Polish-type languages (Section 3.2.1.2.), consider the following example from Serbo-Croatian:[17]

(42) Ko je što kome dao?
who has what to whom given
'Who gave what to whom?'

(43) $[_{CP}$ ko_i je $[_{IP}$ $što_j$ $[_{IP}$ $kome_k$ $[_{IP}$ dao x_i x_j $x_k]]]]$

The problem was: How can *ko* 'who' antecedent-govern its variable despite the presence of intervening *što* 'what' and *kome* 'to whom'? (43), in which *ko* is in SPEC of CP, conforms to (39) and thus is visible to RM. As a result, we would expect the IP-adjoined *wh*-phrases to block the antecedent government by *ko*.

One might hope that IP-adjoined *wh*-phrases in Polish-type languages do not acquire operatorhood at S-structure and can be reconstructed at LF. If this is the case, the variable of *ko* in (43) would satisfy the ECP at LF. Note, however, that Polish-type languages lack superiority effects entirely. In addition to (42), the other conceivable word orders are all possible (Rudin (1988:473)):

(44) a. Ko je kome što dao?
b. Što je ko kome dao?
c. Što je kome ko dao?
d. Kome ko što dao?
e. Kome što ko dao?

(44b-e) are problematic for the reconstruction account since even if the IP-adjoined *wh*-phrases are put back to their D-structure positions, *ko* still intervenes between the *wh*-phrases in SPEC of CP and their variables. Therefore, another explanation must be sought.

To account for the lack of superiority, we present the following condition:

(45) Condition on Relativized Minimality:
The configuration ... X ... Z ... Y ... is invisible to Relativized Minimality if Z is not in its canonical position.

Polish-type languages obey (45) but languages like English do not.[18] Since an IP-adjoined position is not a canonical position for a *wh*-phrase, the examples in (42) and (44) are exempt from RM. Hence their grammaticality.[19]

4.2.1.3. Epistemic Modals

Before leaving this subsection, let us consider epistemic modals and the question left unanswered in Section 3.2.2. which involves epistemic modals.

We assumed above that speaker-oriented adverbs such as *evidently* raise at LF to be licensed by Prop. This assumption enabled us to account for their distribution and the problem of sequencing of adverbs under RM.

Epistemic modals are often semantically characterized as speaker-oriented (cf. Jackendoff (1972), Palmer (1990)). Then it would be expected that they exhibit distributional properties similar to those of speaker-oriented adverbs, an expectation borne out (see Jackendoff (1972:100-105)). A simple-minded extension of the above explanation of speaker-oriented adverbs would be to handle these properties by RM together with the assumption that epistemic modals undergo raising at LF.

The idea that epistemic modals move at LF is not novel. McDowell (1987) argues that epistemic modals raise to SPEC of CP at LF to take scope over the whole proposition. Let us assume instead that epistemic modals, like speaker-oriented adverbs, must be licensed or feature-governed by [+ADV] of Prop, which drives them to raise to SPEC of PropP.[20] Let us further make the reasonable assumption that epistemic modals are assigned the feature [+AO] as speaker-oriented adverbs are.

With these assumptions, the ordering restriction exemplified by (46) follows automatically from the present analysis (Travis (1988:302)):

(46) a. Pete should carefully have crept out of there by now.
b. ?*Pete carefully should have crept out of there by now.

As in the case of speaker-oriented adverbs, epistemic modals cannot be preceded by subject-oriented adverbs. In (46a) *should* is dominated by [+ADV] in INFL since it has been incorporated into INFL. Thus, by definition, [+ADV] in INFL does not intervene between [+ADV] in Prop and the modal. Hence, the modal can be feature-governed by Prop without raising. This is why (46a) is grammatical. In (46b), on the other hand, the modal cannot be licensed in its S-structure position because *carefully*, which acquires [+ADV] from INFL at LF blocks feature government of the modal by [+ADV] in Prop. The LF representations for (46a,b) after the raising of *should* would be as follows:

(47) a. $[_{PropP}$ should$_{[+AO]i}$ Prop$_{[+AO]i}$ $[_{IP}$ Pete t_i $[_{VP}$ carefully$_{[+AO]}$ have crept out of there by now$]]]$
b. *$[_{PropP}$ should$_{[+AO]i}$ Prop$_{[+AO]i}$ $[_{IP}$ Pete carefully$_{[+AO]}$ t_i $[_{VP}$ have crept out of there by now$]]]$

In (47a) the trace is antecedent-governed by Prop. In (47b) Prop with [+AO] cannot antecedent-govern the trace due to intervening *carefully* with [+AO]. In other words, (47b) is excluded in the same way as (14) by the ECP.

We are now in a position to reconsider the following sentences from Chapter 3:

(48) ?*Who must have killed Yuri?

(49) *If John must know the answer, he is lucky.

(48) and (49) both contain negative COMP since NPIs can be licensed in them. In these examples, the epistemic modal *must* cannot be licensed in INFL because [+NEG] in COMP prevents [+ADV] in Prop from feature-governing the modal (see below). After the raising of the modal, (48) and (49) would have the LF representations in (50) and (51) respectively:

(50) *$[_{PropP}$ must$_{[+AO]i}$ Prop$_i$ $[_{CP}$ who $C_{[+NO]}$ $[_{IP}$ t_i have killed Yuri$]]]$

(51) *$[_{PropP}$ must$_{[+AO]i}$ Prop$_i$ $[_{CP}$ if$_{[+NO]}$ $[_{IP}$ John t_i know the answer$]]]$, he is lucky

These representations involve the ill-formed configuration in (74d). They are ruled out

by the ECP.

4.2.2. Minimality in Feature Government

In this subsection, we focus on minimality effects in feature government. We begin by reviewing the relevant cases observed so far. Then we reconsider licensing of NPIs. A way to derive Linebarger's (1987) Immediate Scope Constraint from RM is suggested.

4.2.2.1. Review

Let us summarize the cases where we found minimality effects in feature government. The definition of feature government is recapitulated below:

(52) Feature Government: X feature-governs Y iff
(i) X m-commands Y
(ii) Relativized Minimality is respected

The configuration in question is the following:

(53) ... X ... Z ... Y ...

where the element X is the licenser of the element Y and Z intervenes between X and Y. Intervention in this case is defined in terms of m-command. Recall that there are four licensing features under consideration; [+Q], [+WH], [+NEG], and [+ADV]. The relevant ill-formed configurations examined in the previous chapter are schematically represented below:[21]

(54) a. * ... $[+Q]_i$ ... [+Q] ... *wh*-QP_i ...
b. * ... $[+NEG]_i$ ... [+NEG] ... NPI_i ...
c. * ... $[+ADV]_i$ ... [+ADV] ... $adverb_i$ ...
d. * ... $[+ADV]_i$ ... [+NEG] ... $adverb_i$ ...

(54a-d) are exemplified by (55-58) respectively:

(55) Osamu-wa Akiko-ni $[_{CP}$ dare-ga nani-o sita ka] hanasimasita ka?
-TOP -DAT who-NOM what-ACC did Q told Q
'Does Osamu tell Akiko who did what?'

(56) Nobody didn't see anyone.

(57) Jack will finish the job quickly.

(58) (=(16a)) *Did Frank probably beat all his opponents?

(55) can only be interpreted as a yes-no question, the *wh*-QPs having scope over the embedded clause. (56) can only mean "everyone saw someone"; the NPI *anyone* can only be associated with *not*. In (57) the adverb *quickly*, which is potentially ambiguous between event-modifying reading and process-modifying reading, can only be interpreted as process-modifying. (58) is simply ungrammatical. The LF representations for (55-58) would be as follows (in (62) *probably* is placed in-situ):

(59) $[_{CP}$ Osamu-wa Akiko-ni $[_{CP}$ $[_{IP}$ dare-ga$_i$ $[_{IP}$ x_i $[_{VP}$ nani-o$_j$ $[_{VP}$ x_j sita ka$_{[+Q]i,j}]]]]]$ hanasimasita ka$_{[+Q]}]$

(60) $[_{IP}$ nobody$_{[+NEG]}$ $[_{IP}$ x_i did $[_{VP}$ not$_{[+NEG]}$ see anyone$]]]$

(61) $[_{IP}$ Jack will$_{[+ADV]}$ $[_{VP}$ $[_{V'}$ finish$_{[+ADV]}$ the job$]$ quickly$]]$

(62) *$[_{PropP}$ Prop$_{[+ADV]i}$ $[_{CP}$ did$_{[+NEG]}$ $[_{IP}$ Frank probably $[_{VP}$ beat all his opponents$]]]]$

In (59) the [+Q] in the embedded COMP blocks feature government by the [+Q] in the matrix COMP, making the multiple *wh* question interpretation unavailable. In (60) the [+NEG] in *nobody* cannot feature-govern the NPI due to the intervening [+NEG] in *not*. In (61) the [+ADV] in V blocks feature government by the [+ADV] in INFL, which renders the event-modifying reading of *quickly* impossible. In (62) [+NEG] in COMP disrupts feature government of *probably* by [+ADV] in Prop. (48) and (49) above also involve the configuration given in (54d).

4.2.2.2. Negative Polarity Items

Let us further examine minimality effects in licensing of NPIs. As a point of departure, consider the following sentence involving the VP-idiom NPI *budge an inch* (from Linebarger (1987:337)):

(63) He didn't budge an inch because he was pushed.

Linebarger (1987) notes that (63) lacks the kind of ambiguity detected in examples like (64) (Linebarger (1987:333)):

(64) George doesn't starve his cat because he loves her.

This sentence has the two interpretations in (65):

(65) a. It's not because he loves her that George starves his cat; it's because ...
b. It's because he loves her that George doesn't starve his cat.

Thus in (65) the *because*-clause can be either within or outside of the scope of the negation. On the other hand, (63) allows the reading in (66a) but not (66b):

(66) a. CAUSE (he was pushed, NOT [he budged an inch])
'His not moving was caused by his being pushed.'
b. *NOT CAUSE (he was pushed, he budged an inch)
'His moving wasn't caused by his being pushed.'

Observing examples like (63), Linebarger (1987:338) proposes the following constraint on licensing of NPIs:

(67) The Immediate Scope Constraint (ISC)
A negative polarity item is acceptable in a sentence S if in the LF of S the subformula representing the NPI is in the immediate scope of the negation operator. An element is in the immediate scope of NOT only if (1) it occurs in a proposition that is the entire scope of NOT, and (2) within this proposition there are no logical elements intervening between it and NOT.

According to Linebarger, "logical elements" correspond roughly to propositional operators which include QPs, quantificational adverbs[22], the causal predicate lexically expressed by *because* and so forth. The ISC prohibits the LF in (66b) where the *because*-clause intervenes between the negation and the NPI.

As an illustration of the ISC besides the case of *because*-clause, Linebarger (1987) discusses the interaction among NPIs, QPs, and negation. Consider the following example (from Linebarger (1987:353)):

(68) She didn't wear any earrings to every party.

This sentence can be construed as meaning (69a) but not (69b):

(69) a. NOT $\exists x\ \forall y$ [she wore x to y] where x = earrings, y = a party
'There are no earrings that she wore to every party.'
b. *NOT $\forall y\ \exists x$ [she wore x to y] where x = earrings, y = a party
'It wasn't to every party that she wore any earrings.'

The impossibility of (69b) can be explained by the ISC. (69b) is ruled out since there is an offending universal QP between the negation and the NPI.

It was suggested above that licensing of NPIs is dictated by the general licensing mechanism, i.e., feature government. Assuming that this is on the right track, a natural question is: Can the ISC be derived from some principle to which feature government is sensitive?

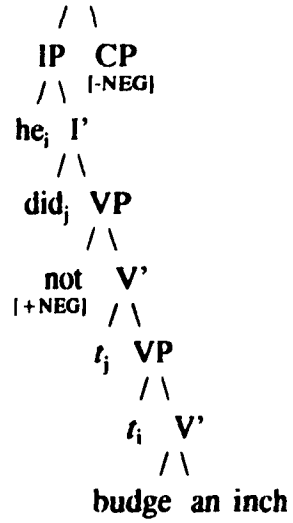
The ISC is reminiscent of the minimality principle under consideration throughout this thesis, i.e., RM. Note that RM is built into the definition of feature government (see (52)). Then it would be plausible to think that the ISC can be subsumed under RM. The idea is simply that "logical elements" as RM inducers block feature government of a NPI by negation.

But it seems that the situation is different from the typical case that RM is meant to cover. At the heart of RM is the intuitive idea that government of an element Y by an element X is blocked by an intervening element Z only if Z is a potential governor of the same kind as X for Y. "Logical elements" are not potential governors for NPIs in this sense since they do not (in most cases) license NPIs.[23]

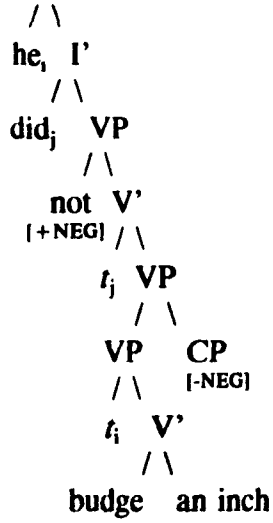
We thus cannot assign "logical elements" the feature [+NEG]. Instead, let us assign them the feature [-NEG], which represents nonnegative logical elements. Let us further assume that [+NEG] and [-NEG] are GT-compatible with each other.

Given these assumptions, NPI licensing facts in examples like (63) and (68) can be accounted for in terms of RM. The LF representations for (63) would be the following:

(70) a.



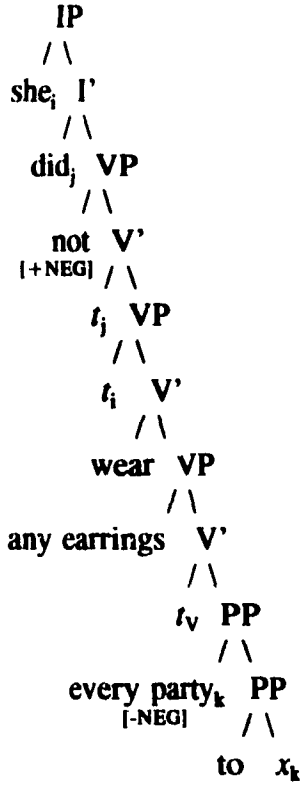
b. *



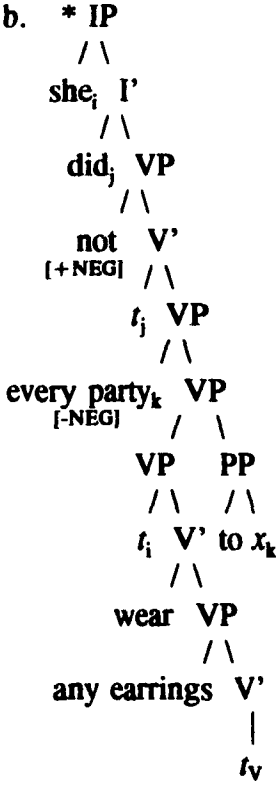
(70a) corresponds to the interpretation (66a), and (70b) to the interpretation (66b). (70a) is well-formed since the feature [+NEG] successfully feature-governs the NPI; the feature [-NEG] does not intervene between the two. (70b) is excluded since the NPI cannot be feature-governed by [+NEG] due to the intervening [-NEG] feature.

The LF representations for (68) would look like the following:

(71) a.



b. *



(71a,b) represent the readings (69a,b) respectively. (71a) is fine since the NPI is feature-governed by [+NEG]. (71b) is ruled out since there is an offending feature [-NEG] between [+NEG] and the NPI.

One may well wonder if elements other than adverbials and standard QPs show the same minimality effects. Observe the following contrast noted by Williams (1988:141):

(72) I know that people have paired off to play tennis, but
a. I don't know who anyone picked as their partner.
b. *I don't know who picked anyone as their partner.

These examples can be explained by the present account. All we need to assume is that syntactic variables, as "logical elements", have the feature [-NEG]. Given this feasible assumption, the LF representations for (72a,b) would be (73a,b) respectively:

(73) a. I do not know $[_{CP}$ who$_i$ $C_{[+NEG]}$ $[_{IP}$ anyone picked $x_{[-NEG]i}$ as their partner$]]$
b. *I do not know $[_{CP}$ who$_i$ $C_{[+NEG]}$ $[_{IP}$ $x_{[-NEG]i}$ picked anyone as their partner$]]$

In (73a) nothing prevents the NPI in subject position from being feature-governed by the negative COMP. In (73b), on the other hand, the variable with the feature [-NEG] in subject position makes licensing of the NPI impossible. Hence the difference in grammaticality in (72).[24]

4.2.3. The Notion of Government Theory Compatibility

The two preceding subsections surveyed the relevant configurations in antecedent government and feature government from the viewpoint of the present feature system. We are now in a position to try to formalize the notion of GT-compatibility. The general ungrammatical patterns that emerged are summarized below:

(74) a. * ... X[+α] ... Z[+α] ... Y ...
b. * ... X[+α] ... Z[-α] ... Y ...
c. * ... X[+AO] ... Z[+NO] ... Y ...
d. * ... X[+ADV] ... Z[+NEG] ... Y ...

Consider the relation of the features [+AO] and [+NO] in (74c). Let us assume that the feature [+NO] is a proper subset of the feature [+AO]. This is a feasible assumption, given the claim that *not* in English is in fact an adverb (Baker (1991), Ernst (1992)). Let us further assume that the feature [+NEG] is a proper subset of the feature [+ADV]. This assumption would imply that only a limited group of adverbs must be licensed by [+NEG]. Since there exist adverbial NPIs such as *ever*, *anymore*, and *yet*, the assumption would not be unreasonable. Then we can collapse (74a, c, and d) into a case where Z has a feature that is a subset of a feature of X.

Recall from Section 2.4. that if our formulation of the SP is correct (see (15)), there must be discrepancy in the relations between a *wh*-phrase and a QP with regard to GT-compatibility, i.e., a *wh*-phrase is GT-compatible with a QP but not vice versa. Consider (75a,b) and their LF representations in (76):

(75) a. What did everyone bring?
b. Who brought everything?

(76) a. $[_{CP}$ what$_i$ did $[_{IP}$ everyone$_j$ $[_{IP}$ x_j $[_{VP}$ t_j bring $x_i]]]]$

b. $[_{CP}$ who$_i$ $[_{IP}$ x_i $[_{VP}$ everything$_j$ $[_{VP}$ t_i brought $x_j]]]]$

As noted above, (75a) is ambiguous, while (75b) is not. If a *wh*-phrase were not GT-compatible with a QP, we would wrongly predict that (75a) should be unambiguous, the QP taking wide scope. This is because in (76a) not all the members of the chain headed by the *wh*-phrase c-command a member of the chain headed by the QP. If a QP were GT-compatible with a *wh*-phrase, we would expect the QP in (76b), which c-commands a trace of the *wh*-phrase, to be able to take scope over the *wh*-phrase. But this is not the case.

The hypothesis that Z is GT-compatible with X only if Z has a feature that is a subset of a feature of X helps us to capture the GT-compatibility of a *wh*-phrase with a

QP and the GT-incompatibility of a QP with a *wh*-phrase. All we need is a natural assumption that the feature [+QO] is a proper subset of the feature [+SO]. The intuition that a *wh*-phrase is a kind of QP has been expressed in the literature by the term "quasi-quantifier" (cf. Chomsky (1975) and others). Given this assumption, a *wh*-phrase is GT-compatible with a QP, but the latter is not GT-compatible with the former; the feature [+SO] is not a subset of the feature [+QO].

Taking the above hypothesis and (74b) into consideration, we put forth the following characterization of GT-compatibility:

(77) Government Theory Compatibility:
An element Z is GT-compatible with an element X only if, for F_Z, a feature of Z, and F_X, a feature of X,
(i) F_Z is a subset of F_X or
(ii) F_Z is the opposite of F_X.

One may well wonder whether or not the following configuration is well-formed:

(78) ... X[-α] ... Z[-α] ... Y ...

(78) is pertinent to antecedent government but not to feature government; [-NEG] is not a licensing feature (77) implies of course that (78) should be ill-formed. Consider (75a) again. *Wh*-phrases, QPs, and variables are all supposed to have the feature [-NEG]. If (78) is ill-formed, the chain (*what*, x_i) in (76a) should be broken at LF because of the QP (and its variable). Then (75a) is incorrectly expected to be unambiguous. Therefore, it must be the case that (78) is well-formed. In order to rule in (78), a condition of the following kind could be invoked:

(79) Condition on Government Theory Compatibility:
A negative-valued feature is visible only to a positive-valued feature.

Notice that according to (77) and (79), the configuration in (80) should be excluded:

(80) * ... X[-α] ... Z[+α] ... Y ...

We will see below that this expectation is fulfilled.

Notice also that (77) is intended to cover both antecedent government and feature

government. Thus, we are suggesting that GT-compatibility is independent of the types of government. Accordingly, we can get rid of "α" from "α-GT compatible" in RM.

Consider the status of the Visibility Condition on Relativized Minimality given in (39). Note that the condition was motivated by examples involving LF movement. Since we do not want to parameterize properties of LF for learnability reasons, (39) is assumed to be universal or part of RM.

Now, RM can be modified as in (81):

(81) Relativized Minimality: X in its canonical position α-governs Y only if there is no Z such that
(i) Z is GT-compatible with X
(ii) Z ß-commands Y and does not ß-command X.

Summarizing, we suggested that GT-compatibility is better expressed by using a set of features.[25] This characterization of GT-compatibility differs sharply from Rizzi's (1990) original characterization in that it is essentially lexical rather than syntactic. The following is a list of the features used in the above discussion ("A $\subset$ B" means "A is a proper subset of B"):

(82) a. operator features --- [+SO] (e.g. *everyone*) $\supset$ [+QO] (e.g. *who*),
[+AO] (e.g. *for this reason*) $\supset$ [+NO] (e.g. *no one*)
b. variable features --- [+SV] (e.g. variable of *everyone*), [+QV] (e.g. variable of *who*)
c. licensing features --- [+Q] (e.g. *ka* in Japanese), [+WH] (e.g. [+WH] COMP in English),
[+ADV] (e.g. INFL in English) $\supset$ [+NEG] (e.g. *not*)
d. negative feature --- [-NEG] (e.g. *because*-clause)

On an empirical level, feature-based RM can cover more data (on A'-dependencies) than Rizzi's RM, as we have seen above. In addition, introduction of features allows us to naturally express cross-linguistic variations in lexicon. Recent studies show that languages can vary substantially in their lexical inventories (cf. Aoun and Hornstein (1985)). By using features in combination, we can capture such variations. For example, the QP status of *wh*-elements in languages like Japanese can be described by

assuming that they possess the feature [+SO]. It can be assumed that Japanese *dare* 'who' and *naze* 'why', for example, have the feature specifications [+QO, +SO] and [+QO, +SO, +AO] respectively.

On a theoretical level, feature-based RM is in accord with the idea that LF is not a locus for parameterization. It is along the lines of the strong hypothesis that lexicon is the only source of cross-linguistic differences (cf. Fukui (1986)).

4.3. The Scope Principle Revisited

In Chapter 2, we dealt with three instances of Move α, i.e., scrambling, NP-movement, and *wh*-movement, and presented the SP in (83):

(83)=((15)) The Scope Principle (SP)
An operator A has scope over an operator B in case
(i) A c-commands a member of the chain containing B ($A \subseteq B$) or
(ii) all members of the chain containing A c-command a member of the chain containing B ($A \not\subseteq B$).

The aim of this section is to briefly reexamine the SP from the viewpoint of the feature-based GT-compatibility presented in the preceding section.

First, let us consider cases involving a standard QP and negation. Observe the following examples from Hornstein (1986):

(84) a. Everyone didn't like the party.
b. John didn't kiss every woman at the party.

These sentences are both unambiguous, but the scope relations between the universal QP and negation are different. In (84a) the QP takes scope over negation, whereas in (84b) negation takes scope over the QP. These interpretations can be explained by the SP. The LF representations for (84a,b) would be as follows:

(85) a. $[_{IP}$ everyone$_{[+SO]_i}$ $[_{IP}$ $x_{[+SV],[\ NEG]_i}$ did not$_{[+NEG]}$ $[_{VP}$ t_i like the party$]]]$

b. $[_{IP}$ John did not$_{[+NEG]}$ $[_{VP}$ every woman$_{[+SO],[-NEG]_i}$ $[_{VP}$ kiss $x_{[+SV]_i}$ at the party$]]]$

Notice that (85a) has the configuration (80). In (85a) the link between the variable x_i and the trace t_i is broken. This is because negation with [+NEG] is GT-compatible with the variable with [-NEG], rendering antecedent government by the latter impossible. In (85b) the QP, being GT-compatible with negation, cannot raise over *not*. If it does, a violation of the ECP will ensue. Since the standard QP and negation are GT-compatible with each other, the relative scope of the two operators in (85a,b) is calculated on the basis of the first clause of the SP. Given (85a,b), the SP accounts for the scope relations in (84a,b) in the familiar fashion. If in (85a) *not* were not GT-compatible with x_i and thus there were a link between x_i and t_i, we would wrongly predict that *not* should be able to take wide scope. Hence, we conclude that the configuration in (80) is actually ill-formed.

Turning to scope interaction of a QP and a modal-like element such as *likely*, consider the following example:

(86) Every student is likely to cheat on the exam.

As noted in Section 2.1.2., this sentence is ambiguous between the reading where the QP takes wide scope and the one where *likely* takes wide scope. The LF representation for (86) would be (87):

(87) $[_{IP}$ every student$_i$ $[_{IP}$ x_i is likely$_j$ $[_{IP}$ t_i to cheat on the exam$]]]$

Assuming the correctness of the SP, we are forced to say that a QP is GT-compatible with *likely*. If this is not the case, the QP in (87) is subject to the second clause of the SP. Then it would be expected that the QP cannot take scope over the modal since not all the members of the chain containing the former c-command a member of the chain containing the latter. How about the GT-compatibility of *likely* to a QP? Consider the following three-way ambiguous example repeated from Chapter 1:

(88) Some senator is likely to speak at every rally.
(i) $\exists$ > likely > $\forall$
(ii) likely > $\exists$ > $\forall$
(iii) likely > $\forall$ > $\exists$

In (88) *every rally* cannot take scope over *likely*. Since we decided that a QP is GT-compatible with *likely*, it must be the case that *every rally* cannot raise over *likely* at LF. If it can, we would wrongly expect the universal QP to be able to take scope over *likely*. This implies that *likely* is GT-compatible with a QP, and blocks antecedent-government by the latter (see Chapter 1, footnote 13). Thus, we hypothesize that both a QP and an operator like *likely* have the feature |+SO|. The ambiguity of (86) is consistent with this hypothesis.

Finally, consider the scope relation of an epistemic modal and *not*:

(89) John must not have done such a silly thing.

In (89) the epistemic modal *must* obligatorily takes wide scope with respect to negation. The LF representation for (89) would look like the following:

(90) $[_{PropP}$ must$_{[+AO]i}$ Prop$_{[+AO]i}$ $[_{IP}$ John $[_{VP}$ not$_{[+AO],[+NO]}$ t_i have done such a silly thing$]]]$

Recall that we are assuming with Ernst (1992) that modals in English are generated within VP and raise to INFL at S-structure. Presumably, in (90) FI forces the intermediate trace of *must* in INFL to delete. The trace t_i has already satisfied the ECP at S-structure. There is no link between Prop and t_i due to the presence of *not* disturbing the antecedent government relation between the two. Given (90), the wide scope reading of *must* in (89) is explainable by the SP.

In short, the SP in (83) in conjunction with RM can explain data other than those examined in Chapter 2.

4.4. Summary

In this chapter, an attempt has been made to seek better characterization of GT-

compatibility. We began by undermining the proviso in Rizzi's RM that adjoined elements will not enter into RM. The inner island phenomena have turned out to lend support to our version of RM. Based on the generalizations on minimality effects in antecedent government and feature government, we have proposed a notion of GT-compatibility which is defined in terms of a set of lexical features. We have also tested the SP against examples not examined in Chapter 2. It has been shown that the SP can account for a wide range of scope (non-)interaction.

FOOTNOTES TO CHAPTER 4

1. Recall that we have replaced Y in (1i) with X, as in (2i).

2. Here we are assuming that reason adverbs are adjoined to IP (though they can be adjoined to VP (see Collins (1991)). Rizzi (1990:50) explicitly assumes that reason adverbs are adjoined to TP dominated by AgrP. Note that even if we follow Rizzi in this regard, both (6a) and (7) are ruled out by RM in (1).

3. But it appears that in the case of non-feature-dependent XP adverbials, only moved ones have the feature |+AO| realized. Consider the following:

(i) It is |for this reason|$_{[+AO]_i}$ that Mary believes with all her heart that Bill was fired t_i.

If the adverbial phrase *with all her heart* has |+AO| at LF, we would incorrectly predict that (i) should be ungrammatical (see below).

4. One may be tempted to say that negative operators are GT-compatible only with certain adverbials, given Ross's (1984) observation on grammatical examples like the following:

(i) It was yesterday that I didn't go to work.

We suggest that in (i) *yesterday* can in fact originate from an IP-adjoined position. To put it differently, a derivation is available for (i) where the negation does not intervene between *yesterday* and its trace. This account predicts that examples like (i) would be ungrammatical if we ensure that the negation disrupts the chain formation, a prediction borne out, as shown in (ii):

(ii) *It is yesterday$_i$ that no one believes that I went to work t_i.

Though there seem to be further complications (see Ross (1984)), we will take a strong stand and hypothesize that all adverbial expressions are labelled |+AO|.

5. Notice that postulation of null operators in cleft constructions along the lines of Chomsky (1977) will not help Rizzi to account for (5a). If we assume with Authier (1989) that null operators adjoin to IP, they will never induce RM effects in Rizzi's system. Thus, both kinds of readings should be possible in (5a). Even if we assume that null operators move to SPEC of CP, (5a) would be expected to be ambiguous, since *no one*, by definition, would not intervene between the null operator and its variable.

6. We suspect that |+QO| and |+NO| can be identified with |+WH| and |+NEG| respectively, but for the sake of exposition, we disregard this potential redundancy in our feature system.

7. For instance, (ia) seems nonexistent, and (ib), which is well-formed, as exemplified in (ii), is amenable to the present analysis (see below):

(i) a. ... [+NO]$_i$ ... [+QO] ... t_i ...
b. ... [+QO]$_i$ ... [+NO] ... t_i ...
(ii) What didn't he fix?

8. (13) and (14) also involve the ill-formed patterns (ia) and (ib) respectively:

(i) a. * ... [+Q]$_i$ ... [+Q] ... t_i ...
b. * ... [+ADV]$_i$ ... [+ADV] ... t_i ...

This is because in (13) the QPs contain the morphemes *ka* and *mo*, and in (14) Prop and *carefully* have [+ADV] at LF. Above, we ruled out (14) with the use of (ib).

9. (17a) can also be excluded as having ill-formed (i):

(i) * ... [+ADV]$_i$ ... [+NEG] ... t_i ...

10. Jackendoff (1972:85) writes "[examples like (18)] are best with the falling intonation of a rhetorical question, which does not solicit information". But he adds that "with rising intonation, [they] seem at least marginal, and certainly better than the corresponding yes-no questions". Apart from examples such as (16a), a semantic account faces a problem posed by the following pair (the judgements are Jackendoff's (1972:85); Travis (1988) marks (ib) with "*"):

(i) a. Bill apparently has never seen anything to compare with that.
b. ??Never has Bill apparently seen anything to compare with that.

The only difference lies in preposing of the negative adverb *never*, which triggers the subject-aux inversion. The contrast in (i) suggests that the restriction in question is syntactic rather than semantic. Under the present account, (ib) is ruled out in the same way as (16a).

11. The following configuration appears to be nonexistent:

(i) ... x[+QV]$_i$ ... x[+QV] ... t_i ...

12. The exact clause-internal structure of Bulgarian is not clear to us.

13. Notice that the MBR of Aoun and Li (1989) cannot rule out superiority violations. In (34b), for instance, *who* does not qualify as a potential governor for the variable of *what* since the coindexation of the latter *wh*-operator with the former one results in a violation of Condition C.

14. Rizzi (1990) would explain the ungrammatical (33b) by stipulating that SPEC of IP can optionally count as A'-specifier (Rizzi (1990:21)). Note that this analysis does not extend to the pure superiority in (35b) because the *wh*-in-situ is within VP.

15. The same effect would be obtained by adjoining the higher VP to IP.

16. Assuming with Huang (1982) among others that Subjacency does not constrain LF movement (but see Pesetsky (1987), Nishigauchi (1990) etc.).

17. The precise clause-internal structure of Serbo-Croatian is not obvious to us.

18. (45) can be learned on the basis of positive evidence and thus does not pose a learnability problem. Note that it does not affect a RM account of *wh*-island effects in Polish-type languages.

19. In Bulgarian-type languages, the order of *wh*-phrases in SPEC of CP is fixed. For instance, orders of *wh*-phrases other than that in (i) are not allowed:

(i) Koj kogo ma kogo e pokazal?
who whom to whom has pointed out
'Who pointed out whom to whom?'

Obviously, the impossibility of the other orders cannot be explained by RM. It may be explained by assuming that (i) adjunction of *wh*-phrases is to the right in Bulgarian-type languages (Rudin (1988)), and (ii) shorter chains must be created before longer ones (Chomsky's suggestion mentioned in Mahajan (1990)).

20. This movement of epistemic modals, which are heads, to a SPEC position goes against the standard X-bar-theoretic assumption. One could assume that epistemic modals adjoin to PropP (see Baltin (1991) for the claim that heads can adjoin to maximal projections at LF). Still, we will tentatively stick to the assumption in the text.

21. The configuration in (i) is assumed to be ill-formed on conceptual grounds.

(i) * ... [+WH]$_i$ ... [+WH] ... *wh*-phrase$_i$ ...

But the *wh*-phrase can raise over the lower [+WH] and be licensed by the higher [+WH] (see (40) above).

22. We are not sure exactly what is meant by "quantificational" adverbs.

23. See Linebarger (1987) for a list of NPI licensers in English.

24. Notice that examples like (i) are correctly ruled in by our account (Linebarger (1987:337)):

(i) He didn't move because anyone pushed him.

In (i) the *because*-clause must be adjoined to VP; if it is adjoined to IP, the NPI inside it will not be m-commanded by [+NEG] in *not*. The feature [-NEG] of the *because*-clause, which we assume to percolate up to the CP node, does not intervene between [+NEG] and the NPI. This is because [-NEG] dominates and, by definition, does not m-command the NPI (see (2)). Since barriers do not block feature government (see Chapter 3), [+NEG] successfully feature-governs the NPI in (i).

25. It is interesting to note that feature-based minimality has been proposed independently by Fanselow (1991) on completely different grounds. Fanselow's (1991:232) minimality condition is provided below:

(i) Relativized feature-based minimality condition:
α cannot enter a relation of government for the feature f with ß at the level L in
$[\theta \ldots \alpha \ldots [\Sigma \ldots \delta \ldots$ ß $\ldots]]$
if Σ is a projection of δ that excludes α, and if δ governs ß for f at L (and if δ c-commands ß).

How RM advocated in this thesis and (i) can be unified is an intriguing question but must be left open for future work.

CHAPTER 5

CONCLUSION

In this thesis, we have been concerned particularly with the following questions and have tried to give them answers:

(1) a. What is the principle governing scope relations among operators?
b. What is the locality principle governing LF movement of operators?
c. What is the mechanism governing licensing of (certain) operators?

In Chapter 2 and later in Section 4.3., we attempted to answer (1a,b) by examining the relation of Move α to the scope of operators. In Chapter 3, we discussed the licensing of *wh*-elements, negative polarity items, and adverbs in order to answer (1c). The claims that have been made above in response to (1a-c) are summarized in (2a-c) respectively:

(2) a. The Scope Principle determines the relative scope of given two operators. This principle is sensitive to Government Theory Compatibility, which is built into Relativized Minimality. It also utilizes chains defined in terms of antecedent government, which is constrained by Relativized Minimality (and barriers).
b. The Empty Category Principle restricts LF movement of operators (as well as other kinds of movement in general). This principle is (partly) defined in terms of antecedent government, which is constrained by Relativized Minimality (and barriers).
c. The Feature-Dependent Item Criterion dictates licensing of (certain) operators (and supposedly other feature-dependent items). This criterion is formulated on the basis of feature government, which is constrained by Relativized Minimality.

What is remarkable in (2a-c) is the importance of Relativized Minimality in the syntax and semantics of operators. As we saw above, Relativized Minimality makes crucial use of the notion of Government Theory Compatibility. The proposals in (2a-c) naturally led us to ask the question in (3):

(3) How should Government Theory Compatibility be characterized?

In Chapter 4, we considered minimality effects in antecedent government and feature government in order to answer (3). In reply to (3), we suggested the following:

(4) Government Theory Compatibility should be defined in terms of lexical features.

Notice that the results in (2) and (4) are in accord with the working hypothesis that LF is not a locus for parameterization (cf. Higginbotham (1983)). From the viewpoint of acquisition, this is of course a welcome outcome because it is commonly held that no positive evidence pertaining to LF properties is available to children. The Scope Principle, the Empty Category Principle, and the Feature-Dependent Criterion are assumed to be universal and thus part of Universal Grammar (the latter two may be derived from an overriding principle such as the Principle of Full Interpretation). Since Relativized Minimality, Government Theory Compatibility in particular, and barriers are incorporated into these principles and criterion, they are also assumed to be part of Universal Grammar. Cross-linguistic variations in scopal and distributional behavior of operators can be ascribed to the differences in lexical inventories among languages and the S-structure parameters and conditions. The lexicon of a particular language determines Government Theory Compatibility in that language. For instance, a standard QP is GT-compatible with a *wh*-element in languages like Japanese but not in languages like English. Also, morphological requirements such as the [+WH] COMP Identification manifest themselves as cross-linguistic differences. Furthermore, parameters such as the Feature Government Parameter and conditions such as the Condition on Relativized Minimality, which constrain S-structure representations, are responsible for certain cross-linguistic variations. Our overall approach does not face a learnability problem since a child has to learn the lexicon of his/her native language anyway and S-structure properties are directly observable.

The results in (2), if they are correct, suggest that semantic interpretation (at least in core cases like those we dealt with above) is heavily dependent on (overt and covert) syntax. As shown in (2), the Empty Category Principle, antecedent government, and feature government, which are all essentially syntactic in nature, are crucial in fixing the scope of operators. This conclusion is far from novel. The direct relevance of syntax to

semantics has often been argued for and emphasized in the GB literature (cf. May (1977, 1985), Hornstein (1984) among others). But what has not been emphasized enough, we believe, is the importance of lexicon (Aoun and Hornstein (1985) is a notable exception). If the claim in (4) is on the right track, lexical properties of operators play an integral part in scope assignment via Government Theory Compatibility and Relativized Minimality. This conclusion makes intuitive sense, which might be the reason why the significance of lexicon to semantics has often been overlooked or taken for granted. In this thesis, we made concrete proposals as to how lexical properties of operators are used in assigning their scope. The general conclusion we can draw from (2) and (4) is that scope interpretation is determined to a great extent by the intricate interaction of lexicon and syntax.

Let us close our discussion by mentioning a few residual questions which are left open for further study. One important problem in the present context is: What are possible lexical features for the purpose of Government Theory Compatibility? There must be some sort of criterion that specifies the possible range of features. Otherwise, the theory would allow undesirable expansion of features. It is quite reasonable to think that the possible features are limited in number so that they are learnable.

In the above discussion, we did not examine three types of government that Rizzi's (1990) Relativized Minimality is intended to cover, i.e., antecedent government in A-movement, antecedent government in head movement, and head government. What does (4) say about these types of government? Let us first take up antecedent government in head movement. Baker and Hale (1990), discussing head movement, argue that Relativized Minimality must be made sensitive to the distinction between lexical and functional categories. Note that their claim, if correct, supports the idea in (4) since the lexical-functional distinction is in the lexicon. In order to account for the data examined in Baker and Hale (1990), all we need to do is to assume that there are features, say, [+LEX] and [+FUN], and that lexical heads and functional heads possess [+LEX] and

[+FUN] respectively.

If antecedent government in A-movement, so-called super raising in particular, is to be explained by Relativized Minimality, (4) may need modification. Perhaps it would be necessary to introduce some Case feature to rule out super raising. If, on the other hand, antecedent government in A-movement is to be accounted for in another way,[1] we can of course maintain (4) as it stands.

It is not clear if (4), as it is, can be extended to head government. According to Rizzi (1990), Relativized Minimality can capture the fact that Exceptional Case Marking cannot take place across CP; COMP, null or overt, will block government of the embedded subject by the higher verb and thus Case assignment. But we suggested above that Government Theory Compatibility is independent of the types of government. Then it would be expected that head government, as in the case of antecedent government, distinguishes lexical categories from functional categories and that COMP, being a functional head, does not block government by a verb, a lexical head. Again, if the facts surrounding Exceptional Case Marking can be handled without Relativized Minimality, we would be able to keep (4) intact.

One major conceptual question concerns the redundancies of Relativized Minimality with the other competing locality principles proposed in the literature. The redundancies between Relativized Minimality and Generalized Binding of Aoun (1985), which we will not go into, especially seem overwhelming.[2] It appears highly unlikely that grammar contains both of these principles.

Despite these and other remaining questions, we hope to have shown that Relativized Minimality in combination with Government Theory Compatibility plays a key role in various scope phenomena and is worth pursuing in future research.

FOOTNOTES TO CHAPTER 5

1. See Lasnik and Saito (1992) for such a proposal.

2. See Fanselow (1991) for much relevant discussion.

REFERENCES

Adams, M. (1984) Multiple Interrogation in Italian. *The Linguistic Review* 4, 1-27.

Aoun, J. (1985) *Generalized Binding*. Dordrecht: Foris Publications.

Aoun, J. and N. Hornstein (1985) Quantifier Types. *Linguistic Inquiry* 16, 623-636.

Aoun, J. and Y. Li (1989) Scope and Constituency. *Linguistic Inquiry* 20, 141-172.

Aoun, J. and Y. Li (to appear) *Wh*-elements in-Situ: Syntax or LF?. *Linguistic Inquiry*.

Authier, J.-M. P. (1989) Two Types of Empty Operator. *Linguistic Inquiry* 20, 117-125.

Baker, C. L. (1970) Notes on the Description of English Questions: The Role of an Abstract Question Morpheme. *Foundations of Language* 6, 197-219.

Baker, C. L. (1970) Double Negatives. *Linguistic Inquiry* 1, 169-186.

Baker, C. L. (1991) The Syntax of English *Not*: The Limits of Core Grammar. *Linguistic Inquiry* 22, 387-429.

Baker, M. (1988) *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Baker, M. and K. Hale (1990) Relativized Minimality and Pronoun Incorporation. *Linguistic Inquiry* 21, 289-297.

Baltin, R. M. (1991) Head Movement in Logical Form. *Linguistic Inquiry* 22, 225-249.

Belletti A. and L. Rizzi (1988) Psych-Verbs and θ-Theory. *Natural Language and Linguistic Theory* 6, 291-352.

Carlson, G. N. (1980) Polarity *any* Is Existential. *Linguistic Inquiry* 11, 799-804.

Chomsky, N. (1973) Conditions on Transformations. In S. R. Anderson and P. Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.

Chomsky, N. (1977) On *Wh*-Movement. In P. Culicover, T. Wasow, and A. Akmajian (eds.) *Formal Syntax*. New York: Academic Press.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Chomsky, N. (1986a) *Knowledge of Language: Its Nature, Origins, and Use*. New York: Praeger.

Chomsky, N. (1986b) *Barriers*. Cambridge: MIT Press.

Chomsky, N. (1989) Some Notes on Economy of Derivation and Representation. *MIT Working Papers in Linguistics* 10, 43-74.

Cinque, G. (1990) *Types of A'-Dependencies*. Cambridge: MIT Press.

Collins, C. (1991) Why and How Come. *MIT Working Papers in Linguistics* 15, 31-45.

Ernst, T. (1984) *Towards an Integrated Theory of Adverb Position in English*. Indiana University Linguistics Club, Bloomington, IN.

Ernst, T. (1991) On the Scope Principle. *Linguistic Inquiry* 22, 750-756.

Ernst, T. (1992) The Phrase Structure of English Negation. ms. University of Delaware, Newark, DE.

Fanselow, G. (1991) Barriers and the Theory of Binding. In H. Haider and K. Netter (eds.) *Representation and Derivation in the Theory of Grammar*. Dordrecht: Kluwer Academic Publishers.

Fiengo, R., C.-T. J. Huang, H. Lasnik and T. Reinhart (1988) The Syntax of Wh-in-Situ. In *Proceedings of WCCFL* VII, 81-98.

Fukui, N. (1986) *A Theory of Category Projection and Its Application*. Ph.D. dissertation, MIT, Cambridge, MA.

Fukui, N. and M. Speas (1986) Specifiers and Projection. *MIT Working Papers in Linguistics* 8, 128-172.

Haegeman, L. (1991) Negative Concord, Negative Heads. ms. Université de Genéve.

Higginbotham, J. (1983) Logical Form, Binding, and Nominals. *Linguistic Inquiry* 14, 395-420.

Hoji, H. (1985) *Logical Form Constraints and Configurational Structures in Japanese*. Ph.D. dissertation, University of Washington, Seatle.

Hoji, H. (1986) Scope Interpretation in Japanese and its Theoretical Implications. In *Proceedings of WCCFL* V. Stanford University, Stanford

Hoji, H., S. Miyagawa, and H. Tada (1989) NP-Movement in Japanese. ms. USC, OSU, and MIT.

Hornstein, N. (1984) *Logic as Grammar*. Cambridge: MIT Press.

Huang, C.-T. J. (1981) Move *Wh* in a Language without *Wh* Movement. *The Linguistic Review* 1, 369-416.

Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. dissertation, MIT, Cambridge, MA.

Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Jaeggli, O. (1987) ECP Effects at LF in Spanish. In D. Birdsong and J.-P. Montreuil (eds.) *Advances in Romance Linguistics*. Dordrecht: Foris Publications.

Johnson, K. (1991) Object Positions. *Natural Language and Linguistic Theory* 9, 577-636.

Kayne, R. (1983) Connectedness. *Linguistic Inquiry* 14, 223-249.

Katz, J. J. and P. M. Postal (1964) *An Integrated Theory of Linguistic Descriptions*. Cambridge: MIT Press.

Kim, S.-W. (1989) *Wh*-phrases in Korean and Japanese are QPs. *MIT Working Papers in Linguistics* 11, 119-138.

Kim, S.-W. (1991) *Chain Scope and Quantification Structure*. Ph.D. dissertation, Brandeis University, Waltham, MA.

Kim, Y. and R. Larson (1989) Scope Interpretation and the Syntax of Psych-Verbs. *Linguistic Inquiry* 20, 681-688.

Kiss, K. É. (1986) The Order and Scope of Operators in the Hungarian Sentence. In W. Abraham and S. de Meij (eds.) *Topic, Focus, and Configurationality*. Amsterdam: John Benjamins.

Kitagawa, Y. (1986) *Subjects in Japanese and English*. Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Klima, E. (1964) Negation in English. In J. Foder and J. Katz (eds.) *The Structure of Language*. Englewood Cliffs: Prentice-Hall.

Kuroda, S.-Y. (1970) Remarks on the Notion of Subject with Reference to Words like *Also*, *Even*, or *Only*, Part II. *Annual Bulletin* vol.4, Logopedics and Phoniatrics Research Institute, University of Tokyo.

Kuroda, S.-Y. (1988) Whether We Agree or not: A Comparative Syntax of English and Japanese. In W. J. Poser (ed.) *Papers from the Second International Workshop on Japanese Syntax*. Stanford: CSLI.

Laka, I. (1991) Negative Complementizers. Paper presented at NELS 22, Newark, DE.

Larson, R. (1988) On the Double Object Construction. *Linguistic Inquiry* 19, 335-392.

Lasnik, H. and M. Saito (1984) On the Nature of Proper Government. *Linguistic Inquiry* 15, 235-289.

Lasnik, H. and M. Saito (1992) *Move α: Conditions on Its Application and Output*. Cambridge: MIT Press.

Liejiong, X. (1990) Remarks on LF Movement in Chinese Questions. *Linguistics* 28, 355-382.

Linebarger, M. (1980) Polarity *Any* as an Existential Quantifier. In *Papers from the Sixteenth Regional Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago.

Linebarger, M. (1987) Negative Polarity and Grammatical Representation. *Linguistics and Philosophy* 10, 325-387.

Mahajan, A. (1990) *The A/A-bar Distinction and Movement Theory*. Ph.D. dissertation, MIT, Cambridge, MA.

May, R. (1977) *The Grammar of Quantification*. Ph.D. dissertation, MIT, Cambridge, MA.

May, R. (1985) *Logical Form: Its Structure and Derivation*. Cambridge: MIT Press.

McCloskey, J. (1979) *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. Dordrecht: Kluwer Academic Publishers.

McConnell-Ginet, S. (1982) Adverbs and Logical Form: A Linguistically Realistic Theory. *Language* 58, 144-184.

McDowell, J.P. (1987) *Assertion and Modality*. Ph.D. dissertation, University of Southern California, Los Angeles, CA.

Nakamura, M.-N. (to appear) Scrambling and Scope in Japanese. In P. Clancy (ed.) *Proceedings of the Second Southern California Japanese/Korean Linguistics Conference*. Stanford: CSLI.

Nakayama, M. and M. Koizumi (1991) Remarks on Japanese Subjects. *Lingua* 85, 303-319.

Nishigauchi, T. (1990) *Quantification in the Theory of Grammar*. Dordrecht: Kluwer Academic Publishers.

Palmer, F. R. (1990) *Modality and the English Modals*. Second Edition. London: Longman.

Pesetsky, D. (1987) *Wh*-in-Situ: Movement and Unselective Binding. In E. J. Reuland and A. ter Meulen (eds.) *The Representation of (In)definiteness*. Cambridge: MIT Press.

Pritchett, B. (1990) A Note on Scope Interaction with Definite Plural NPs. *Linguistic Inquiry* 21, 646-654.

Progovac, L. (1988) *A Binding Approach to Polarity Sensitivity*. Ph.D. dissertation, University of Southern California, Los Angeles, CA.

Rizzi, L. (1982) *Issues in Italian Syntax*. Dordrecht: Foris Publications.

Rizzi, L. (1990) *Relativized Minimality*. Cambridge: MIT Press.

Rizzi, L. (1991) Residual Verb Second and the *WH*-Criterion. ms. Universitè de Genéve.

Rochette, A. (to appear) The Selectional Properties of Adverbs. In *Papers from the Twenty Sixth Regional Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago.

Ross, J. (1967) *Constrains on Variables in Syntax*. Ph.D. dissertation, MIT, Cambridge, MA.

Ross, H. (1984) Inner Islands. In *Proceedings of the Tenth Annual Meeting of Berkeley Linguistic Society*. University of California, Berkeley.

Rudin, C. (1988) On Multiple Questions and Multiple *Wh* Fronting. *Natural Language and Linguistic Theory* 6, 445-501.

Saito, M. (1985) *Some Asymmetries in Japanese and Their Theoretical Implications*. Ph.D. dissertation, MIT, Cambridge, MA.

Saito, M. (1990) Long Distance Scrambling in Japanese. ms. University of Connecticut.

Stroik, T. (1990) Adverbs as V-Sisters. *Linguistic Inquiry* 21, 654-661.

Riemsdijk, H. van and E. Williams (1981) NP-Structure. *The Linguistic Review* 1, 171-217.

Tancredi, C. D. (1990) Not Only EVEN, but Even ONLY. ms. MIT, Cambridge, MA.

Taraldsen, T. (1986) *Som* and the Binding Theory. In L. Hellan and K. K. Christensen (eds.) *Topics in Scandinavian Syntax*. Dordrecht: Kluwer Academic Publishers.

Travis, L. (1988) The Syntax of Adverbs. *McGill Working Papers in Linguistics*, 280-310.

Travis, L. (1991) Derived Objects, Inner Aspect and the Structure of VP. Paper presented at NELS 22, Newark, DE.

Travis, L. (1992) Negation within the VP. Talk given at the Workshop on Lexical-Syntactic Relations at McGill University. Montreal, Québec.

Wahba, W. (1991) LF Movement in Iraqi Arabic. In C.-T. J. Huang and R. May (eds.) *Logical Structure and Linguistic Structure: Cross-Linguistic Perspectives*. Dordrecht: Kluwer Academic Publishers.