A THEORY OF CATEGORY PROJECTION AND ITS APPLICATIONS

by

Naoki Fukui

B.A. International Christian University
(1979)

M.A. International Christian University
(1982)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS OF THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1986

© Naoki Fukui 1986

The author hereby grants M.I.T. permission to reproduce and to distribute copies of this thesis document in whole or in part.

Signature of Author: _____
Department of Linguistics and
Philosophy. May 23, 1986

Certified by: _____
Kenneth Hale
Thesis Supervisor

Accepted by: _____
Morris Halle
Chairman, Department Committee

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

MAY 29 1986

LIBRARIES

ARCHIVES

A THEORY OF CATEGORY PROJECTION AND ITS APPLICATIONS

by

Naoki Fukui

Submitted to the Department of Linguistics and Philosophy in partial fulfillment of the requirements of the Degree of Doctor of Philosophy in Linguistics

ABSTRACT

This thesis proposes a new system of category projection where Lexical categories and Nonlexical (or "Functional") categories project in different ways, which is crucially different from the standard views in which all categories project in the same fashion.

In Chapter 1, I introduce some of the basic notions of Government-Binding Theory within which all of the discussion in this thesis takes place. The aim of Chapter 2 is to show the fundamental difference between Lexical categories and Functional categories. That is, Lexical categories have Lexical Conceptual Structures (LCS) in the sense of Hale and Keyser (1985), whereas Functional categories do not have Lexical Conceptual Structures comparable to the ones Lexical categories have, and the latter type of categories only have the function of "connecting" two syntactic units via some sort of "binding" and "agreement." Based on this fundamental difference, a new projection system is introduced, in which Lexical categories project up to a single-bar level, allowing free recursion at that level, while Functional categories can project up to a double-bar level, taking a unique specifier and a unique complement.

Chapter 3 explores various consequences of the projection system introduced in Chapter 2. One important consequence is that the proposed projection system, combined with a "bottom-up" θ-marking mechanism, predicts that the so-called "external argument" appears within the projection of a Lexical head at D-structure, receiving the external θ-role in that position, and then moves outside the Lexical projection to its S-structure position, for Case reasons. This move makes possible the explicit syntactic representation of what has been called the "implicit argument" both in noun phrases and in clauses (in the case

of passives).

In Chapter 4, I proceed to focus on Japanese and propose a new phrase structural configuration for this language in the light of the projection system introduced in Chapter 2. It is argued that Japanese lacks the Functional categories DET and COMP, and has a very defective INFL which contains no agreement features. From this, it immediately follows that Japanese has no specifiers, which close off the category projection. I argue there that this is indeed the case, i.e., Japanese has no specifiers and every phrase in this language is always "open." Other consequences of my proposal, including the derivability of overt wh movement in Japanese, are also discussed in this chapter.

Thesis Supervisor: Kenneth Hale
Title: Ferrari P. Ward Professor of Linguistics

ACKNOWLEDGMENTS

With this thesis ends my exciting four years in the MIT Linguistics Department. It is extremely difficult, almost impossible , to decide where to start and stop giving thanks to people, since so many have had an impact on me in various ways, making it possible to complete this thesis. The following list is not at all exhaustive. I hope I have not forgotten anyone who should have come among the first in the acknowledgments.

To my committee members, Ken Hale, Noam Chomsky, and Howard Lasnik, I am forever grateful (at least for the rest of my life). Ken is an ideal thesis supervisor. He gave me warm encouragement, linguistic and otherwise, whenever I most needed it. His keen insight into linguistic phenomena, whatever the language in question is, always let me have the feeling of pleasure and privilege to work with him. I don't bother to repeat incredibly many good things about Ken which have already been said in the acknowledgments of MIT theses. I just agree with all of them. Noam, although on sabbatical in my thesis year, kindly agreed to be on my committee and gave me many insightful comments. My intellectual indebtedness to him is enormous and cannot be described here. A small part of

his influence on me can hopefully be detected through the following pages. I was extremely fortunate to have Howard as a member of my committee. I thank him especially for sharing his wealth of knowledge and clarity of thought. His concern and eagerness to explore the consequences of my small ideas always made me feel what I was doing was not out of the question.

Mamoru Saito, an unofficial member of my committee, deserves special thanks. I shared with him my first two years at MIT, and the constant support he gave me during that period was invaluable. The telephone bills for the long distance calls to California were certainly worthwhile. Mamoru played a big role in my education as a theoretical syntactician.

The other faculty members at MIT challenged and encouraged me, sharing their knowledge generously. I am grateful to Jim Higginbotham, Morris Halle, Haj Ross, Luigi Rizzi, Donca Steriade, Jim Harris, Richard Kayne, Tarald Taraldsen, Jay Keyser, Wayne O'Neil, Jim McCloskey, and K.P. Mohanan for making a linguist out of me.

It is the well-known truth that the education at MIT is a collective effort and the chief educators of MIT students are their fellow students. I especially thank my classmates Andy Barss, Myung-Yoon Kang, Tova Rapoport, Betsy Sagey, and Peggy Speas for their everlasting friend-

ship. Peggy Speas deserves a special mention. Peggy and I have shared an office for the last two years and have complained to each other about every aspect of life. As for the linguistic aspect, many of the core ideas of this thesis have grown out of the discussions I have had with her in our office.

I have gained something valuable from every one of my fellow students. If I were to list everything that each person has contributed to my education, it would require another thesis. I do want to single out Kyle Johnson for thanks. His comments on my ideas and answers to my questions were always so helpful that I couldn't resist imposing on him again and again, and he always responded to my queries with generosity.

Special thanks go to the community of Japanese syntacticians: Nobuko Hasegawa, Hajime Hoji, Kazuko Inoue, Akio Kamio, Yoshihisa Kitagawa, Susumu Kuno, Kiyoshi Kurata, Kiyoko Masunaga, Shigeru Miyagawa, Masaru Nakamura, Yukio Oba, Shin Oshima, Carol Tenny, Masanobu Ueda, and Hiroshi Yamada.

Everything in this world has its starting point and my professional career as a linguist certainly had its own with two excellent teachers in Japan, Kazuko Inoue and Masatake Muraki. Thanks to Inoue-sensei and Muraki-sensei, I had the best possible education in linguistics in a quite

liberal atmosphere. I was also fortunate to have many friends with whom I could talk about linguistics without limit: Yasutomo Arai, Yasuaki Abe, the late S.-I. Harada, K.-I. Harada, Yukio Otsu, Hiroshi Yamada, Takayasu Namiki, and many others. All of these people have contributed to my education in one way or another.

My study at MIT has been made possible through a "Fulbright" fellowship by the Japan-United States Educational Commission, as well as a research/teaching assistantship from the Department of Linguistics and Philosophy at MIT, for which I am grateful.

Many thanks to my family back home; my parents Tatsuo and Masako Fukui, and my brother-in-law Akio Kaneko and my sister Yoko Kaneko, for their constant encouragement.

And finally to Meg, a special salute. She probably spent as much time and energy (or even more!) for this thesis as I did. Without her patience, encouragement, and above all love and friendship, this thesis could have never been completed.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

This thesis is a study of the system of projection, which, in various important respects, is different from the standard versions of what has been called "the X-bar theory" (Chomsky (1970, 1985), Bresnan (1977), Jackendoff (1977)). Throughout the following discussion, I will assume, as a background theoretical framework, a theory of generative grammar that has been called "Government-Binding" (GB) Theory and will presuppose that the reader has basic familiarity with this theory.[1] In this introductory chapter, I will first introduce a set of basic concepts assumed in GB Theory which are minimally necessary for the understanding of the discussion in the following chapters. I will then briefly outline the contents of each chapter of the thesis.

1.1 Basic Notions

As mentioned above, our discussion in this thesis takes place within the general theoretical framework of GB, which contains various technical notions that I will assume, in most cases, without discussion. In this section I present very cursorily, some of the more essential notions. I will do this in many cases by simply giving definitions; for illustration and discussion, I refer the reader to the above-mentioned literature and references therein.

As a starting point I will follow Lasnik and Kupin's (1977) theory of phrase markers, where phrase markers are defined in a set-theoretic way, although the adoption of this particular theoretical stance will not have a direct bearing on the discussion that follows. The basic vocabulary used in Lasnik and Kupin's (1977) system is as follows.[2]

(1) N : set of non-terminals

Σ : set of terminals

abc... : single terminals (elements of Σ)

...xyz : strings of terminals (elements of Σ^*)

ABC... : single non-terminals (elements of N)

...XYZ : strings of non-terminals (elements of N^*)

$\alpha\beta\gamma$... : single symbols (elements of $\Sigma \cup N$)

...$\chi\psi\omega$: strings of symbols (elements of $(\Sigma \cup N)^*$)

$\mathcal{A},\mathcal{B},\mathcal{C}$ ... : arbitrary sets (ordered or unordered)

(Lasnik and Kupin (1977:174-175))

We then define "monostrings" (Lasnik and Kupin (1977:176)):

(2) φ is a <u>monostring</u> with respect to the sets Σ and N if $\varphi \in \Sigma^* \cdot N \cdot \Sigma^*$

Based on the notion "monostring" just defined, we now define the basic predicates "is a*," "dominates," and "precedes" in the following way (Lasnik and Kupin (1977:176-177)).

(3) Let φ = xAz, $\varphi \in \mathcal{P}$, $\psi \in \mathcal{P}$

a. y <u>is a*</u> φ <u>in</u> $\mathcal{P}$ if xyz $\mathcal{P}$.

b. φ <u>dominates</u> ψ <u>in</u> $\mathcal{P}$ if ψ = xχz, $\chi \neq \phi$, $\chi \neq$ A

c. φ <u>precedes</u> ψ <u>in</u> $\mathcal{P}$ if y is a*φ in $\mathcal{P}$, and ψ = xyχ , $\chi \neq$z.

Lasnik and Kupin's (1977) definitions of these basic predicates in terms of the notion "monostring" (and their theory of "Reduced Phrase Markers") has various important theoretical consequences (cf. Goodall (1984)), which are largely irrelevant to our present concern. Thus, in what follows, we will freely translate their definitions to ones in terms of "nodes" rather than "monostrings."

An important qualification should be made with respect to the notion of dominance. May (1985), having investigated various properties of adjunction operations, particularly those in LF, proposes a distinction between "categories" and their "segments" (the latter term is from Chomsky (1985)). According to this distinction, a category β is assumed to consist of a sequence of nodes ("segments") (β_1, ..., β_n), where β_i immediately dominates β_{i+1}.[3] Although in most cases a category consists of only one segment, a structure of the form (4), a typical adjunction structure in which α is adjoined to β, presents a crucially different case.

(4) $[_{\beta 1}\ \alpha\ [_{\beta 2}\ \ldots\ ]]$

The distinction became relevant when the notion "dominates" is considered. May (1985) proposes a definition of "dominates" in (5) in order to ensure that α is not dominated by β in an adjunction structure such as (4).[4]

(5) α is dominated by β only if it is dominated by every segment of β.

Thus, in (4), α is not dominated by a "category" β which consists of two segments β_1 and β_2, since a segment of β, namely β_2, does not dominate α.

Whether or not May's distinction between categories and segments holds for every structure of the form (4) is an open question. I will make some suggestions concerning this problem in Chapter 3. I will also discuss the general characterization of "adjunction" in Chapter 4.

We now define some of the fundamental configurational notions of GB Theory as follows.

(6) a. c-command:

α c-commands β iff α does not dominate β and every α that dominates α dominates

b. m-command:

α m-commands β iff α does not dominate β and every γ, γ a maximal projection, that dominates α dominates β

c. government:

α governs β iff α m-commands β and there is no γ, γ a barrier for β, such that γ excludes α

(Chomsky (1985:6-7))

The term "excludes" used in the definition of "government" above is defined as follows.

(7) α excludes β if no segment of α dominates β

(Chomsky (1985:7))

To define the notion of "barrier," which is also used in the definition of "government" above, we first define "Blocking Category" (BC).

(8) For γ, a maximal projection, γ is a <u>BC</u> for β iff γ is not L-marked and γ dominates β

(Chomsky (1985:12))

Based on the notion BC, we define the concept of "barrier."

(9) γ is a <u>barrier</u> for β iff (i) or (ii):

(i) γ immediately dominates δ, δ a BC for β

(ii) γ is a BC for β, $\gamma \neq$ IP

(Chomsky (1985:12))

Here, "immediately dominates" (cf. fn. 3) is restricted to a relation between maximal projections (in the sense of the X-bar theory[5]), so that γ immediately dominates δ even if a nonmaximal projection intervenes between γ and δ. The notion of "L-marking" in (8) is defined in terms of "θ-government":

(10) α <u>θ-governs</u> β iff α is a zero-level category that θ-marks β, and α, β are sisters

(11) α <u>L-marks</u> β iff α is a lexical category that θ-governs β

(Chomsky (1985:12))

We will also assume that "government" must meet the Mimimality Condition of Chomsky (1985).

(12) The Minimality Condition

In the configuration : ... α ... [$_\gamma$ ... δ ... β ...]

α does not govern β if γ is a projection of excluding α and γ immediately dominates β

(adapted from Chomsky (1985:33-34))

The intuitive content of the Minimality Condition is that δ protects β from government by α, regardless of whether γ is a barrier for β, i.e., δ . a "closer" governor for an element β, serves to protect government from outside (cf. also Reuland (1984)).

Also of importance for the following discussion is the Binding Theory. In this thesis, we will assume the version of the Binding Theory proposed in Chomsky (1986).

(13) The Binding Theory

Suppose that we have an expression E with the indexing I, where an indexing is an association of indices with phrases of E. Let α be a category, β be a local domain, and γ be a lexical category that governs α, then:

a. I is BT-compatible with (α, β) if:

(A) α is an anaphor and is bound in β under I
(B) α is a pronominal and is free in β under I
(C) α is an r-expression and is free in β under I

b. Licensing Condition for α

For some such that (i) or (ii),
I is BT-compatible with (α, β):

(i) α is an r-expression and (a) if α heads its chain or (b) otherwise

(a) β = E
(b) β is the domain of the head of the chain of α

(ii) α is an anaphor or pronominal and β is the least CFC (="Complete Functional Complex") containing γ for which there is an indexing J BT-compatible with (α, β)

(adapted from Chomsky (1986:171-172))

CFC (="Complete Functional Complex") is roughly defined as the category, α, in which all grammatical functions compatible with α's head are realized.

We assume the following version of the ECP, although the reference to θ-government in the statement of proper government might be eliminable, as suggested by Chomsky (1985).

(14) The ECP : a nonpronominal empty category must be properly governed.

(15) α properly governs β iff α θ-governs or antecedent governs β.

(Chomsky (1985:13-14))

We also assume the basic principle of the "Bounding Theory" (cf. Chomsky (1981)) of the form such as follows:

(16) Given a chain $(\alpha_1, \ldots, \alpha_n)$:
If (α_i, α_{i+1}) is a link of a chain,
then α_{i+1} is 1-subjacent to α_i.

(17) β is n-subjacent to α iff there are less than n+1 barriers for β that excludes α.

(cf. Chomsky (1985:24))

In the following chapters, we will investigate particular details of the notions and principles briefly sketched so far, when they are applied to concrete examples. Also, other principles of grammar will be introduced as the discussion proceeds. For a schematic exposition of the principles of GB and various technical notions assumed in

that theory, we again refer the reader to the above-mentioned literature.

1.2 Outline of the Thesis

The organization of this thesis is as follows: In Chapter 2, I introduce a new system of projection in which two types of categories, Lexical categories and Nonlexical categories, project in different ways. Behind this "non-uniform" view of category projection is the observation that there is a fundamental difference between Lexical categories and Nonlexical categories, i.e., the former type of categories have "meaning" ("Lexical Conceptual Structure" in the sense of Hale and Keyser (1985), "θ-grid" in the sense of Stowell (1981), etc.), whereas the latter type of categories do not have comparable "meaning" and only have the function of "connecting" two elements via "θ-binding" (cf. Higginbotham (1985)), "selection," and "agreement." In the system of projection proposed there, Lexical categories project up to a single-bar level, allowing free recursion at that level, while Nonlexical categories can project up to a double-bar level, taking a

unique specifier and a unique complement. It is also argued in Chapter 2 that this projection system, coupled with Brame's (1981, 1982) idea that determiners are heads of "noun phrases," makes it possible to capture the basic structural parallelism between "clauses" (CP and IP) and "noun phrases," which has been pointed out in the literature but has hitherto been unable to receive a natural account.

Chapter 3 explores various consequences of the system of projection proposed in Chapter 2. Among those consequences, the most notable one is that the so-called "external argument" (Williams (1980)) can now be allowed to occur within the projection of a Lexical category. For example, in our projection system, the external argument of a verb appears within a projection of the verb at D-structure, receiving a θ-role under the strict sisterhood condition, and then moves up to the specifier of IP position in order to avoid a violation of the Case Filter. This move opens up a possibility of representing the so-called "implicit argument" (Roeper (1983, 1984)) in passives explicitly in a structural configuration. Also, the hypothesis that "external arguments" occur within a projection of a Lexical category implies that the specifier of IP position, for instance, is always an "A'-position." Thus, in our projection system, the θ/θ' distinction and A/A' distinction overlap completely. This suggests that

we can eliminate the A/A' distinction entirely from the theory of grammar. The elimination of the A/A' distinction from grammar has many consequences, especially for the treatment of "crossover" phenomena and for the status of the Binding Theory. I argue that traces of NP-movement can no longer have the status of anaphors, and suggest that the Binding Theory has nothing to do with chain-internal relations. Instances of "illicit" movement are independently excluded by other principles of grammar, such as the ECP/subjacency.

I turn to Japanese in Chapter 4 and propose a new phrase structural configuration for this language in the light of the projection system proposed in Chapter 2. Various observations, particularly with respect to the existence of the "VP" node in Japanese, that have been made in the literature are examined in the first section. I, then, examine the status of Nonlexical categories in the language and conclude that Japanese does not have the Nonlexical categories COMP and DET, and that a Nonlexical category INFL in this language is very defective having no agreement feature with it. From this, it immediately follows from our conception of the projection system that Japanese lacks specifiers. I then argue that this is indeed the case; none of the elements that have been assumed to be specifiers in Japanese have the characteristic

properties of specifiers, i.e., the function of "closing off" the category projection. Based on these conclusions, a new phrase structure for Japanese is proposed. It is shown that this proposed phrase structure is quite consistent with the facts observed in the literature which are summarized in the first section of this chapter. Some consequences of the proposal, including the derivability of the lack of overt wh movement in Japanese, are also discussed.

Notes to Chapter 1

1. For a detailed exposition of GB Theory, the reader is referred to, among others, Chomsky (1981, 1982, 1985, 1986), Riemsdijk and Williams (1986), and references cited there.

2. I will assume the basic notations used in formal grammar/automata theory. For a detailed explanation of those notations, see Hopcroft and Ullman (1979) and Lewis and Papadimitrion (1981), among others.

3. We assume the following definition of "immediately dominates."

   α <u>immediately dominates</u> β iff α dominates β and there is no γ such that γ dominates β but does not dominate α.

4. The formulation (5) is taken from Chomsky (1985:5).

5. That is, XPs. The notion of "maximal projection" in the system of projection to be proposed in this thesis will be discussed in Chapter 3.

CHAPTER 2

PROJECTION TYPES : LEXICAL VS. FUNCTIONAL CATEGORIES

This chapter introduces the basic background framework within which various analyses to be presented in the following chapters are carried out.[1] As a starting point, we focus on what we take to be a fundamental difference between lexical (N, V, etc.) and nonlexical or "Functional" categories (COMP, INFL, etc.) with respect to the way they project. Functional categories project to X" (henceforth XP, thus NP for N", IP for I", etc.), and are limited to a unique specifier position and a single complement position. By contrast, lexical categories project up to a single-bar level, X', allowing free recursion (or 'iteration' in the sense of Harris (1946, 1951) at that level, limited only by the Projection Principle and other independent licensing conditions. This amounts to rejection

of the 'uniform bar-level hypothesis' according to which the number of bars for the maximal projection is uniform across categories, and which has been assumed explicitly or implicitly in almost every work on X-bar theory (with some notable exceptions. See fn. 2) since it was first proposed in Chomsky (1970)[2]. What I would like to propose in this chapter is, then, to 'relativize' the notion of maximal projection based on the well-founded distinction between Lexical and Functional categories.

2.1 Introductory Remarks

Following Chomsky (1970, 1972), I will assume that the primitive terms of UG include the category features [± N] and [± V], and that these features allow a partition of lexical items into four categories. It is not clear to what extent the above features may be labels for some semantic or other property of the categories, but there is an important distinction between categories which bear these features and those which do not: the categories bearing these features are those which may take arguments. In the theory of Higginbotham (1985), these and only these are

the categories which have a θ-grid as part of the lexical entry. Following the longstanding tradition, I will call these four categories the Lexical Categories.

(1) Lexical Categories: [+N, -V] (noun)
[+V, -N] (verb)
[+N, +V] (adjective)
[-N, -V] (preposition)

In English at least, the Lexical categories do not exhaustively partition the set of items in the lexicon. In particular, the items such as COMP and INFL, which have been called Nonlexical Categories, act as syntactic heads but do not appear to have these features nor do they have θ-grids or "Lexical Conceptual Structures" in the sense of Hale and Keyser (1985).[3]

In the framework of GB (cf. Chomsky (1981, 1982, 1985, 1986)), the relationship between the lexical and the syntactic levels, in particular D-structure, is one of projection from the former to the latter; properties of lexical items, including θ-marking properties, are projected from the lexicon into syntax, constrained by the Projection Principle and the schematic "X-bar" well-formedness conditions on phrase markers.

(2) **The Projection Principle** (informal statement): lexical properties are maintained at all syntactic levels.

(3) **The X-bar Schema:**

(i) X' = X X"* (order irrelevant)

(ii) X" = X"* X'

where X"* stands for zero or more occurrences of some maximal projection.

(Chomsky (1985:2))

My proposal is based on several empirical observations about structure across categories. It has long been observed that the cross-categorial generalizations captured by the X-bar schema were fuzzy in certain respects; even Jackendoff (1977) resorted to some alternative features (specifically [±subject], [±object], [±comp] and [±det]) to get the generalization to work out right. Until Chomsky (1985), it was thought that the categories IP and CP (especially the latter) were defective in some way; Chomsky suggests extending the X-bar schema so that CP and IP would both have specifier positions.

In the following discussion, I will be taking the position that the determiners found in noun phrases are Functional heads, on a par with the Functional heads COMP and INFL. To the best of my knowledge, the first to

advocate such a view of determiners was Brame (1981, 1982), who developed the idea within his own theoretical framework. (Brame called determiners "head selectors.")[4] Abney (1985) argues within the framework of GB that determiners can be considered as heads of a constituent Determiner Phrase (DP). I will, in what follows, mainly discuss Abneys' observations just for the sake of exposition. This should not be confused to mean that I am ignoring Brame's pioneering work. In fact, Abney's work should best be regarded as an extension of Brame's original idea within the GB framework.

Abney points out that Functional heads are special in that they are closed-class items, that they lack the sort of semantic value associated with Lexical categories, and that they always select a unique complement. This proposal that DET, INFL and COMP[5] constitute a natural class allows parallel structures to be assigned to DP (=Determiner Phrase), IP and CP. We call this class of categories <u>Functional Categories</u>[6].

In addition to Abney's observations, I state the following observations concerning the Functional categories.

(4) (i) Functional heads have one and only one (i.e. non-iterable) specifier, while the specifiers of Lexical heads may be iterable ones.

(ii) The specifiers of Functional heads are often (in our model, always--see below) moved from within their complement.

(iii) All Functional heads can have specifier positions; it is not at all clear that all Lexical heads have specifier positions.

(iv) Languages which lack Functional heads also lack specifier positions.

In the following, I will show how these properties, as well as those observed by Abney, of the Functional categories (or the difference between Functional and Lexical categories) can receive principled explanations under the system I am proposing.

Before we proceed, let us be clear about exactly what we mean by "specifier". Chomsky (1985) emphasizes that the notion "specifier" is strictly a relational one, used as a label for whichever maximal projections happen to appear in a given category as immediate daughters of X". That is, there is no node label 'specifier', and the righthand X" which appears in the X-bar schema (3ii) above is 'relationally' defined as the 'specifier' of X', whatever the node label of the X" might be. However, this version of

the X-bar schema per se does not give us an explanation for the contrast between (5) and (6).

(5) (a) the very very old man
(b) Mary's big red book
(c) Susan never could have been eating cabbage.

(6) (a) *the the old man
(b) *yesterday's Chomsky's book.
(c) *it Mary ate a bagel.
(d) *the John's cat
(e) *every the book
(f) *what who did buy?

These data show that there are some types of "specifiers" which may iterate (e.g., very, big, have, been, etc.) and others which may not (e.g., the, Chomsky's, what, etc.). It is of course not a priori necessary under the modular approach we are assuming (cf. Chomsky (1981, 1986, etc.)' that the ill-formed examples be ruled out by X-bar theory alone. For example, cases (6a) and (6e) might be ruled out as violations of vacuous quantification (cf. Chomsky (1982)) or by some generalized version of the θ-criterion (cf. Higginbotham (1985)), and cases like (6b), (6c), and (6d) could be excluded by the Case Filter.[7] (6f), an instance of "doubly-filled COMP" effect, is probably to be

excluded as an ECP violation although it is not entirely clear how cases like this could be ruled out in terms of the ECP in the most current framework (Chomsky (1985)), which assumes wh movement to be a movement into the specifier of CP (Recall that the X-bar schema given in (3) allows the iteration of specifier position).[8] However, it is desirable to give a principled reason, in terms of the position which a given "specifier" occupies in a syntactic structure, why certain "specifiers" may iterate while others may not. The X-bar schema given in (3) above allows any number of "specifiers" in any type of category projection and hence cannot in principle capture the basic difference between the two types of "specifiers," iterable ones and non-iterable ones.

It should also be pointed out that the presence of apparent subjects across categories (cf. Stowell (1982)) does not provide evidence that each category has some unique subject position given by X-bar theory, since extraction data reveals an underlying difference in the status of the "subject" from category to category, as shown by the examples below:

(7) (a) we saw Bill's book.
(b) we saw Bill drunk.
(c) *whose did you see book?
(d) who did you see drunk?
(e) whose book did you see?
(f) *who drunk did you see?

The subject of the adjective can be extracted as in (7d), while the subject of the noun cannot (7c). The noun plus its subject can move as a constituent as in (7e), while the adjective plus its subject cannot as shown in (7f). These examples indicate that the status (or structural position) of the "subject" of the adjective drunk in (7b), (7d) and (7f) differs in some fundamental way from the status of the "subject" of the noun book in (7a), (7c) and (7e).

Based on various observations made above, I would like to propose a way of looking at how categories project, which is different from the standard X-bar theory in which every category project in the same fashion. The crucial distinction for this view is between Functional and Lexical categories: Functional categories have a unique specifier, but Lexical categories may iterate "specifiers," as long as all "specifiers" are fully licensed and can be interpreted at LF. I maintain that only the specifiers of Functional categories "close off" their projections, which

I take to be a characteristic property of specifiers, and the projection of Functional category moves up to an XP level, a "closed" category level, due to the existence of a specifier, a "closing" element. On the other hand, all projections of a Lexical category are X', since there is no inherent limit to their iteration. In order to avoid terminological confusion, I will use the term "specifier" to refer to an element that closes off a category projection. Thus, only Functional categories can have specifiers in this sense. Note incidentally that according to this definition, the 'iterated' elements in (5), for example, are <u>not</u> specifiers.

I have been assuming that the iteration is allowed in some structural position, namely at the X' level of a Lexical category's projection (cf. the possibility of iterative adjectives and pre-verbal 'auxiliary' elements exemplified in (5)). To further clarify this point, it may be helpful at this point to briefly summarize the proposed arguments (and add several new arguments) for the iteration possibility at the X' level.

The possibility of 'iteration' (or 'recursion') at the single-bar level has been noted by various linguists (cf. Harris (1946, 1951), Baker (1978), Hornstein and Lightfoot (1981), Radford (1981), etc.). The following discussion is based on Radford (1981, Chapter 3). Consider

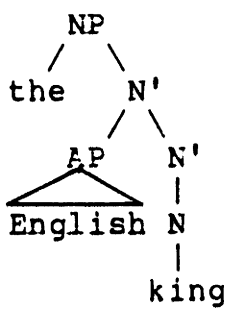
a phrase such as (8). This phrase is structurally ambiguous between the two interpretations (9a) and (9b).

(8) the English king

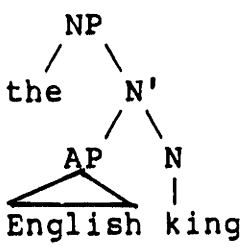
(9) a. the king who is English
b. the king of England

The interpretations (9a) and (9b) correspond to the following structures (10a) and (10b), respectively (Radford (1981:96) with adaptations).[9]

(10) a.



b.



The crucial difference between these two structures has to do with the status of king. In (10a), king has the status of N', while in (10b) it has the status of N. A piece of evidence for the postulation of these two structures can be obtained from the fact about 'one substitution.' If we assume, following Baker (1978), that the pro form one replaces uniquely an N', then it should be predicted, given the structures in (10), that the phrase the English one in (11) can only have the interpretation (12a), but can never have (12b).

(11) I like the French king, but not the English one

(Radford 1981: 96)

(12) a. the English one = the king who is English
b. the English one = the king of England

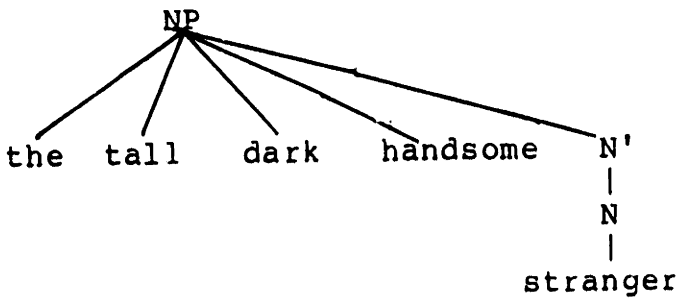
This prediction is actually borne out. The phrase the English one in (11) has only the meaning corresponding to (12a). Now if the structure (10a) is attested, as seems plausible in view of the fact about one substitution, then we have to allow the 'recursion' of N's. Another piece of evidence for the 'recursion' of N's is obtained from the possibility of phrases like (13), which is taken from

Radford(1981:104). (See also the examples in (5)).

(13) the tall, dark, handsome stranger

Given the X-bar schema in (3), the structure of (13) should appear as in (14).

(14)



The same is true for the example (8), which I deliberately ignored in the above discussion, since this possibility does not affect the argument. However, it seems that the structure (14) fails to capture the basic difference between a determiner <u>the</u> and other prenominal modifiers. Notice that in the configuration (14), a determiner <u>the</u> is exactly on a par with prenominal adjectives; they are all relationally defined as 'specifiers' of NP, if we assume them to be XPs, or if not, they are totally outside the scope of the X-bar schema in (3). Assuming for the sake

of argument that the and the prenominal adjectives fall under the X-bar schema (3), the distinction (in terms of structure) should be made between them in view of the different behaviors each exhibits with respect to ordering restrictions. Among the prenominal adjectives, the ordering restriction imposed on them is basically a semantic one (See, among others, Ziff (1960) and Martin (1968)). So, even if we change the linear order of the prenominal adjectives in (13), the resulting forms are not as bad as those which violate some syntactic constraint.

(15) a. ??the tall, handsome, dark stranger
b. ?the dark, tall, handsome stranger
c.???the dark, handsome, tall stranger
d.???the handsome, tall, dark stranger
e.???the handsome, dark, tall stranger

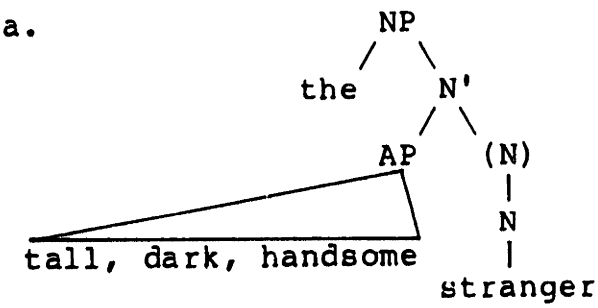
Judgments may vary concerning the relative 'oddness' between the examples in (15). But the point here is that none of the forms in (15) is as bad as the following examples in which the determiner the intervenes the prenominal adjectives.

(16) a. *tall, dark, handsome, the stranger
b. *tall, the, dark, handsome stranger
c. *dark, tall, the, handsome stranger
d. *handsome, the, tall, dark stranger
etc.

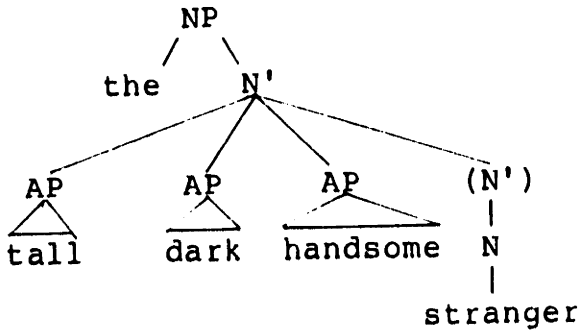
This indicates that there is a grammatical (formal) requirement that a determiner such as <u>the</u> precede (or in hierarchical terms 'be outside of the c-domain of') all the prenominal modifiers. Thus, we might conclude that the difference between determiners and other prenominal elements must be somehow syntactically represented.

Suppose that this distinction can be made by putting the determiner in a position outside of N', while putting other prenominal modifiers inside the N'. Then, there are three possible structures for a phrase like (13).[10]

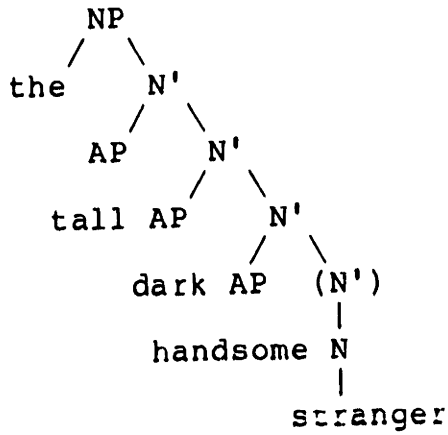
(17) a.



b.



c.



The structure (17a) is immediately precluded by the ungrammaticality of (18) (pointed out to me by Howard Lasnik).

(18) *The stranger is tall, dark, handsome.

If <u>tall, dark, handsome</u> constituted a single constituent as represented in (17a), there would be no way of accounting

for the ill-formedness of (18).

The choice between (17b) and (17c) is not straightforward. However, a binding fact about the internal structure of Japanese noun phrases gives a piece of evidence (though not decisive, see note 12) in favor of the structure (17c). Consider the following examples in which an anaphor-like element zibun and its antecedent appear in a noun phrase.

(19) a. John$_i$-no zibun$_i$-no hihan
-Gen criticism

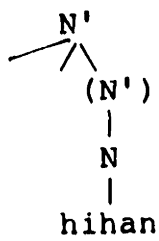
Lit. 'John$_i$'s zibun$_i$'s criticism'

b. *zibun$_i$-no John$_i$-no hihan

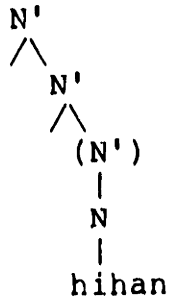
Lit. 'zibun$_i$'s John's criticism'

There are two possible structures for Japanese noun phrases like those in (19), namely (20a) and (20b), corresponding to (17b) and (17c), respectively.

(20) a.

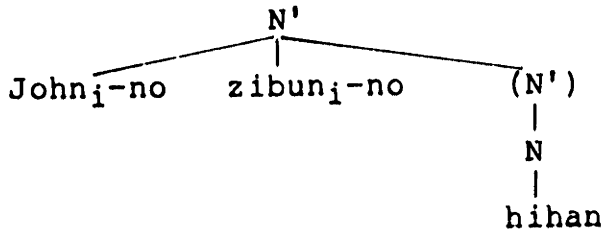


b.

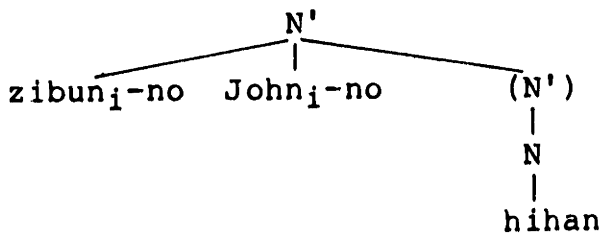


If (20a) is the correct structure, the examples in (19) have the following structure[11].

(21) a.



b.

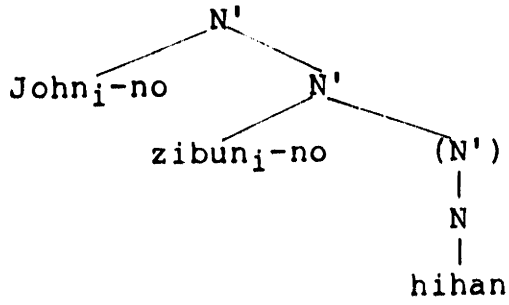


In both (21a) and (21b), a name <u>John</u> is bound by (c-commanded by and coindexed with) <u>zibun</u>, which is a direct violation of the clause (C) of the Binding Theory that requires names to be free. Therefore, the 'flat'

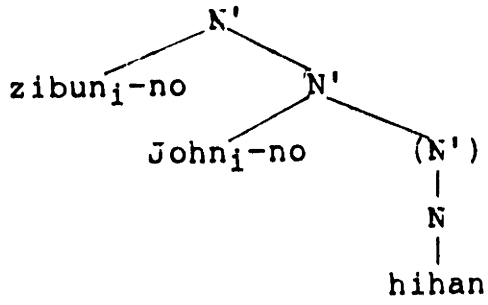
structure (20a) cannot account for the contrast in (19). It incorrectly rules out both (19a) and (19b) as a violation of the Binding Theory (C).

The structure (20b), on the other hand, accounts for the contrast without problem. The examples in (19) have the following structures under this assumption.

(22) a.



b.



In (22a), John is not c-commanded by zibun, while in (22b) it is c-commanded by zibun. Thus, we correctly predict, given the structures in (22), that (19a) is grammatical, excluding (19b) as a violation of the Binding Theory (C).

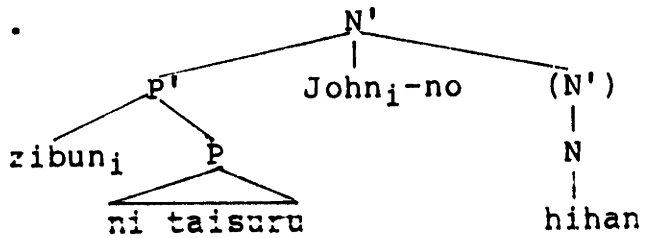
One might object here that the structure (20a) could

account for the contrast in (19) if we state the relevant condition on names in terms of precedence: A name cannot be preceded by an element it is coindexed with. This condition, which is solely based on the precedence relation, is immediately falsified by the grammaticality of the following examples.

(23) a. zibun$_i$ ni taisuru John$_i$-no hihan
toward

Lit. 'John$_i$'s criticism toward <u>zibun$_i$</u>'

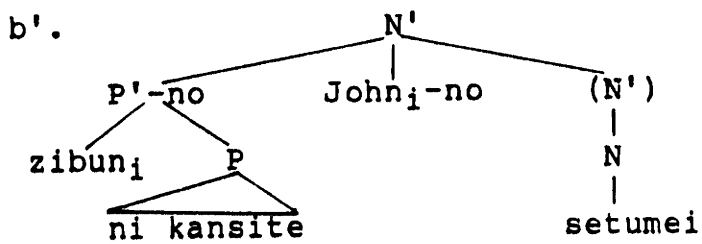
a'.



b. [zibun$_i$ ni kansite]-no John$_i$-no setumei
about explanation

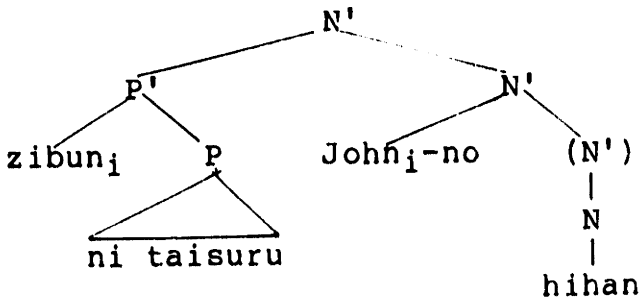
Lit. 'John$_i$'s explanation about <u>zibun$_i$</u>'

b'.

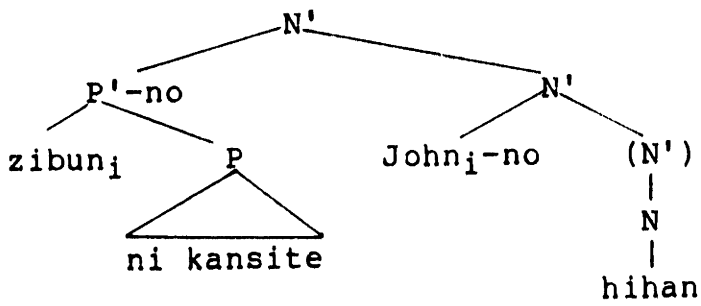


Given the 'flat' noun phrase structure (21a), the examples (23a) and (23b) should have the structures (23a') and (23b'), respectively. These structures minimally differ from the structure (21b) in that zibun does not c-command its antecedent John, due to the postpositions, in the former, whereas zibun does c-command, due to the absence of such a postposition, in the latter. Thus, in order to account for the grammaticality of the examples in (23), the incorporation of the hierarchical relation 'c-command' into the binding condition on names is unavoidable even if we assume that precedence plays a role in such a condition: A name cannot be both preceded and c-commanded by an element it is coindexed with. However, this seems to be an unnecessary roundabout of handling what can be dealt with in a straightforward fashion by the Binding Theory (C), which makes use of only c-command relation, under the 'hierarchical' structure (20b) for Japanese noun phrases.[12] The structures of (23a) and (23b) under the 'hierarchical' approach should look like (24a) and (24b), respectively. And in neither structure does zibun c-command its antecedent. Thus, the structures are ruled in, without violating the Binding Theory (C), as desired[13].

(24) a.



b.



Another piece of evidence for the 'hierarchical' structure (17c) comes from the fact about '<u>one</u> substitution' again. Consider the following examples.

(25) John bought a big expensive red car;

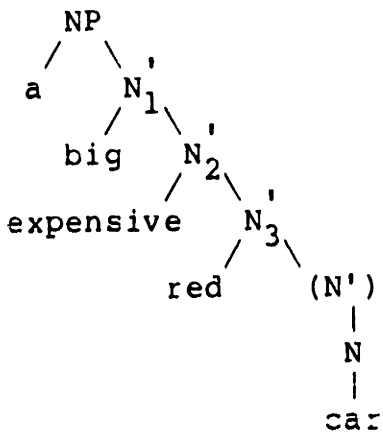
a. and Mary bought a small cheap one.

b. and Mary bought a small one.

A <u>pro</u> form <u>one</u> in (25a) means 'red car,' in addition to another possible reading in which <u>one</u> means 'car' (Recall the earlier discussion on the 'the English king' above).

And in (25b), one can mean 'expensive red car.' These facts can be accounted for straightforwardly if we assume, following Baker (1978), that one substitutes for an N', and that the noun phrase a big expensive red car has the following 'hierarchical' structure, which corresponds to (17c).

(26) a.

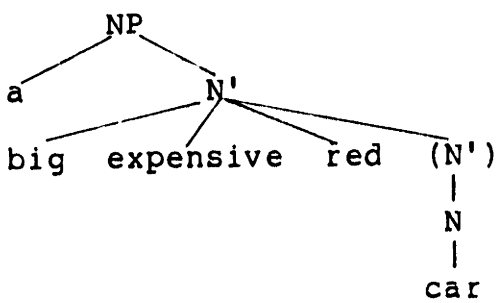


Replacing the smallest (other than the one in parentheses) N'(=N'_3) by one, we get (25a), and (25b) can be obtained if we substitute one for N'_2. Even the substitution of one for the biggest N'(N'_1) is possible. In this case, the following expression will be produced (after the application of the rule which deletes a before one, presumably for some semantic reason).

(25) c. and Mary bought one, too.

One the other hand, it seems impossible to give a coherent account for the '<u>one</u> substitution' facts, given the 'flat' noun phrase internal structure like (17b). The noun phrase <u>a big expensive red car</u> will have the structure (26b) under this approach.

(26) b.

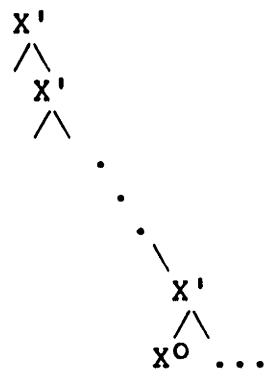


In (26b), none of the prenominal adjectives forms a constituent with the head noun <u>car</u>. Thus, the account of the '<u>one</u> substitution' phenomena under the 'flat' approach would be, to say the least, much more complicated than the one under the 'hierarchical' approach given above.

We have seen that there is good reason to assume the possibility of 'iteration' (or 'recursion') at the single-bar level of the projection of Lexical categories. We have also argued that there are at least two pieces of

evidence that the internal structure of Lexical categories, in particular noun phrases, is 'hierarchical,' rather than 'flat.' In the following discussion, we will assume that the structure of Lexical category's projection is as follows (linear order irrelevant).

(27)



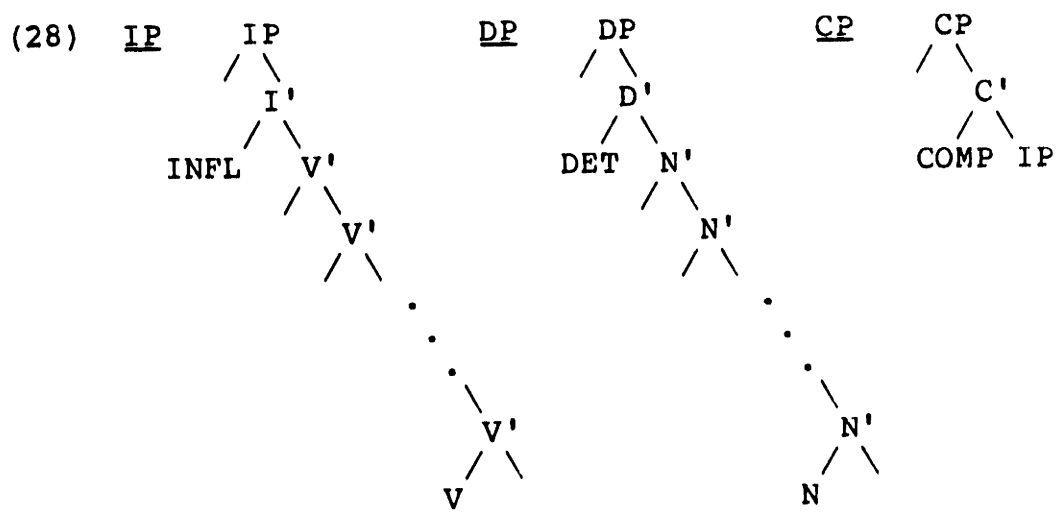
2.2 The Structure of IP and DP

The Projection Principle itself allows any number of arguments (and modifiers) of Lexical categories, as long as no violation of other principles of UG, say the θ-criterion, results and they are all fully licensed and can be interpreted at LF, as required by the Principle of Full Interpretation (Chomsky (1986)). Functional

categories, on the other hand, are restricted to have only one specifier, if any, and one complement for the reasons to be discussed below. This move captures the fact pointed out by Abney (1985) that Functional categories differ from Lexical categories in that they take unique complements. Further, it allows us to encode the distinction between iterable pseudo-specifiers and non-iterable specifiers: the elements in Lexical categories which are neither head nor complement are iterable if they meet all licensing conditions of other modules of UG, while Functional categories have a unique specifier, if any, as required by the principles to be introduced below.

Based on our discussion so far, I would like to propose the following basic schematic structures for IP, DP, and CP.

(28)



In the above structures, I am suggesting that Brame's (1981, 1982) proposal that the determiner heads a constituent DP be adopted, and I am proposing that the maximal projection of a Lexical category is X', with free 'recursion' at that level of projection. Given the structures of IP and DP in (28), it is no longer necessary (nor desirable) to say that 'subject' of V and N, i.e., 'external argument' of these Lexical heads, is present in the specifier position of IP/DP at D-structure, since there is a structural position available for the external arguments, namely one of the 'iterated' positions at the single-bar level of these categories. In fact, I will claim that the external argument starts out under the projection of Lexical categories (V'/N'/A') and then later (in English at least) is moved to a specifier position of IP/DP by Move-α for Case reason. (See Section 2.5 below).

2.3 Function Features

I adopt the standard analysis of the elements of the category INFL: i.e., that Tense/AGR assigns nominative Case, while <u>to</u> does not. I further extend this analysis,

proposing that each Functional category includes some elements which assign what I will call Function Features, or F-Features, and other elements which do not assign these features.[14] F-Features include nominative Case, assigned by Tense/AGR, genitive Case, assigned by 'S, and +WH, assigned by a WH-COMP (for the latter two cases, see below). I now introduce the term Kase to mean both Case in the standard sense (i.e., Case assigned by Lexical Categories, in particular Objective Case assigned by V) and F-Features assigned by Functional Categories.

(29) Kase = Case $\cup$ F-Features

The specifier position of a Functional category can appear only when Kase is assigned to that position. Otherwise, the projection of a Functional category stops at the single-bar level. (This is what I will later call the "Functional Projection Theorem" derived from the general principle called The Saturation Principle to which I will turn directly.) The Kase assignment which licenses the element in specifier position may come either from the Functional head itself (this would be licensing by F-Features), or, as in Exceptional Case Marking (ECM) environments, from a Lexical element (this would be licensing by Case

assignment). See below for details on ECM.

In DET position, articles are in complementary distribution with 'S, the genitive Kase assigner. Therefore, I will suppose 'S, like tensed INFL, assigns Kase, and that the, a, etc., like to, do not assign Kase. The only possible filler for specifier position of COMP is a WH-phrase,[15] so I suggest that the feature [+WH] be considered as an F-Feature, a member of the set of Kase, so that the alternation between +WH and that in COMP is parallel to the Tense(AGR)/to alternation in INFL and the 'S /determiner alternation in DET. This gives the following paradigm.

(30)

	C	I	DET
Kase assigner	WH	Tense/AGR	'S
non-Kase assigner	that	to	the

We now have a way of explaining the doubly-filled COMP effect, which, as Abney (1985) points out, seems to be parallel to the fact that determiners do not appear with other specifiers.[16] The reason that examples in (31) are

all ungrammatical is that the Functional heads underlined in these examples do not have F-Features to assign, so the pre-head position is unlicensed.

(31) a. *I wonder who <u>that</u> arrived.
b. *I think that Susan <u>to</u> leave.
c. *I enjoyed Mary <u>the</u> book.

2.4 The Saturation Principle

Notice that by associating the presence of the position of specifier of a Functional category with the presence of Kase, we are disassociating totally the existence of specifiers from the Projection Principle. This means that the "Extended" part of the Extended Projection Principle (cf. Chomsky (1982)) really has nothing to do with the Projection Principle, if the former is interpreted as a requirement that IP have a specifier position. We differ, then, from Rothstein (1983), who suggests that the requirement that the specifier of IP be filled (in English) can be explained in terms of a general requirement that

predicates must be predicated of something and thus must have subjects. In our view, this condition on predication may be true, but since saturation of a predicate takes place within the projection of a Lexical head, i.e., the external and the internal arguments are both within a projection of Lexical category at D-structure (see the discussion above), the condition on predication has nothing directly to do with the licensing of the specifier position of IP. This view of the "Extended" part of the Extended Projection Principle is further supported by the fact of "There-insertion" phenomenon and the cases of pleonastic it, since it can hardly be claimed that there is a predicational relation in any normal intuitive sense involved between these pleonastic elements and the predicate phrase. Then, what is the reason for the obligatoriness of that position in languages like English? The requirement that we adopt, which is also independently necessary in Rothstein's theory, is the following:

(32) THE SATURATION PRINCIPLE: All grids must be saturated.[17]

Here, "grids" include not only the θ-grid of a lexical entry, but also Kase grids (F-Features and Case). Thus,

the Saturation Principle collapses the θ-criterion with a requirement that if an element has a Kase to discharge, it must be discharged. Note in passing that what I am claiming here is that the existence of the specifier position of Functional categories is determined by the presence of F-Features (and Case, in the case of ECM), rather than by the existence of the "external argument" of a Lexical category. The obligatory existence of the external argument of a predicate (and the optionality of such an external argument in the case of noun phrases) is probably due to Rothstein's principle of predication. The important difference between my approach and Rothstein's is that in my system, the existence of an external argument and the existence of the specifier position are totally disassociated, whereas in Rothstein's theory, they are equivalent to each other. It seems to me that the existence of pleonastic mentioned above and the cases of "non-argument" genitive phrases (e.g., yesterday's lecture) to be discussed below provide evidence for my approach.

Higginbotham (1985), states the θ-criterion in (33).

(33) (a) Every thematic position is discharged.
(b) If X discharges a thematic role in Y, then it discharges only one.

(Higginbotham (1985:561))

As we have stated previously, Functional heads do not have θ-grids, while Lexical heads do have θ-grids. Both may have Kase grids. Notice that the assumption that lexical items have Case grids is not an innovation; in fact it is implicit in most theories of lexical representation and explicit in most studies of languages with richer overt case marking than English. (In such work, what we are calling a "Case grid" is usually called a "case array".) See, for example, Ostler (1979), Levin (1983), Nash (1980) and Simpson (1983). See also Chapter 4 for some evidence that Japanese verbs must have Case grids which are, although related, independent of their θ-grids.

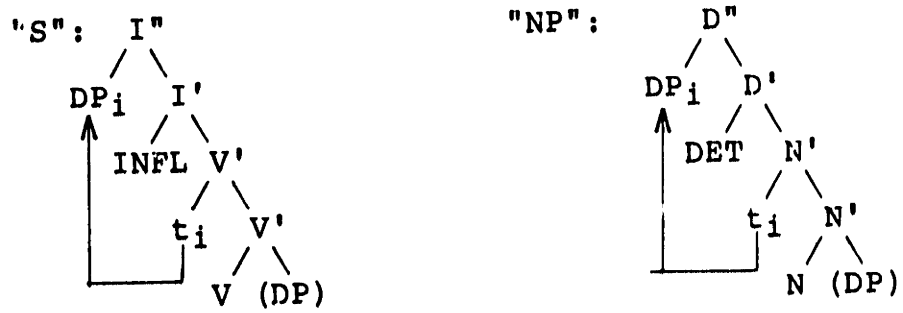
A slight modification of Higginbotham's (1985) statement of the θ-criterion gives us the appropriate Saturation Principle:

(34) (a) Every grid position is discharged.
(b) If X discharges a grid position in Y, then it discharges only one.

2.5 Deriving the Surface Order

There are several ways that ways to derive the surface order of English from the D-structures which I am proposing. I suggest adopting the standard assumption that nominative and genitive Kase are assigned leftward under government. Therefore, an external argument ('subject') of a verb, for example, must move to get Kase in order to avoid a Case (Kase, in our terms) Filter violation, because it cannot be assigned Kase in its D-structure position. This property of assigning Kase leftward extends to all Functional categories, thereby making these categories different from Lexical ones, which assign Case rightward. Under such an analysis, a movement operation parallel to that in the standard Raising cases takes place in ordinary tensed sentences and DPs[18].

(35)



An interesting difference between IP and DP is captured under this analysis: since the verb may assign structural Kase (recall that in our terms, Kase includes Case as well as F-Features) to its direct object, only the "subject" (external argument except for 'ergative' cases) may raise to get Kase from INFL (Tense/AGR) because the movement of the object to a Kase-marked position results in a violation of a condition on chains which would rule out Case (and Kase) conflict (Chomsky (1986)). Nouns, on the other hand, do not assign structural Case, according to Chomsky (1986), therefore either argument may move, and the other argument will be Kase-marked by an inserted preposition <u>of</u>.

(36) (a) $[_{DP}$the Romans$_i$ $[_{D'}$'S $[_{N'}t_i$ $[_{N'}$destruction of the city]]]]

(b) $[_{DP}$the city$_i$ $[_{D'}$'S $[_{N'}$destruction t_i by the Romans]]][19]

Since nouns do not assign structural Kase, any NP (in our system, these are actually DPs) may move to receive the Kase assigned by D, regardless of whether that NP (DP) is an argument of N. Thus, in addition to (37a) and (37b), (37c), in which an 'adjunct' is moved to a specifier of DP, is a possible option for movement within a DP.

(37) (a) the Romans' destruction by the city
(b) the city's destruction of the Romans
(c) yesterday's destruction of the city by Romans.[20]

Under our system, the "subject" of a clause is required in the specifier of IP position only by the Saturation Principle. If INFL has F-features to discharge, some DP must move to the sister of I' position so that those features may be discharged. We can also explain ECM in terms of the Saturation Principle. An ECM verb has accusative Kase to discharge, so the argument of a subordinate verb which does not otherwise get Kase, i.e. an external argument, is moved into a position where it may get that accusative Kase. Notice that another difference between Lexical and Functional heads is that Lexical heads may govern and Kase-mark into their complements, while a Functional head may not. We speculate that this difference is attributable to the directionality of F-feature assignment: the direction of F-Feature assignment (at least in English) is uniformly to the left, while the direction of Case assignment is uniformly to the right. We may further attribute this difference to the nature of F-Feature assignment by a Functional head to its specifier position and Case-assignment by a Lexical head: the former is basically an 'agreement' phenomenon, a 'SPEC-head' agreement

in the sense of Chomsky (1985), whereas the latter is based on the θ-related head-complement relation. In view of this basic difference between F-Feature assignment and Case-assignment, it is not surprising to observe that each of these processes is subject to different parametric variation, thus exhibiting different properties with respect to directionality. That Case-assignment is to the right in English is derived from the fact that English is a head initial language. That F-Feature assignment is to the left, however, does not directly derive from the head initial/final parameter. Maybe there is another parameter connected with the X-bar theory that determines the linear position of the specifier ("SPEC initial/final" parameter), thus deriving the direction of F-Feature assignment. Or, perhaps, there is a universal relation with respect to the positions of complements and specifiers, as claimed by Lightfoot (1979), that specifiers and complements are always on opposite sides of the head.[21] If this is true, then the 'leftwardness' of F-Feature assignment ("SPEC-head" agreement) is the direct consequence of the 'rightwardness' of Case-assignment of a Lexical head.

An alternative way of deriving the correct order of the subject and predicate at PF (suggested by Noam Chomsky (personal communication)) would be to assume that there is a rule of PF which fronts the subject to the specifier of

IP position, and that INFL assigns Kase (nominative Kase) to the subject within V', prior to the application of the fronting rule. One advantage of this approach is that Case-assignment (in the standard sense) may be considered to be uniformly rightward (in English), under government defined in terms of strict 'c-command', rather than 'm-command' (see Chapter 1 for definitions of these notions). Although this approach is attractive in some respects, I will not take this position in what follows for the following reasons. First of all, in view of the fundamental difference between nominative Kase assignment and objective Case assignment, i.e., the former process is an instance of agreement phenomena ('SPEC-head' agreement), whereas the latter process is an instance of 'head–complement' relation, it is not entirely clear that the integration of these processes under the name 'Case-assignment,' imposing the same conditions (rightward directionality and c-command) on both of them, is the right way to go. Secondly, this 'PF fronting' analysis seems to have some disadvantages: (i) we must assume that the PF fronting rule applies obligatorily although nothing forces it to apply, which is in conflict with the general assumption of GB that application of Move-α is optional; (ii) we must assume an equivalent PF rule within DP, which may only apply if the DET is the genitive Kase assigner

'S; and (iii) in ECM case like John believes Bill to be honest, we have to apply the PF fronting rule before Case-assignment (recall that to is not a Kase assigner) in order for Bill to get Case, which is again in direct conflict with the standard assumption that Case-assignment takes place no later than S-structure.

2.6 On the Specifier Position

Given the projection system that I just proposed, a question naturally arises as to the status of the A/A' distinction in UG. Recall that in the standard version of GB theory (Chomsky (1981, 1982, 1985, 1986), among others), there are three cases where the A/A' distinction and the θ/θ' distinction do not coincide: (i) the subject position of passive; (ii) the subject of a raising predicate; and (iii) the subject of NP. Consider, for example, the following.

(38) a. Mary$_i$ was kissed t_i
b. John$_i$ seems t_i to be honest
c. the city$_i$'s destruction t_i

The underlined positions are A-positions, since A-positions are defined as the positions in which an argument may appear in D-structure (cf. Chomsky (1981:47)), and obviously arguments appear in the specifier of IP (e.g., John kissed Mary) and the specifier of NP (e.g., Romans' destruction of the city) at D-structure. But, θ-roles are not assigned to the underlined S-structure positions in the examples of (38); rather, they are assigned at D-structure to the positions indicated by the traces coindexed with their antecedents (If θ-roles are assigned to the underlined positions, it results in a θ-criterion violation). Thus, these positions underlined in (38) are A-positions but not θ-positions, i.e., θ'-positions.

However, in the system I am proposing, the specifier of IP position and the specifier of DP (in our terms), the positions in which the underlined phrases in (38) appear at S-structure, are never filled by an argument at D-structure. Those positions are filled only by an application of Move-α (or perhaps by insertion, in the case of expletives). Therefore, A-positions are equivalent to θ-positions, and consequently A'-positions are equivalent to θ'-positions, i.e., the A/A' distinction and the θ/θ' distinction completely overlap in our system. And if these distinction overlap completely, there is no reason to postulate the A/A' distinction in addition to the θ/θ'

distinction in UG, which makes possible a significant conceptual simplification of the theory of grammar. We can eliminate the A/A' distinction entirely from the theory of grammar by replacing the reference to the A/A' distinction with the reference to the θ/θ' distinction. For example, the Binding Theory would be a theory of "θ-binding," rather than that of "A-binding." Notice that in the system without the A/A' distinction, traces are 'uniform' in the sense that they are all θ'-bound, since there is no empty θ-position at D-structure and consequently there is no movement to a θ-position. And nothing else can distinguish different types of traces in our system. Specifically, we cannot distinguish between "A-bound" and "A'-bound" traces, simply because there is no distinction between A-positions and A'-positions. Thus, the difference between 'variable' traces (=traces of wh movement) and 'anaphor' traces (=traces of NP movement, these are actually not 'anaphors' in the standard sense. See the discussion below) is minimal (but crucial): the former is operator-bound but the latter is not (cf. Chomsky (1982)). Note that if this is the correct approach, it has an important implication for the status of the Binding Theory in relation to traces, i.e., the Binding Theory, conceived as a theory of "θ-binding" would have nothing to do with the traces of NP-movement, which have been assumed in the

standard literature to be subject to the Binding Theory (A). This is because all traces are θ'-bound and no A/A' distinction exists in the system I am proposing. Therefore, for example, the "super-raising" case like *John seems that it is certain t to win (cf. Chomsky (1985)) cannot be accounted for by the Binding Theory, but rather, should be handled by some other principle of grammar, perhaps by the ECP; the intermediate CP dominating that it is ... is a barrier for the trace, by inheritance from IP. This move seems desirable on conceptual grounds since it eliminates certain redundancies between the modules of grammar, namely, the ones between the Binding Theory and the conditions on chains (ECP/subjacency), making the former irrelevant as a condition on chain links. Unless the case arises where only the Binding Theory can rule out an illicit movement, the move suggested in the present discussion seems to be supported in view of this conceptual advantage. In fact, it seems to me that all the cases where a violation of the Binding Theory (A) is involved also involve a violation of some other principle of grammar. For example, the "super-raising" cases like the one mentioned above arguably involves a violation of the ECP; cases like *$John_i$ is believed that Mary likes t_i involves a violation of the chain condition (the chain ($John_i$, t_i) contains two Case positions); etc. Also, as

Howard Lasnik pointed out to me (cf. also Lasnik (1985)), there are cases where only chain conditions can rule out the illicit movement, e.g., *John$_i$ seems that Mary regrets his$_i$ belief t_i to be intelligent (no Binding Theory (A) violation with the designated coindexing). Thus, it is clear that the Binding Theory alone does not suffice to handle the distributional restriction on traces and that chain conditions can handle the cases (as far as I know) that can be (redundantly) accounted for by the Binding Theory. This situation seems to indicate that the direction implied by the system I am proposing is right, but we should not hasten to draw a definite conclusion in this regard until all the relevant cases have been subjected to exhaustive scrutiny.

One immediate problem for the theory of grammar without the A/A' distinction would be a treatment of the so-called 'crossover' phenomena. Consider the following contrast.

(39) a. John$_i$ seems to his$_i$ friends to be t_i intelligent.
b.*?who$_i$ does it seem to his$_i$ friends that Susan likes t_i.

In the theory with the A/A' distinction, the presence of 'weak crossover' effect (cf. Postal (1971), Wasow (1972))

in (39b) and the lack of such effect in (39a) can be reduced to the positional difference between $\underline{\text{John}}_i$ and $\underline{\text{who}}_i$: In (39a), $\underline{\text{John}}_i$, the antecedent of the trace t_i, is in an A-position, whereas $\underline{\text{who}}_i$ in (39b) occupies an A'-position. Thus, the contrast in (39) is accounted for if we claim that crossover effect is obtained only when the antecedent of the relevant trace is in an A'-position. On the other hand, in the theory without the A/A' distinction, such an account would not be possible, since the position of $\underline{\text{John}}_i$ (the specifier of IP) and the position of $\underline{\text{who}}_i$ (the specifier of CP) in (39) are indistinguishable by means of the θ/θ' distinction. They are both θ'-positions and we have no additional A/A' distinction. One possible way to overcome this difficulty is to make reference to the "content" of the antecedent, rather than to its position (cf. Barss (forthcoming) for a similar approach), and to say that only a movement of an 'operator-like' element (e.g. wh phrases and quantifiers) invokes a crossover effect. Then, the contrast in (39) can be attributed to the difference between John and who: John is a name and is not an 'operator-like' element, whereas who is clearly an 'operator-like' element. Hence, a crossover effect is invoked in (39b) which involves a movement of who, but (39a) does not exhibit the crossover effect since what is moved (crossing over the pronoun his)

non-operator. Notice that the reference to the 'content' of a moved phrase is necessary even in the theory with the A/A' distinction. Consider the following examples.

(40) a. John$_i$, his$_i$ mother loves $\underline{t}_i$
b.*?who$_i$ does his$_i$ mother love $\underline{t}_i$

In a topicalization example (40a), it can arguably be said that the landing site of a topicalized phrase is an A'-position, since it is inconceivable that some argument appears at D-structure in that position. (40b) is a regular wh movement case in which who$_i$ is moved to an A'-position, exhibiting weak crossover effect. This fact indicates that the reference to the 'content' of the moved phrase is necessary even in the theory with the A/A' distinction. Therefore, making reference to the 'content' of the moved element is not the price only the theory without the A/A' distinction has to pay.

The account of crossover facts based on the 'content' of the moved phrase cannot handle all the relevant cases, however. Consider the following examples (provided by Ken Hale (personal communication)).

(41) a. everyon$_i$ seems to his$_i$ friends to have been treated t_i badly
b. John told me who$_i$ seemed to his$_i$ friends to have been treated t_i badly

Both (41a) and (41b) involve a movement of an 'operator-like' element; in the former, a quantifier <u>everyone</u>$_i$ is moved by raising, and in the latter, a <u>wh</u> element <u>who</u>$_i$ is moved by raising and subsequently by <u>wh</u> movement. If the 'content' of the moved phrase is the only relevant factor to invoke crossover effect, then these exmaples should exhibit the effect of weak crossover. The fact is that neither of the examples in (41) exhibits crossover effect. The theory with the A/A' distinction can account for the lack of crossover effect in (41) by saying that the relevant movement in these examples is an A-movement (raising) and that A-movement simply does not invoke crossover effect. In the theory without the A/A' distinction, however, it is impossible to make an account based on the A/A' distinction because there is no such a distinction. Then, what is the possible way of handling the examples in (41) in the theory without the A/A' distinction I am proposing? Modifying slightly the suggestion of Ken Hale's, we might say that the relevant factor distinguishing the cases where crossover effect is invoked and

those in which no such effect is observed is the type of the Kase assigned to the position to which an element is moved.[22] That is, if the Kase assigned to that position is a Case (in the standard sense, i.e., objective Case, nominative Case, etc.), then no crossover effect is observed. This accounts for the lack of crossover effect in the examples of (41). In both (41a) and (41b), the antecedent of the relevant trace (<u>everyone</u>$_i$ in (41a) and <u>who</u>$_i$ in (41b) (But see footnote 23 for the latter case)) is assigned nominative Case. Thus, by the above condition, no corssover effect is invoked in (41). This account can be extended to the contrast in (39). In (39a), <u>John</u>$_i$, the antecedent of t_i, is assigned nominative Case; hence no crossover effect. On the other hand, in (39b), the position of <u>who</u>$_i$ gets an F-Feature [+WH] but does not get any Case, therefore a crossover effect can be invoked (and in fact must be invoked because the moved element is a <u>wh</u> phrase. See below).

Maintaining the reference to the 'content' of the moved phrase, we have the following descriptive characterization of the crossover phenomena (I am restricting my attention to the weak crossover case here) in the theory without the A/A' distinction.

(42) In a configuration [X_i [... pronoun$_i$ ... t_i]], where neither the pronoun nor the trace t_i c-commands the other (weak crossover), the crossover effect is observed iff

(i) Case is not assigned to the position of X_i, and

(ii) X_i is an 'operator-like' element (e.g., wh elements or quantifiers)

The possibility of eliminating (42ii), i.e. the reference to the 'content' of the moved phrase, depends on the analysis of topicalization. If topicalization is analyzed as an instance of adjunction (perhaps to IP, see Baltin (1982)), then it is possible that no Kase is assigned to the position to which a topicalized element is moved. Thus, all the 'crossover' cases we have been considering so far can be divided into three cases as summarized below.

(43)

the position of X_i	Kase	crossover effect	
specifier of IP	Case	No	(raising)
specifier of CP	wh	Yes	(wh movement)
adjoined	none	No	(topicalization)

Then, it becomes possible to collapse the clauses (42i) and (42ii), eliminating the reference to the 'content' of the moved element, as follows.

(42') ... iff Case is not assigned (i.e., discharged) to the position of X_i.

On the other hand, if topicalization is analyzed as a substitution to, say, the specifier of CP (cf. Chomsky (1977b)), then some Kase must be assigned (discharged) to that position (Recall that in the system I am proposing, a specifier position is licensed only by the existence of Kase which is discharged to that specifier position). Call this feature (Kase) a "Topic" feature, a member of the F-Feature set, and assume that some empty COMP bears this feature just like certain empty COMP has the feature [+WH]. Then, the facts concerning (weak) crossover phenomena we have discussed so far is summarized as follows.

(44)

the position of X_i	Kase	crossover effect	
specifier of IP	Case	No	(raising)
specifier of CP	wh	Yes	(wh movement)
specifier of CP	Topic	No	(topicalization)

If (44) is the correct characterization, then (42') cannot be maintained, since in topicalization case, Case is not assigned to the position of X_i ("Topic" is an F-Feature, but not a Case), but the corresponding sentence does not exhibit the crossover effect (cf. (40b)). Instead, the relevant condition should be something like:

(42") ... iff wh Feature is assigned (i.e. discharged) to the position of X_i.

(42") eliminates clearly the redundancy inherent to the conditions (i) and (ii) in (42), and thus simplifying the condition on (weak) crossover, but it is not clear whether (42") truly eliminates the reference to the 'content' of the moved element. Thus, the condition (42') seems to be more desirable than (42") on conceptual grounds. But, the issue here is clearly an empirical one, depending on the

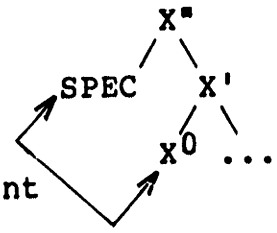
analysis of topicalization. I just leave open the choice between (43) and (44), and hence between (42') and (42"), waiting for further research on topicalization.

The preceding discussion was not intended to provide a fully adequate account of the (weak) crossover phenomena within the theory which lacks the A/A' distinction. But it has been shown, I belive, that a rather reasonable account can be given to (weak) crossover cases even if we do not postulate the A/A' distinction, and consequently that the (weak) crossover phenomena do not cast a critical doubt on the approach I am proposing, at least no more than they do on the standard approach.[23] If the move suggested by the model I am proposing in this chapter is on the right track, then we have a theory of grammar without the A/A' distinction and hence with traces whose distribution is constrained only by conditions on chains (ECP/subjacency) and has nothing to do with the Binding Theory, eliminating the redundancy we now have in the standard GB theory between the conditions on chains and the Binding Theory.

Let us now turn our attention to the status of the specifier position, in particular, its uniqueness. Why should it be the case that the Functional categories can have one and only one specifier position? Why not two, five, any odd number, etc.? While my answer to this

question at this point can be little more than speculation, I can make some suggestion. Recall that I pointed out above (2.3) that the specifier position of a Functional category is licensed only when Kase is discharged to that position. Implicit in this claim is the assumption that no (non-head) position in syntactic structure can be licensed by some well-formedness condition like the X-bar schema alone. Thus, for some syntactic entity to be present in syntactic structure, it is necessary that the entity be licensed by some syntactic relation. Suppose now that the relation between the specifier position of a Functional category and its Functional head is basically an agreement relation (the "SPEC-head" agreement in the sense of Chomsky (1985), see above), and that no other relation can possibly hold between these two positions.[24]

(45)



agreement

X^0 : Functional head

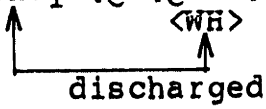
Then, we have the following condition on the projection of Functional categories as a direct consequence of the Saturation Principle introduced in 2.4 above.

(46) **FUNCTIONAL PROJECTION THEOREM:**

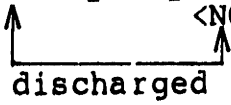
A Functional head projects to the X" level if and only if there is Kase to be discharged to its specifier position. (Otherwise, it projects only to X')

The relevant Kase to be discharged to the specifier position of a Functional category comes either from a Functional head as in (47) or from a higher verb as in the ECM case (48), which is quite consistent with Chomsky's (1985) conclusion that "if α governs β, it also governs the specifier and head of β; in particular, that CP is not an absolute barrier to government" (Chomsky (1985:9)).[25] To illustrate, consider the following representations (Kase grid is represented as "<Kase>") (irrelevant details omitted):

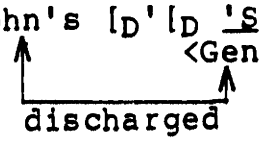
(47) a. [$_{CP}$ who$_i$ [$_{C'}$[$_C$ WH] did you see t_i]]



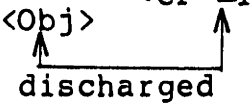
b. I think that [$_{IP}$ John [$_{I'}$ [$_I$ Tense/AGR] see (=saw) Mary yesterday]]



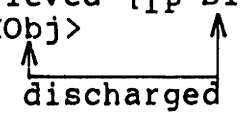
c. [$_{DP}$ John's [$_{D'}$[$_D$ 'S pictures]]] are on sale



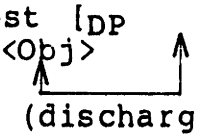
(48) a. who$_i$ do you believe [$_{CP}$ t_i [$_{C'}$[$_C$ that] John kissed t_i]]]



b. I believed [$_{IP}$ Bill [$_{I'}$[$_I$ to] be intelligent]]]



c. I lost [$_{DP}$ [$_{D'}$[$_D$ an] interesting book]] last night



A few remarks on the above quote from Chomsky (1985) are in order. First, it should be the case that the government of a specifier position from outside of its own projection is always an "exceptional" phenomenon.[26] (Recall that ECM is "Exceptional" Case Marking). Such government (and Case assignment) is marked in that it is possible only when there is no Kase to be discharged to the specifier position within the maximal projection containing that specifier position. That is, the unmarked Kase discharge within a single maximal projection has priority over the marked one crossing the maximal projection boundary. Thus, the assignment of Objective Kase (Case) by the verb to the specifier position of a DP is blocked when the Functional head of the DP is a Kase assigner.

(49) V [DP ____ [D' [D 'S] ...]]
<Obj> <Gen>
discharged discharged

(50) a. I read { John's / his } book last night.

b. *I read { John / him } book last night.

Secondly, we should ask at this point why the exceptional Case marking comparable to that in the IP case (48b) is not possible in the DP case (cf. (48c)). Namely, why is (51) ungrammatical?

(51) *I lost [DP John [D'[D an] interesting book]] last night

Apparently, the fundamental difference between (48b) and (51) is that in (48b), I' (to be intelligent) is not a possible Case receiver, whereas in (51), D' (an interesting book) not only is a possible Case receiver but also is to be obligatorily Case-marked. Therefore, if the Case (Kase)-grid of the verb lost were totally 'discharged' being assigned to John in (51), it would result in a violation of the Case Filter (Chomsky (1981)), with the D'

(an interesting book) Caseless.[27] While the problem as to why I' (and C', perhaps) does not need Case whereas D' must be assigned Case still remains unanswered, I tentatively adopt the above account of the impossibility of the ECM into the specifier of DP, leaving the principled explanation for the asymmetry between IP (and CP) and DP in this regard for future research.[28] And if a DP does not allow ECM from outside as indicated by the ungrammaticality of (51), then the specifier of the DP is present only when the Functional head is a Kase-assigner; otherwise, the projection of D is always D' (unlike that of I or C, which allows Case-assignment from outside). Thus, the structure of (48c) should be:

(48) c'. I lost <Obj> [D' [D an] interesting book] last night

discharged

Assuming the discussion so far, let us go back to the original problem, i.e., the uniqueness of the specifier of a Functional category. Putting aside the ECM cases, the problem can be restated as follows, given the Functional Projection Theorem.

(52) Why is it that Functional heads can have one and only one Kase?

Recall that I suggested that the relation between a Functional head and its specifier is basically an agreement relation (the "SPEC-head" agreement in the sense of Chomsky (1985)). Then, the uniqueness of Kase per Functional head should be a reflection of more general phenomenon, i.e., the "one-to-one" status of agreement phenomenon. I state this property of agreement as:

(53) If X, a Functional head, agrees with Y, then there is no Z such that Z ≠ Y and X agrees with Z.

It is not clear at this point whether the property of agreement stated in (53) can be deduced from some other more general property of UG. Note, however, that it might be possible to extend this "uniqueness" requirement to Kase-assignment (including Object Case assignment) in general (modulo the so-called "double object" construction).[29] If this turns out to be true, then the "bijective" relationship observed in agreement phenomenon can be an instance of the general property of Kase-assignment.

Let us now discuss briefly the difference between Functional categories and Lexical categories with respect to the possibility of iterating their "specifiers." As we observed before, Lexical categories allow 'free recursion' at the single-bar level, whereas Functional categories do not allow such recursion but rather have one and only one (if any) specifier. It is now clear that this difference is rooted in a fundamental difference between the relationship between a Functional head and its specifier and that between a Lexical head and its iterated "specifiers." The former relationship is, as we have seen, necessarily an instance of agreement ("SPEC-head" agreement), which is (perhaps universally) required to be one-to-one, whereas the latter relationship is not an instance of agreement but one of "modificational relation." And it is known that modificational relation is not limited to one-to-one as long as each modificational relation is appropriately interpreted, i.e., licensed, in LF.

Notes to Chapter 2

1. This chapter is a substantially developed version of a work I am undertaking in collaboration with Margaret Speas. A somewhat condensed presentation of the system to be introduced in this chapter is to appear as Fukui and Speas (forthcoming).

2. The most explicit statement of this assumption can be found in Jackendoff's (1977) 'uniform three level hypothesis.' The assumption that the number of bars for maximal projections is uniform across categories is kept intact even in the most recent version of X-bar theory (Chomsky (1985)). To the best of my knowledge, George (1980) first casts doubt on this assumption for the reasons quite different from ours. See also diSchiullo (1980) and Emonds (1985).

3. There have been various proposals in the literature that INFL weakly bears these features in one way or another, but even these proposals have not attributed a θ-grid to INFL.

4. Reuland (1984) also proposes that noun phrases, especially gerunds, contain an INFL-like element. Aoun (1982), too recognizes the INFL-like nature of determiners.

5. Abney (1985) also considers preposition to be a Functional category. However, the status of preposition as a (pure) Functional category does not seem to me to be entirely clear as other three Functional categories.

6. Abney (1985) calls them "Functors." I do not adopt this term since the same name is used in a different sense in the framework of Montague Grammar.

7. It is not clear, however, how examples like (6d) can be ruled out by the Case Filter, since the NP status of the is quite dubious and the Case Filter applies only to NPs. Vacuous quantification explanation cannot handle this case because John is not a quantifier.

8. Incidentally, the Saturation Principle to be proposed later in this chapter gives a natural explanation for these cases by making use of the notion of "Kase discharge". See below.

9. Radford gives the node AP to English. Whether English should be labeled as Adjective or AP (or A') is immaterial to our discussion here.

10. Again, AP's in (17) may well be Adjectives, rather than phrasal projections of Adjective. This is irrelevant

to our present discussion. Cf. the previous footnote.

11. The existence of N' node dominating a noun hihan is not relevant to the present discussion. I am assuming that no 'Gen' is not a postposition having the node label P, but rather is adjoined to a noun phrase as a Case-realizer. I am also assuming that there is no NP (=N") node in Japanese and that N' is a 'maximal projection' of N in this language. These points will be discussed in detail in Chapter 4.

12. At least in this particular case. The general significance of the precedence relation for the Binding Theory is another matter. See Lasnik (1976), Kuno (1986), and Barss and Lasnik (1986) for relevant discussion.

13. There is a problem for the 'hierarchical' approach. That is, how is it possible that zibun is c-commanded by its antecedent in (24)? (The c-command requirement on zibun is widely discussed. See, among others, Inoue (1976) and McCawley (1976)). There are two possible ways to deal with this problem. First one is to formulate the Binding Theory (A) in terms of 'm-command' (rather than 'c-command') so that the antecedent John can c-command zibun (For the definitions of 'c-command' and 'm-command', see Chapter 1. Cf. also Chomsky (1985)). This seems to be a wrong move, since it is generally observed (cf. Chomsky (1985)) that the Binding Theory makes use of the

stricter structural relation 'c-command,' although the looser relation 'm-command' plays a role in the case of government. Another possible way of handling this problem is to derive the structures in (24) by an application of a noun phrase internal scrambling, and to invoke a 'chain binding' (Barss (1984)) or 'reconstruction' Riemsdijk and Williams (1981), Hornstein (1984), among others) mechanism, parallel to the way English sentences such as which picture of himself$_i$ does John$_i$ like most?, himself$_i$, John$_i$ hates, etc. are treated in the literature. This approach seems to be on the right track, although I have no strong argument supporting this analysis. But to the extent that reasonable treatment of the c-command requirement of zibun in (24) is possible along these lines without appealing to precedence relation at least in the core cases, the argument for the 'hierarchical' structure for noun phrases in the text goes through.

14. This may be subject to some parametrization. For example, it is possible that in some language a Functional category φ_i includes only F-Feature assigner, whereas another Functional category φ_j includes only non-F-Feature assigner, while in some other language the situation is opposite, etc.

15. And empty operators. I assume that empty operators are also licensed by abstract [+WH] feature.

16. An immediate problem with this approach to the doubly-filled COMP effect is how to deal with the languages that do not exhibit such an effect (e.g., Polish). Although it might be possible to handle those cases by parametrizing the F-Feature assignment across languages, I leave this problem open here, pending future research.

17. We extend the use of the term 'saturated' here to mean : a grid X is saturated iff every position in X is discharged. Since Higginbotham (1985) uses the term 'saturated' for constituents ('a constituent such that every role in its associated grid is discharged is saturated' (Higginbotham (1985:561)), and does not use it for 'grids,' the Saturation Principle in the text should be stated as 'Every position in a grid must be discharged,' if we strictly follow his terminology.

18. The spirit of this "subject raising" in the clausal case can be traced back to Fillmore's (1968) "subject-ivalization" rule in the framework of Case Grammar, and McCawley's (1970) proposal that English is underlyingly VSO. Within the GB framework, similar proposals have been made by various people. Ken Hale suggested the idea in his classnotes at MIT in 1978. Fukui (1984) and Lumsden (1985) have suggested that the subject of a clause should be considered as an A' position. Kitagawa (1984), Koopman and Sportiche (1985, 1986), Kuroda (1985) and Johnson (1985) have independently proposed subject-raising analyses, but in orientations quite different from ours. See Koopman and Sportiche

(1986) for some arguments for the "subject raising" rule.

19. We ignore here the structural position of an agentive by phrase. For reasons related to θ-marking, it should be outside of the minimal N' (in a DP case). In fact, by phrase (and any other "adjunct" (or modifier) phrases) should be "higher" than all arguments (including "external" argument). See the discussion in Chapter 3.

20. See Anderson (1984) and Larson (1985) for suggested accounts of the apparent caselessness of certain NP adverbs in phrases like the destruction of the city yesterday.

21. Hawkins (1982) presents some counterexamples to Lightfoot's claim that specifiers and complements are always on opposite sides with respect to the head. However, we should be careful about accepting his 'counterexamples,' since all of his examples have to do with the relative ordering of determiners/demonstratives and the complements of a noun. Notice that in the system I am proposing, these elements (determiners/demonstratives) are not specifiers. More detailed crosslinguistic studies in the light of our system should be done to evaluate Lightfoot's (1979) proposal.

22. More precisely put, the type of the Kase assigned to the position which immediately precedes the relevant trace in the chain formed by the movement in question. For example, in (41b), the chain formed by the movement of $\underline{\text{who}}_i$ is : C = ($\underline{\text{who}}_i$, t_i^1, t_i^2), where $\underline{t}_i^2$ is left by the application of raising, and $\underline{t}_i^1$ is created by the application of wh-movement from the subject of the seem clause to the specifier of CP. And what is relevant here is the type of the Kase assigned to the position of $\underline{t}_i^1$, which is an immediate 'antecedent' of the relevant trace $\underline{t}_i^2$.

23. I confined myself to the discussion of "weak" crossover cases. The failure of the Binding Theory (C) type approach to the "strong" crossover cases is argued, convincingly I believe, by Higginbotham (1981) based on the examples like *which picture of which daughter of which man$_i$ does he$_i$ like, *which picture of which daughter of which friend of which man$_i$ does he$_i$ like, etc. Crossover cases involving LF movement such as *?his$_i$ mother loves everyone$_i$ (QR, cf. May 1977, 1985)), *?his$_i$ mother loves JOHN$_i$ (Focus movement, cf. Chomsky (1976)), etc. show somewhat different property than the one I described in the preceding discussion. It seems that in LF any element adjoined in that component (assuming the 'adjunction' analysis of May (1977, 1985)) can assume an 'operator-like' status, regardless of its content. Thus, not only quantifiers like everyone and wh elements like who, but also the otherwise non-operatorlike element John, can function as an operator in LF (cf. topicalization case discussed in the text).

24. The node SPEC is given in (45) just for the sake of exposition. This does not mean that I am claiming the existence of the node whose label is "SPEC." Rather, I am assuming (following Chomsky (1985)) that the notion 'specifier' is relationally defined (see 2.1).

25. Note that Case (or Kase) assignment takes place under "government." Thus, this conclusion is still not incompatible with Lasnik and Saito's (1984) argument against the "proper government" of the trace in COMP from the higher verb.

26. Note in this connection that there are many languages (e.g. French, Spanish) that do not allow ECP. See Chomsky (1981) and reference therein.

27. I am assuming that only "maximal projections" are visible to Case-marking (as well as to other grammatical processes like θ-marking). Thus, the D' in (51) must be a "maximal projection." The notion "maximal projection" will be made precise in Chapter 3.

28. Note incidentally that Stowell's (1981) Case Resistance Principle neither provides a principled account of the difference between IP (and CP) and DP nor is consistent with my system in this respect, since I am assuming DET to be a (potential) Kase-assigner.

29. This extension is suggested to me by Ken Hale (personal communication).

CHAPTER 3

CONSEQUENCES OF THE PROJECTION SYSTEM

This chapter takes up and discusses some consequences of the theory of projection proposed in Chapter 2. In the following discussion, we will assume the configurational notions which follow, as introduced in Chapter 1 and repeated here for convenience.

c-command:

α c-commands β iff α does not dominate β and every γ that dominates α dominates β

m-command:

α m-commands β iff α does not dominate β and every γ, γ a maximal projection, that dominates α dominates β

dominance:

α is dominated by β only if it is dominated by every segment of β

government:

α governs β iff α m-commands β and there is no γ, γ a barrier for β, such that γ excludes α

3.1 The Simplification of θ-marking

Among other results, the theory of projection proposed in Chapter 2 makes it possible to greatly simplify the definition of sisterhood, and consequently the mechanism of θ-marking (to the subject). This is because, in my system θ-marking takes place only within the projection of a Lexical head. More specifically, subject stands within a lexical projection at D-structure, the level at which θ-marking takes place, as opposed to, for example, in the specifier of IP position as assumed in the standard version of GB theory.

Let us first take up the definition of "sisterhood" given in Chomsky (1985).

(1) α and β are <u>sisters</u> if they are dominated by the same lexical projections.

(Chomsky (1985:10))

Compare (1) with the simple, most restricted (and also "traditional") definition of "sisterhood" given below, which is the equivalent of the mutual c-command relation between α and β (assuming that no dominance relation holds between α and β).

(2) α and β are <u>sisters</u> if they are dominated by the same nodes

The additional part in the definition of sisterhood (1) i.e., the stipulation that only the projection of a <u>Lexical</u> category is relevant to the definition of sisterhood, is necessary in Chomsky's (1985) system in order to make it possible to define the θ-marking of subject in terms of sisterhood, so that we get the following simple characterization of direct θ-marking.

(3) α <u>directly θ-marks</u> β only if α and β are sisters

(Chomsky (1985:11))

Consider the following schematic representations for sentences and noun phrases.

(4) a. $[_{IP}$ <u>NP</u> $[_{I'}$ I $[_{VP}$ V ... $]]]$

b. $[_{NP}$ <u>NP</u> $[_{N'}$ N ... $]]]$

The underlined NP's are subjects; it is the subject of a clause in (4a) and in (4b), it is the subject of a noun phrase. In (4a), the subject NP and the VP are sisters even though there is a node dominating VP but not the subject NP, namely I', since I' is not a projection of a Lexical category. Thus, VP directly θ-marks the subject NP. But the verb in (4a) does not directly θ-mark the subject NP because the VP, a lexical projection, dominates the verb but does not dominate the subject NP; the verb only indirectly θ-marks the subject, mediated through the VP, a desired result as argued in Chomsky (1981) and Marantz (1984). Similarly in (4b), the underlined subject NP and the head N are not sisters, since a lexical projection N' intervenes between them. Thus, the N in (4b) does not directly θ-mark its subject; it only indirectly θ-marks its subject mediated through N'. (But N' itself directly θ-marks the subject NP.) Notice that in the noun phrase case (4b), the specification of whether or not the

projection is a lexical projection is unnecessary; we can obtain the right result based on the simpler notion of sisterhood (2), i.e., the subject NP and the head N are not sisters due to the existence of N' which dominates the head N but not the subject NP, so that the head N only indirectly θ-marks its subject (presumably mediated through N'). The crucial case for the specification of whether or not the projection in a lexical projection in Chomsky's (1985) system is the clausal case (4a). If we define sisterhood as in (2) without reference to the "lexical projectionhood" of a given category, then the subject NP and the VP cannot be sisters since there is a node dominating the VP but not the subject NP, namely I'. Thus, in Chomsky's (1985) system it is crucial to take into consideration the lexical/nonlexical distinction of a given category, thus extending the notion of sisterhood so that the subject and VP can be sisters, in order to state the condition on direct θ-marking uniformly in terms of sisterhood.

This extension of the notion of sisterhood not only renders the notion less restrictive and less simple, but also creates a serious problem in regard to the θ-marking relation between a verb and its clausal complement. Consider the following representation where V takes CP as its clausal complement.

(5) V [$_{CP}$ [$_{C'}$ C [$_{IP}$ ...

Given the weakened definition of sisterhood (1), V and IP are sisters since the intervening nodes CP and C' are not lexical projections. Nevertheless, V should not (directly) θ-mark the IP, because it is the CP, not the IP, that is the complement of a verb. Chomsky notices this problem and states that we can in fact allow this possibility and can still get the right result: Suppose that V mistakenly θ-marks IP in (5), then a violation of the θ-criterion will result since the argument CP will not receive a θ-role (assuming that the verb has only one θ-role to assign). Notice that it is crucially assumed in this account that IP need not get a θ-role. Consider now the following example.

(6) I believe [$_{IP}$ John to be intelligent]

The <u>believe</u> type verbs are generally assumed to take IP as their clausal complements (cf. Chomsky (1985), Massam (1985)). In (6), <u>believe</u> and the subject of the complement clause <u>John</u> are sisters, given the definition (1), since the intervening IP is not a lexical projection. A θ-criterion account given above cannot handle this case, since we

just decided that IP need not get a θ-role. One might reconcile this difficulty by assuming that IP need not get a θ-role when it is not an argument but it has to get a θ-role when it is an argument. But this seems to me to be completely circular. Another possible way of resolving this difficulty is to appeal to the other half of the θ-criterion. That is, if John gets a θ-role from the verb believe, it will result in a θ-criterion violation since it also gets a θ-role from a predicate (be) intelligent, producing a θ-role conflict. Consider then, the following examples.[1]

(7) a. *I believe [$_{IP}$ John to rain]

b. *I believe [$_{IP}$ John to seem that Bill is crazy]

It is in the first place not entirely clear that we can claim that in (7a) the verb rain assigns a θ-role to the subject John. Therefore, no θ-criterion account similar to the one given above could easily be proposed in this case. Even if we somehow assume that rain assigns a "quasi θ-role" to its subject (cf. Chomsky (1981)), it is completely impossible to assume that the raising predicate seem assigns a some sort of θ-role to its subject in (7b), since the inability of the raising predicate (as well as

the passive predicate) to assign a θ-role to its subject is the crucial basis for any analysis of the raising construction. Thus, nothing prevents an "exceptional θ-marking" of the subject of IP by the matrix verb <u>believe</u>, and hence nothing accounts for the ungrammaticality of the examples in (7). One final possibility to get over this difficulty in Chomsky's (1985) system is to assume that <u>believe</u> takes CP, rather than IP, as its complement at D-structure, where θ-marking takes place, even if its clausal complement is infinitival.

(8) I believe [$_{CP}$ [$_{IP}$ John to be intelligent]]

Given (8), we can extend Chomsky's original account to say that if the verb <u>believe</u> mistakenly θ-marks the subject of IP, <u>John</u>, then there will be a θ-criterion violation, since the argument CP remains θ-less. In this account, the alleged D-structure (8) must be converted to the S-structure (9) (=(6)) to account for the exceptional Case-marking of <u>John</u> by the matrix verb <u>believe</u> (we are assuming that Case-marking takes place at S-structure).

(9) I believe [$_{IP}$ John to be intelligent]

That is, we have to posit a rule that "deletes" the CP node in the course of derivation from D-structure to S-structure, a rule similar to the one called "S'-deletion" (cf. Chomsky (1981)). However, the postulation of this rule creates at least two serious problems which would not arise under the analysis that believe type verb takes IP as its clausal complement (when it is infinitival) from the beginning, i.e., at D-structure. First, in the current framework of GB theory we are assuming here, the rule "S'-deletion" can no longer have any intuitive content which it used to have in, say, Chomsky (1981). For example, it can no longer be a mechanism of capturing the 'bar-reduction' phenomena, since what the rule is supposed to do in the current framework is to "delete" the node CP (presumably with COMP), a node totally independent from IP, given the extension of the X-bar schema to nonlexical categories proposed in Chomsky (1985), i.e., "S'" is not "S plus one bar" any more. Secondly, the claim that believe takes CP as its complement at D-structure, and the same verb takes IP at S-structure (due to the "S'-deletion"), which is crucially assumed in this account, is in direct conflict with the Projection Principle which states, informally, that the θ-marking properties of each lexical item must be represented categorially at each syntactic level. A somewhat formal statement of the

Projection Principle is adapted from Chomsky (1981).

(10) The Projection Principle

Given $[_{\gamma} \ldots \alpha \ldots \beta \ldots]$
$[_{\gamma} \ldots \alpha \ldots \beta \ldots]$

where α is an immediate constituent of γ:

(i) if β is an immediate constituent of γ at L_i, and $\gamma = \alpha'$, then α θ-marks β in γ

(ii) if α selects (where "selection" means θ-marking, direct or indirect) β in γ as a lexical property, then α selects β in γ at L_i

(iii) if α selects β in γ at L_i, then α selects β in γ at L_j

(cf. Chomsky (1981:36-38))

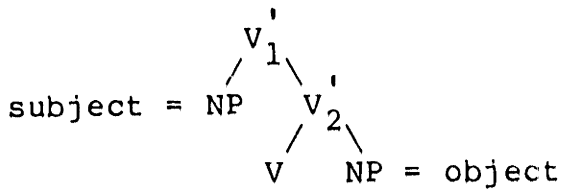
The important part of the Projection Principle for the present discussion is (10iii) which dictates that α's selectional property be preserved at every linguistic level (D-structure, S-structure, and LF). It is clear that the "S'-deletion" account would create a violation of this requirement, since, in such an account, believe selects CP as its complement at D-structure, but it selects IP, a different category, at S-structure, due to the "S'-deletion" rule. (Recall that under the current framework (cf. Chomsky (1985)) these two categories are not related in terms of

projection : they are independent categories.)

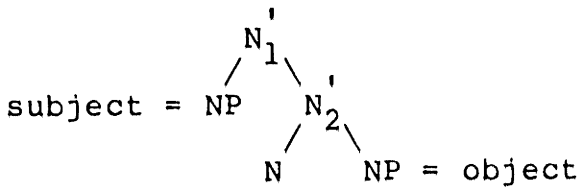
I have shown that the assumption that the subject of a clause appears in the specifier of IP at D-structure requires the extension of the notion of "sisterhood" as proposed in Chomsky (1985) in order to maintain θ-marking as a process taking place under sisterhood. I have also pointed out that this extension of sisterhood not only is conceptually undesirable but also creates a serious problem, namely, it wrongly allows "exceptional θ-marking" into a clausal complement of a verb such as <u>believe</u>. Various possibilities have been entertained in the previous discussion to resolve this difficulty in Chosmky's (1985) framework, but nothing turned out to be satisfactory. It seems now clear that there is no straightforward way of overcoming the difficulty under the standard assumption.

On the other hand, in the system I proposed in Chapter 2, there arises simply no problem comparable to the one pointed out above. Recall that in this system subject of clause appears at D-structure in one of the base-generated "adjoined" positions of a projection of V. In other words, the position of subject of a clause at D-structure is exactly parallel to that of noun phrase subject.

(11) a.



b.

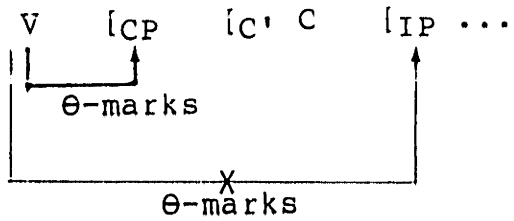


Given the configurations in (11), it is clear that the extension of the notion of sisterhood as stated in (1) is no longer necessary to handle the θ-marking of subject in clauses. We can maintain the more restricted definition of sisterhood given in (2). In (11a), as well as in (11b), the subject NP and the verb (the noun in (11b)) are not sisters since V'_2 dominates the verb but not the subject NP. Thus, the verb does not directly θ-mark its subject. But the V'_2, a direct projection of the verb, and the subject NP are sisters since they are dominated by the same nodes, i.e., there is no node that dominates one but not the other. Therefore, the V'_2 directly θ-marks the subject NP (and hence the verb indirectly θ-marks its subject mediated through its projection V'_2).[2] Notice incidentally that the direct θ-marking of the subject NP by V'_2 is still "compositional" in the relevant sense.

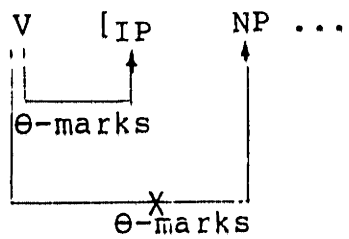
Thus, the way of θ-marking of subject I am proposing is consistent with the claim that the θ-role of subject should be determined compositionally by the VP (cf. Chomsky (1981), Marantz (1984)), replacing the node label "VP" by "V'" as required in my system.

Let us now consider the θ-marking relation between the <u>believe</u> type verbs and their clausal complement, which creates a problem of "exceptional θ-marking" in the system that involves the notion of sisterhood as defined in (1). As I have just discussed, in the system I am proposing there is no need to extend the notion of sisterhood in order to handle the θ-marking of subject: we can maintain the simple definition of sisterhood as in (2). With this in mind, consider the following configuration.

(12) a. (=(5))



b. (cf. (6))



In (12a), direct θ-marking of IP by the matrix verb is straightforwardly ruled out because they are not sisters given the definition of sisterhood (2). Thus, the verb can only directly θ-mark CP, which is its sister; they are dominated by the same nodes (assuming that dominance is irreflexive. See Chomsky (1985)). In (12b), a typical "exceptional θ-marking" construction, IP is a sister of V so that it gets a θ-role from the verb, i.e., V directly θ-markes IP. However, the NP in (12b) is not a sister of V due to the existence of IP that dominates the NP but not the V. Thus, the direct θ-marking of the NP by the matrix verb is impossible, as desired. In this system, the explanation for the contrast in grammaticality between (6) and the examples in (7) is straightforward. In (6), John, the subject of IP, gets a θ-role from the predicate (be) intelligent but it does not get a θ-role from the matrix verb believe, since believe and John are not sisters due to the intervening IP. Therefore, no violation of the θ-criterion ensues (Recall, incidentally, that unlike direct θ-marking, government, and hence Case-marking, do not require sisterhood. Thus, the "exceptional" Case-marking of John in (6) is possible). On the other hand, the subject of IP in the examples of (7) cannot get a θ-role at all, since rain in (7a) and seems in (7b) both lack a θ-role to assign, and the θ-marking

from the matrix verb <u>believe</u> is impossible because <u>believe</u> and the subject of IP are not sisters. Thus, <u>John</u>, the subject IP, remains without any θ-role assigned to it, resulting in a θ-criterion violation. Hence the ungrammaticality of (7a) and (7b).

I have argued in this section that the extended definition of sisterhood given in (1) is not only conceptually undesirable but also creates a serious empirical problem, i.e., it wrongly allows the possibility of "exceptional direct θ-marking." Under the system of projection proposed in Chapter 2, on the other hand, the θ-marking of subject is straightforwardly carried out under the restricted notion of sisterhood defined in (2), without creating such problems as "exceptional θ-marking." Thus far, I have not discussed the exact mechanism of θ-marking, however. Let us briefly see how θ-marking takes place in my system. The following discussion on θ-marking mechanism is rather sketchy and is not intended to be comprehensive.

I have been assuming that θ-marking takes place under the strict sisterhood in the sense defined in (2). Let us further assume that an argument structure, a "θ-grid" in the sense of Stowell (1981), is more than just an unordered list of θ-roles: it is structured according to the "closeness" of a θ-role to the predicate.[3] I will represent this by the linear order of the θ-role in a

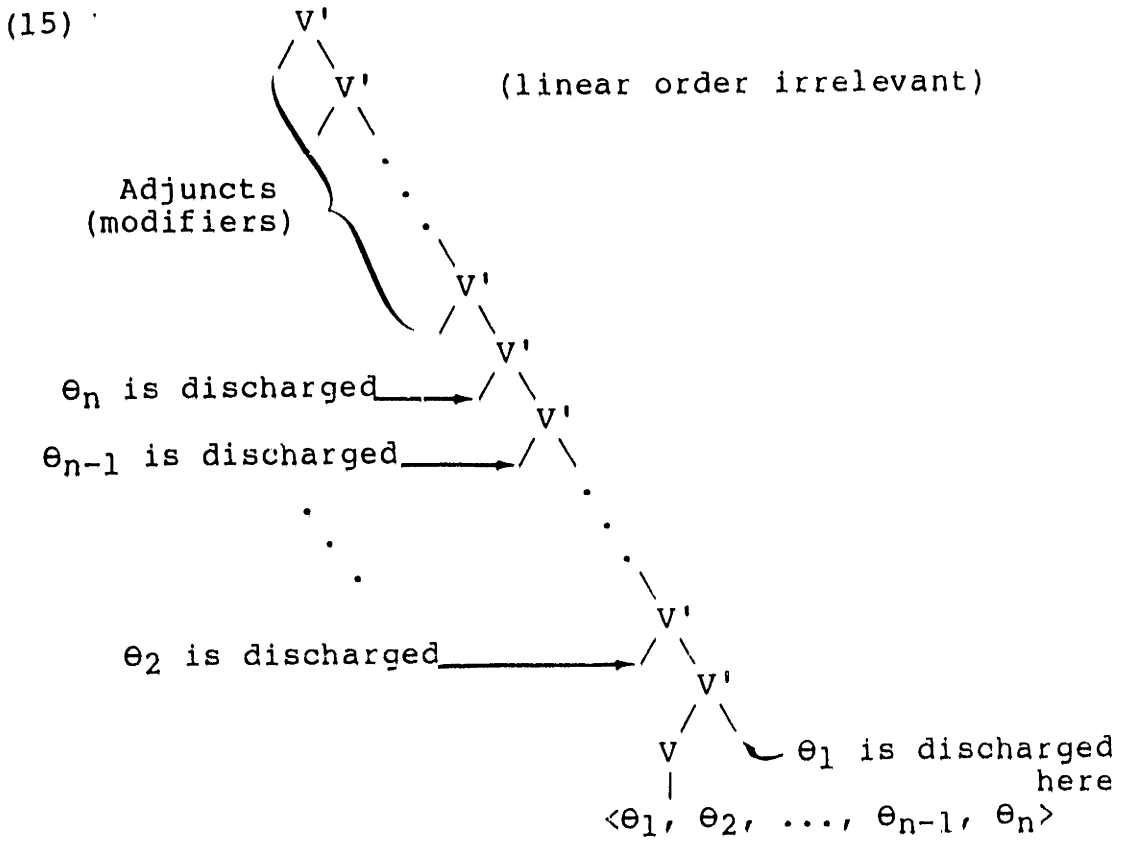
θ-grid, i.e., the lefthand θ-role is "closer" than the one to its right in a θ-grid of the lexical head. For example, θ_i is "closer" than θ_{i+1} to the lexical head to which the θ-grid is associated, and θ_1 is the "closest" θ-role.

(13) θ-grid = $< \theta_1, \ldots, \theta_i, \theta_{i+1}, \ldots, \theta_n >$

And I will also assume that the "discharge" of the θ-roles (in the sense of Higginbotham (1985)) takes place sequentially from left to right under the strict sisterhood without skipping over a non-θ-marked position. This mode of θ-marking, coupled with Higginbotham's (1985) version of the θ-criterion introduced in Chapter 1, which is repeated here as (14) here, gives us the following schematic D-structure representation (15) for, say, a verbal projection.

(14) a. Every thematic position is discharged
b. If X discharges a thematic role in Y, then it discharges only one.
(Higginbotham (1985:561))

(15)

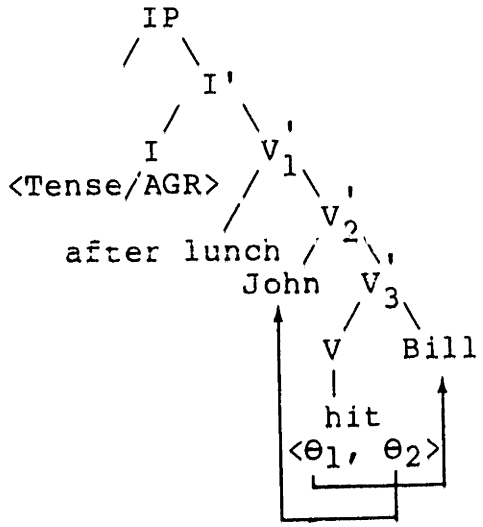


In (15), the only position which gets θ-marked directly by the V is the sister of the verbal head V. In other words, θ_1 is the only "internal" θ-role, and all the other θ-roles in a given θ-grid are assigned compositionally from the bottom up under the sisterhood relation. After every thematic position in a θ-grid nas been discharged, non-arguments (adjuncts/modifiers) may appear, and these modifiers will be placed in appropriate positions for their interpretation later in the derivaton, presumably in LF (see Czaykowska-Higgins (1986) for details of adverbial

modifications.)[4] Thus, for example, the D-structure representation for (16) under the present assumptions will be (17), where Bill is directly θ-marked by the verb hit, John is compositionally θ-marked by the verb mediated through V_3', and after lunch is a modificational phrase. (Recall again that linear order in (17) is irrelevant.)

(16) John hit Bill after lunch

(17)



In the course of derivation, as I mentioned above, John will be moved into the position of specifier of IP to receive Case, and after lunch will be placed in an appropriate position for its interpretation.

The above discussion is not at all conclusive. We will obviously have to specify more closely the mechanism of

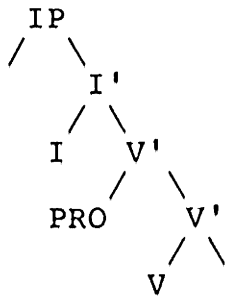
θ-marking, e.g., obligatoriness/optionality of θ-marking, etc. For more details, see, among others, Chomsky (1985, 1986), Higginbotham (1985), Williams (1985a), and Czaykowska-Higgins (1986). I will briefly return to the related matters in Japanese in Chapter 4.

3.2 The Position of PRO

Another consequence of the system of projection proposed in the previous chapter is the position of PRO. As I discussed earlier, the subject of a sentence starts out within a projection of V and then is moved into the specifier of IP to get Case in my system. I also claimed that the D-structure position of the subject does, which is within a projection V, is the position to which an external θ-role is assigned. This implies that PRO can also appear in the same position as the lexical subject, the only difference being that the lexical subject must move to the position of the IP specifier to receive Case, whereas PRO stays in the original position because it need not be, in fact it cannot be, Case-marked. Thus, the schematic clausal structure in which PRO can appear is

such as follows.[5]

(18)



If I in (18) does not contain Tense/AGR, i.e., if it is _no_, and no Kase is discharged from outside the IP, nothing happens and PRO will be controlled if there is a possible controller; if there is no such a controller, it gets a so-called "arbitrary" interpretation as in _it is unclear what to PRO do_. If I in (18) contains an F-Feature to be discharged in the specifier of IP position, then something must be moved to that position to avoid a violation of the Saturation Principle. In most cases, movement of a phrase under these circumstances would result in a violation of some principles of grammar, for example, Case conflict. Therefore, PRO may normally appear only in an infinitival clause where I does not contain an F-Feature. However, there is a case where movement of a phrase into the position of the IP specifier would be allowed, namely a

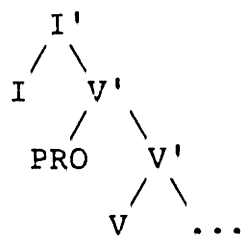
case of passive.

Before going into the discussion of passive, let us make sure that the appearance of PRO within a projection of V does not violate any principle of UG. The crucial requirement imposed on the distribution of PRO that has been widely assumed in the literature is the so-called "PRO theorem" which dictates that PRO be ungoverned. This theorem can be derived from the "classical" Binding Theory (cf. Chomsky (1981, 1982)) by the following reasoning. Observe first that PRO is like an overt pronoun in that it never has an antecedent within its clause (or NP), and that it is also similar to anaphors in that it has no intrinsic referential content. Suppose then that PRO is a pronominal anaphor whose feature specification is [+anaphor, +pronominal] (cf. Chomsky (1982)). If PRO has these features, it has to obey both the Binding Theory (A), which requires [+anaphor] element be bound in its governing category, and the Binding Theory (B) which stipulates that [+pronominal] element be free in its governing category, a contradiction. Therefore, PRO has no governing category and is therefore ungoverned. The same reasoning can naturally hold with some modifications even within the most recent version of the Binding Theory (Chomsky (1986)) that we are assuming in this study: It gives a licensing condition for a category governed by a lexical element.

From this, it follows that a pronominal anaphor (PRO) meets contradictory conditions, the clauses (A) and (B) of the Binding Theory. It then follows that a pronominal anaphor (PRO) is ungoverned (or "not lexically governed") if license.

Then our task is to show that in the following structure PRO is indeed ungoverned.

(19)



I have already suggested before that I does not govern the position PRO now occupies and that the inability of government across a maximal projection is a property of Functional categories in general, presumably because of the directionality of F-Feature assignment. No problem arises as to the government by the lexical head V, either, as long as V' is the maximal projection of V proposed above. I have suggested (cf. footnote 2) that May's (1985) distinction between categories and segments does not hold for the base-generated "adjunction" structure such as (19).[6] If this is the case, then PRO is not

m-commanded by, hence not governed by V, since a maximal projection, namely the lower V', dominates the V but not PRO in (19).[7] Thus, if (19) is not considered to be an "adjunction" structure relevant to May's distinction, no problem arises as to the government of PRO by V, since it is clear that PRO is not m-commanded by V (Recall that "m-command" is the necessary, though not sufficient, condition for government.) This move, namely the move toward the distinction between "real adjunction structure" and base-generated structure that "looks similar" to adjunction structure, seems to me to be the right one. See Chapter 4 for further discussion.

Furthermore, even if May's distinction turned out to be applicable to the base-generated structures like (19), PRO would still not be m-commanded, and hence not be governed, by the lexical head V, because the lower V', now interpreted as a segment of a category V', does not dominate PRO. Thus, the category V' which consists of two segments, the higher V' and the lower one, does not dominate PRO, but it does dominate V, since every segment of the V' dominates V. Therefore, there is a maximal projection that dominates V but not PRO, namely the category V'. Hence, the V does not m-command PRO even if May's distinction holds in (19).[8] Incidentally, it goes without saying that PRO in (19) <u>does</u> govern and hence

m-command the elements dominated by the lower segment of the category V', including V, as generally required for the antecedent government of a trace by its "VP-adjoined" antecedent (cf. Chomsky (1985)). This does not entail, of course, that the "reverse" government relation, i.e., government of an "adjoined" element by something under (dominated by) V' must hold. In fact, I have just argued that the "reverse" government relationship cannot hold.

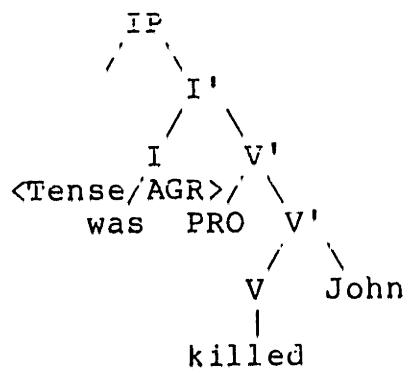
I have thus shown that the postulation of PRO does not violate the requirement (the "PRO theorem") that PRO be ungoverned, under either interpretation of the "adjunction" structure (19).

The postulation of PRO within the projection of a Lexical category has various consequences. Among these is the fact that we now may make overt a representation of these "implicit arguments" (cf. Roeper (1983, 1984)) which behave as though they were syntactically present.[9] Consider in particular the passive. Under previous analyses, if the passive morphology absorbed accusative Case, it was necessary to assume that it also absorbed the subject θ-role, so that the subject position could be an available landing site for NP movement. Under my proposal, this is not necessary: the passive morphology absorbs accusative Case, but it need not absorb the subject θ-role.[10] Thus, there must be a position to which the

subject θ-role is assigned (when it is not absorbed), or a violation of the θ-criterion (a part of the Saturation Principle) results. PRO appears in that position, receiving the external (subject) θ-role from the verb mediated by V'. This PRO is what has been called the "implicit argument." To illustrate, consider the passive (20), whose D-structure representation in our system is (21) (ignoring the V-raising operation, which is, strictly speaking, to take place later in the derivation).

(20) John was killed.

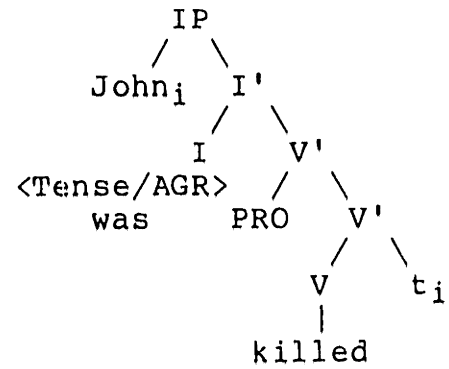
(21)



In the course of derivation PRO remains in its D-structure position, whereas the object John, which cannot be assigned Case in its D-structure position due to passive morphology, must move to the IP specifier position to receive Case

(Kase) from I, yielding the following S-structure.

(22)



Thus, the status of implicit arguments are now no longer a mystery; they are overtly represented in a phrase structural configuration within the projection of a Lexical category.

One interesting result of this approach to implicit arguments is that we now make explicit what we should do for the explanation of the well-known observation that the implicit argument (a "hidden agent") in a passive is obligatorily disjoint from its S-structure subject: our task now is to determine how to block the coindexing of PRO with John in a structure like (22). If the chain formed by the movement of John, namely (John$_i$, t_i), is an A-chain, as is widely assumed, then the obvious way to prohibit PRO from being coindexed with John (and with t) is to resort to the following property of A-chains, which Chomsky (1981) considers as one of four defining characte-

ristics of A-chains (see Chomsky (1981:333)).

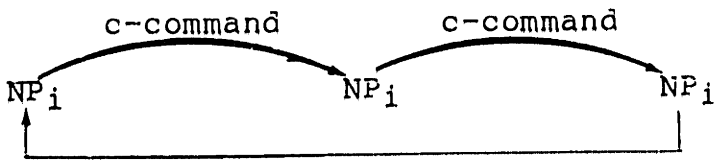
(23) α_i locally A-binds α_{i+1} (where α_i are members of the chain)[11]

Coindexing PRO with John (and with t) would clearly make the chain ($John_i$, t_i) an impossible A-chain, since under this coindexation John would no longer locally A-bind t_i. And if the chain ($John_i$, t_i) is ruled out as a violation of (23), then the structure is also excluded as a violation of the θ-criterion since John does not receive a θ-role. Notice that this explanation is valid only if the chain ($John_i$, t_i) is an A-chain. However, as we saw before, the A/A' distinction is not at all straightforward in the system I am proposing. It is probably possible even in our system to make a necessary distinction between chains without recourse to the A/A' distinction, presumably in terms of whether or not Case is assigned to the head of a given chain as suggested in Chapter 2 (and with reference to the "content" of the head of a chain, which is independently necessary). If this is the case, then we can distinguish different types of chain appropriately, and can make use of the condition (23) to account for the impossbility of coindexing PRO with John (and with t),

modifying the condition accordingly so as to refer to a particular type of chains. However, even a more straightforward way of handling this problem seems to be possible. Namely, to generalize the condition (23) to all chains. In fact, such an extension is proposed by Lasnik (1985) (cf. also Rizzi (1982)) as a generalized strong crossover constraint.

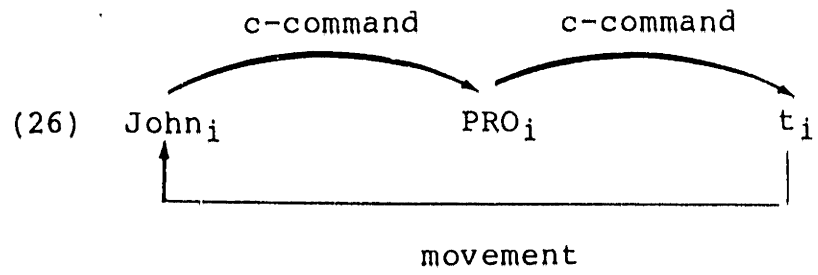
(24) If α_i and α_{i+1} are successive links in an A/A' chain, then α_i locally A/A' binds α_{i+1}.
(Lasnik (1985:488))

This generalized locality condition on chains would rule out all derivations with the following abstract property.

(25)


(Lasnik (1985:488))

It is clear that coindexing PRO with John (and consequently with t) in (22) would create the structure with exactly this abstract property.



Thus, to the extent that Lasnik's generalization of the locality condition on A-chains (23) to all chains is supported,[12] the impossibility of the coindexation of PRO with John in (22), i.e., the obligatory disjointness between PRO and John, is basically accounted for.

The preceding discussion is not at all conclusive as to what is a principled explanation for the fact that the S-structure subject and the "hidden agent" in a passive are always disjoint to each other. However, one thing is clear. Given ;the system of projection we are proposing, we can now relate this observed fact to the general caseS-structure subject and the "hidden agent" in a passive are always disjoint to each other. However, one thing is clear. Given the system of projection we are proposing, we can now relate this observed fact to the general case of "crossover" facts, whatever the real explanation for the "crossover" facts turns out to be.

There are several problems that should be discussed in connection with the proposed structure (22) for

passives. First of all, one might object that the structure (22) looks very similar to the "obligatory control" structure, in which PRO is obligatorily controlled by its controller. Even if the impossibility of coindexing the S-structure subject with PRO is accounted for along the lines just suggested, the structure must be ruled out since PRO in (22) must be obligatorily controlled and it cannot be controlled by John in (22). However, the superficial similarity between the "obligatory control" structure and the structure (22) is not an apparent one. Recall that I have suggested that the Binding Theory should be formulated as a theory of "θ-binding" in the projection system I am proposing. If this is the correct characterization of the Binding theory in my projection system, it can naturally be extended to "control" relation, in view of the fundamental similarity between the binding relation and the control relation (cf. Manzini (1983)). Then, we can characterize the control relation as follows.

(27) Control relation holds between two θ-positions.

In fact, all the typical "obligatory control" relations hold between two θ-positions. Consider, for example, the following.

(28) a. $John_i$ t_i tried [to PRO_i pass the exam]

θ θ

control

b. $John_i$ t_i promised Mary [to PRO_i came early]

θ θ

control

In (28), $\underline{t}_i$ is the trace of $\underline{John}_i$ left by the "subject raising" rule discussed in Chapter 2. And the control relation holds between the trace, a terminal element of a chain ($\underline{John}_i$, $\underline{t}_i$), and PRO, both being in θ-positions, as required by (27). Now consider (29), which is a reproduction of the relevant portions of (22).

(29) $John_i$ was PRO_i killed t_i

θ' θ θ

In (29), $\underline{John}_i$ is in a θ'-position as discussed before, so that control relation cannot hold between the position of $\underline{John}_i$ and that of PRO since (27) requires all of the positions that enter into control relation be θ-positions. Control of PRO by the trace $\underline{t}_i$, which is a terminal element

of the chain(John$_i$, t$_i$) left behind by the "passive movement" is also impossible because of the anti-c-command condition on control such as follows.

(30) A controllee cannot c-command its controller.

Although conditions like (27) and (30) are stipulative and should be derived from more general principles of grammar, they seem to be at least descriptively adequate. And if the obligatory control structure is characterized in terms of the conditions like (27) and (30), which seems reasonable, the structure (22) does not fall under the general obligatory control cases. Therefore, the impossibility of the coindexation between PRO and the S-structure subject John will not make the structure ill-formed.

Next problem has to do with the position of PRO in (22). Notice first that the position that PRO occupies in (22) is also the position in which a lexical NP appears at D-structure, and I proposed in Chapter 2 that this lexical NP is moved into the specifier of IP position to get Kase, leaving a trace in its D-structure position. It has occasionally been proposed (see, in particular, Jaeggli (1982), Rizzi (class lectures, MIT, 1984), Wahl (1985), and Chomsky (1986)) that trace must be governed as well as

antecedent-governed. If this is a correct condition on traces, then the position that PRO occupies in (22) must be governed since it is the position where the trace is left behind by the "subject raising" rule proposed in Chapter 2. On the other hand, the "PRO theorem" requires that PRO be in an ungoverned position. This is a contradiction since the same position is required to be governed and ungoverned, from considerations of different principles of grammar. There are various ways of overcoming this problem. Let us briefly consider some of those possibilities.

Note first that the above-mentioned paradox crops up only if the two conditions, the condition on traces "trace is governed," and the PRO theorem "PRO is ungoverned," are true in the forms exactly as they are stated in these quotes. However, it is not entirely clear whether this is the case. Consider the "PRO theorem" first. As I have argued before, government of PRO by V in the structure in question is impossible, since V does not m-command the position of PRO (even if we assume May's (1985) distinction between category and segments). Therefore, the only possibility of government of the position occupied by PRO is government by I from outside of the projection of V. Let's suppose the government of the position of PRO by I is indeed possible. That is, even if I (when it contains

Tense/AGR) assigns Nominative Case only to its specifier position and not to a position inside the projection of V (as argued before), I nevertheless governs into the projection of V, in particular, it governs the position occupied by PRO.[13] Now PRO is governed by I in the position we are considering. Suppose further that the "PRO theorem" is stated in terms of "lexical government," i.e., government by a Lexical category, rather than "government," which seems reasonable in view of the fact that the notion "governing category" is defined (in part) by government by a Lexical category (cf. Chomsky (1986:171)).

(31) PRO is not lexically governed.

Assuming the core ideas of "V-raising" analysis in Chomsky (1985), the condition (31) seems to correctly determine the distributional property of PRO, i.e., essentially it occurs only in the subject position of infinitives and gerunds, while allowing the position we are considering to be governed, as desired.

Consider next the condition that trace must be governed as well as antecedent-governed. Evidence supporting the claim might be obtained from the contrast between

(32) and (33).

(32) *how raw did John eat the meat t.

(33) a. how clean did John pick the bone t.
b. how red did you paint the house t.
c. how angry did John make his friends t.
(Chomsky (1985:71))

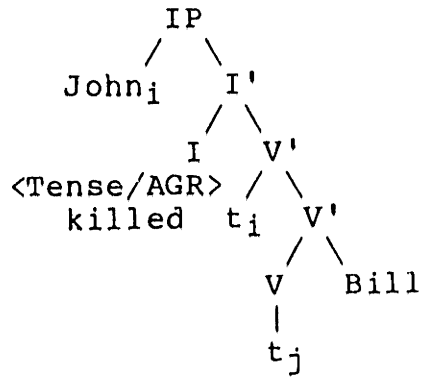
Extraction of how raw from the position designated by t in (32) is impossible, whereas extraction from similar positions is possible in (33). Assuming that the positions designated by t in (32) and (33) are all antecedent-governed (by, for example, the "VP-adjoined" trace with the notion of government defined in terms of exclusion in Chomsky's (1985) system. See Chomsky (1985) for details and background assumptions.), the only conceivable difference between cases like (32) and those like (33) would be that there is a relation between the matrix verb and the wh phrase in (33), while there is no such relation in (32). As Chomsky (1985:71) puts it: "in (iii)(=(33c);N.F.) under a small clause analysis with the fronted AP as the selected head, in (i)(=(33a);N.F.) because of the lexical character of "pick clean" (as distinct from "eat raw"); and in (ii)(=(33b);N.F.), possibly for the same reason

(note that picking-clean is something that can be done to a bone, and painting-red something that can be done to a house, but eating-raw is not something that can be done to meat)." Stating the condition on trace in terms of government is obviously one possible way of capturing this difference between (32) and (33) (and other relevant constructions). But it is by no means the only one. Suppose tentatively that the relevant condition on traces might be something like the following.

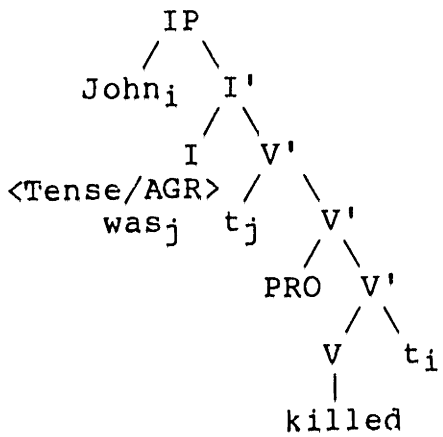
(34) Trace is "licensed" by a Lexical head.

"Licensing" in (34) include θ-marking, predication, and modification (presumably, an instance of predication). Notice that in our system (and in part in the standard system as well) these relations do not necessarily require government relation to hold between the two elements involved. What our system does require is that the two elements involved in such relations be within the same Lexical projection. Now compare the following S-structure representations.

(35) a.



b.



In (35a), the trace of the "subject raising," namely $\underline{t_i}$, is not governed,[14] as argued before. However, it is "licensed" by V, i.e., the position of the trace is θ-marked by a Lexical category V (mediated by V'). Thus, the trace of $\underline{John_i}$ satisfies the condition (34) even if it is not governed. As for (35b), assuming the "helping verb" <u>be</u> in passive construction is not a Lexical element, PRO is not lexically governed in its S-structure position,

satisfying the requirement imposed by the "PRO theorem."

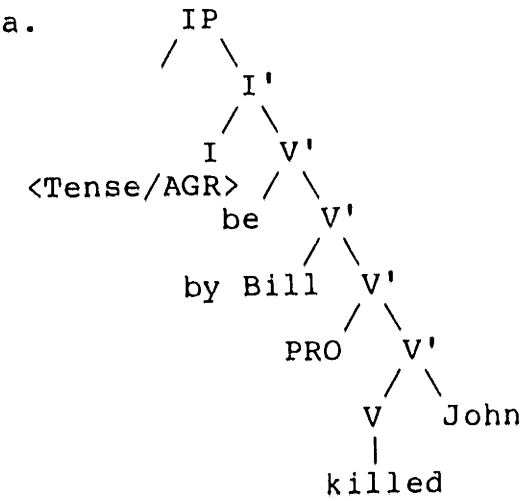
Although the previous discussion is rather sketchy it is now clear that the apparent paradox created by the "PRO theorem" and the condition on traces is not a serious one for our analysis of passives.

Let us now turn to the discussion of passives with an overt agent, namely by phrase. There are two approaches to by phrase in passives; one is to assume that by phrase is an adjunct of the passive predicate (cf. Zubizarreta (1985)), the other is to assume that it is an argument of the passive predicate (cf. Roberts (1985) and Jaeggli (1986)). Let's see how our system works under each of these assumptions.

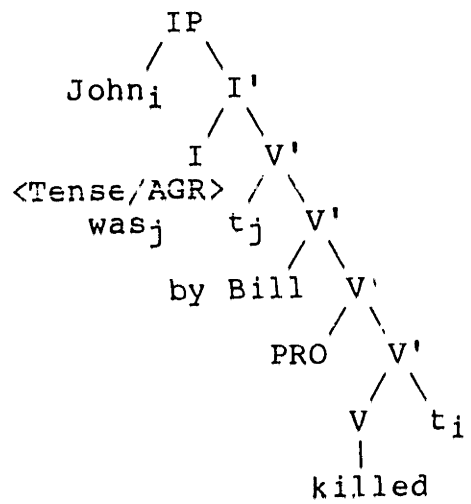
If the by phrase is an adjunct, then our system predicts that a sentence like (36) would have the D-structure representation (37a) and the S-structure representation (37b), ignoring the linear order.

(36) John was killed by Bill.

(37) a.



b.



First of all, in the D-structure representation (37a), the existence of PRO as an external argument of killed is required, because otherwise a violation of the Saturation Principle (the θ-criterion, in particular) would result. For we are assuming, for the sake of argument, that the

external θ-role of the verb killed is not absorbed. It is also required by the θ-marking mechanism proposed before that an adjunct by phrase occupy a "higher" position than every argument, in particular, PRO. Representation (37b) is the S-structure representation of (36) after "V-raising" of be and the movement of John (which is "forced" by the Case Filter since the passive morphology absorbs the Case-assigning power of a verb) have applied. The obligatory disjointness between PRO and the S-structure subject John can be accounted for in the way suggested before. Thus, the only problem that remains is to account for the obligatory control of PRO by Bill in (37b). Although nothing explicit can be said at this point due to the lack of substantial theory of control, it is clear that there is no seriously problematic factor for the control of PRO by Bill in (37b). If by does not create any branching structure relevant to c-command, then Bill is the "optimal" position for controlling PRO, i.e., Bill is the "closest nominal element c-commanding PRO" (cf. Rosenbaum (1969), Huang (1984), etc.). Even if by creates a branching structure visible for c-command relationship, it is still possible for Bill to control PRO, since control relation does not necessarily require a controller to c-command its controllee as exemplified by examples like the following (Examples are shown in (38) with the

standard structures, not with the one I am proposing).[15]

(38) a. [PRO_i to clear myself$_i$ of the charges] is important to me$_i$

b. [PRO_i finishing my$_i$ work on time] is important to me$_i$

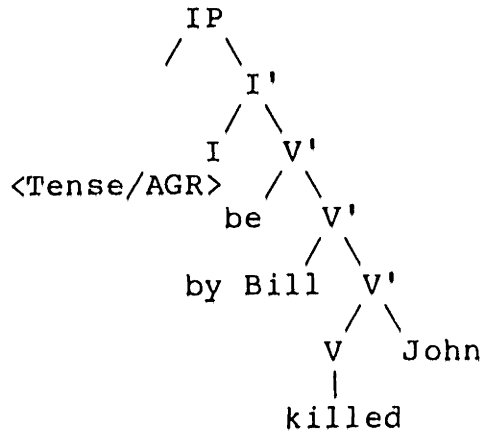
(Chomsky (1981:77))

c. [PRO_i losing the race] will upset John$_i$

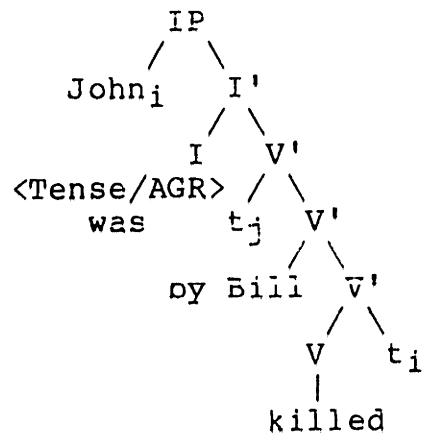
Therefore, there will be no serious problem with the structure (37b) with respect to the control relation between an overtly represented agent Bill and the postulated "hidden agent" PRO.

If, on the other hand, by phrase is an argument of the passive verb, receiving an external θ-role from the verb, then PRO cannot show up in D-structure, since if it did, no θ-role would be assigned to it, violating the θ-criterion. Thus, the D-structure representation and the S-structure representation for (36) should be (39a) and (39b), respectively.

(39) a.



b.



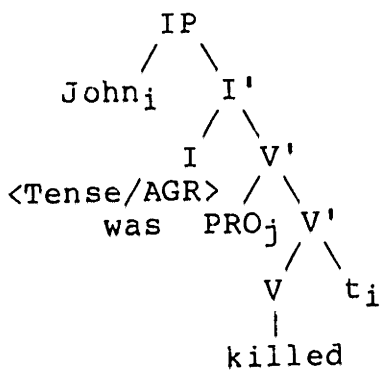
The problem of control does not arise under the assumption that by phrase is an argument, simply because PRO does not show up in the structure.

We have thus seen that there is no serious problem for our proposal about passives with respect to the overt "agentive" phrase. Let us go on to discuss another potential problem for my analysis, namely, the problem

related to the Specified Subject Condition (SSC) (cf. Chomsky (1973, 1981, etc.)).

Given the proposed S-structure (22) for passives, one might suspect that this structure would violate the SSC, or the Binding Theory (A), if PRO is not coindexed with the trace of John as in the following (we ignore here the trace of the V-raising as irrelevant).

(40)



If, as is widely assumed, the trace left by NP movement (movement to an A-position) should be regarded as an anaphor, then it must obey the Binding Theory (A). In (40), the least Complete Functional Complex containing the anaphor t_i and its lexical governor killed is (the higher) V'. Thus, the Binding Theory (A) requires, roughly, that the trace t_i be bound in V'. But the trace is not bound if PRO is not coindexed with it. Hence, the structure would violate the Binding Theory (A) under the standard

assumptions. Recall, however, that I have already discussed that the A/A' distinction is not straightforwardly defined and does not seem to have real content in the system I am proposing. As a result, I have suggested that the Binding Theory, and Control Theory for that matter, should be reinterpreted as a theory of θ-binding (with same qualifications on Case), i.e., a theory about the relationship between two θ-positions. I have also pointed out that all traces are θ'-bound in my system. From these considerations, I tentatively concluded that the Binding Theory does not apply to traces. Let us be clear about the scope of the Binding Theory and the status of traces with respect to the Binding Theory.

(41) The Binding Theory applies only to distinct chains.

Suppose now we have the following supplementary convention.

(42) The status of a chain with respect to the Binding Theory is determined by its head.

If these statements are true, then "chain internal" relations such as the one between $\underline{\text{John}}_i$ and its trace, $\underline{t}_i$,

in (40) are not subject to the Binding Theory. Specifically, the traces like $\underline{t}_i$ are no longer "anaphors": traces are just traces, whose distribution is constrained only by conditions on chains (and possibly, a condition like (34)). This division of labor seems to be conceptually desirable, eliminating the redundancies observed between the Binding Theory and other modules of grammar such as ECP/subjacency. If this is indeed the case, then the problem of SSC concerning (40) simply does not arise; the Binding Theory is inapplicable to the chain internal relation between $\underline{\text{John}}_i$ and $\underline{t}_i$ in (40). Or, if (42) is correct, coindexing PRO with the trace $\underline{t}_i$ is ruled as a violation of the Binding Theory (C). This is because the status of the chain ($\underline{\text{John}}_i$, $\underline{t}_i$) is an R-expression since its head is an R-expression, $\underline{\text{John}}_i$, and the terminal element of the chain, $\underline{t}_i$, is bound by PRO_i, assuming that the Binding Theory governs the relationship between terminal elements of distinct chains (with some modification concerning the position of the head of a chain, i.e., if the position of the head is a Case-marked position, then the head is also "visible" to the Binding Theory. This complication is independently necessary for the account of weak crossover phenomenon. See 2.6), which directly follows from (41) and our assumption that the Binding Theory is a theory of θ-binding, combined with the general

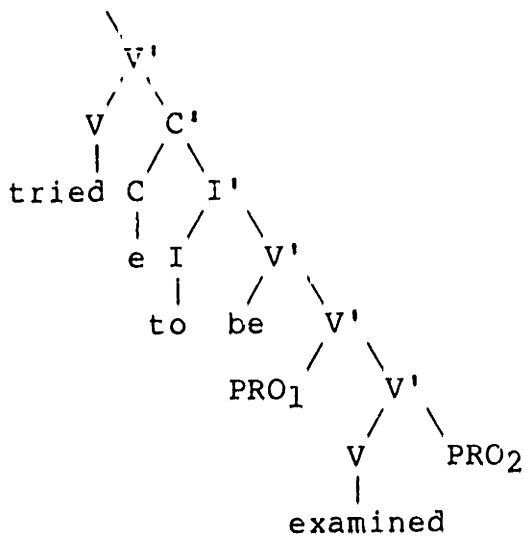
condition that all movements are to θ'-positions.[16]

Let us now discuss another apparently problematic case for the analysis of passive I am proposing. Consider the following example.

(43) John tried to be examined.

Under the analysis of passive proposed above, (43) should have the D-structure such as follows (irrelevant portions omitted.)[17]:

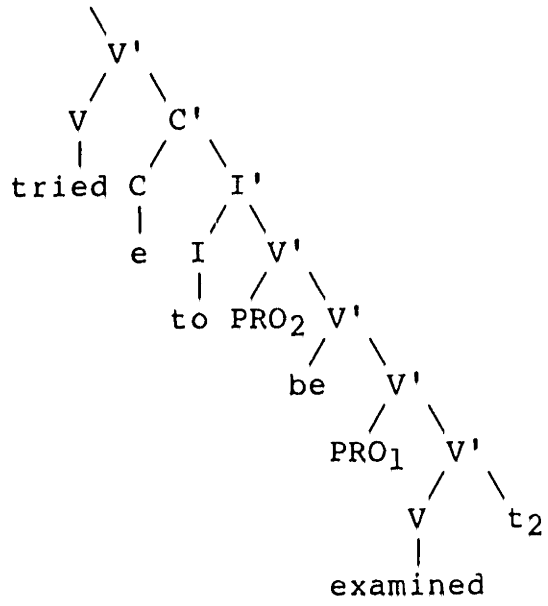
(44)



In (44), there are two PROs, PRO_1 and PRO_2. PRO_1 is a postulated "hidden agent" which we have been discussing.

PRO_2 is the one to be obligatorily controlled by the matrix subject John in (43). However, this PRO_2 is in the position governed (and furthermore "lexically governed") by the verb examined. Therefore, the "PRO theorem" (or the condition (31)) "forces" the PRO_2 to move into some other position which is not (lexically) governed. In the standard GB theory, this D-structure object PRO is assumed to be moved into the specifier of IP position within an embedded clause. But the theory of projection I am assuming here does not allow such a movement, since no Kase is assigned to the specifier position IP (to does not have F-Features and the "exceptional" Case assignment by the higher verb tried is blocked presumably by the minimality condition, with the empty complementizer being the minimal governor (but not Kase-assigner)), and hence that specifier position is not licensed. Therefore, we have to look for another option to avoid a violation of the "PRO theorem." Here, I will tentatively propose that PRO_2 is "adjoined" to V', yielding the following S-structure.[18]

(45)



Now in (45), PRO_2 is in the "ungoverned" position, as required by the "PRO theorem." Note that this movement of PRO_2 is a "licit" movement in that it does not violate any principles of grammar, in particular ECP/subjacency. Thus, although I have no strong argument for this particular movement, it is now clear that there is a way of getting the right result in our system with respect to the cases like (43).

Before closing the discussion of passives, let us briefly see that our analysis of passive, in particular, the postulation of PRO within a Lexical projection, V', can naturally be extended to noun phrases. Consider first the following paradigm taken from Chomsky (1986).

(46) a. they$_i$ told [stories about each other$_i$]
b. *they$_i$ heard [my stories about each other$_i$]
c. *they$_i$ told [stories about them$_i$]
d. they$_i$ heard [my stories about them$_i$]

(Chomsky (1986:166))

These are typical cases where we observe the SSC effect. The subject my blocks the binding of each other by they in (46b), whereas it allows the occurence of them coindexed with they, due to the Binding Theory (A) and (B), respectively. On the other hand, the SSC does not apply to (46a) and (46c) because there is no subject in a noun phrase. Thus, binding of each other is permitted in (46a) by the Binding Theory (A), and binding of them is blocked by the Binding Theory (B) in (46c).

Let us now consider the following paradigm, in which tell and hear in (46) are interchanged.

(47) a. they$_i$ heard [stories about each other$_i$]
b. *they$_i$ told [my stories about each other$_i$]
c. they$_i$ heard [stories about them$_i$]
d. they$_i$ told [my stories about them$_i$]

(Chomsky (1986:167))

Examples in (47) are well in accord with what the SSC

predicts except for (47c). There is no subject in a noun phrase in (47c), so binding of them by they should be blocked by the Binding Theory (B), but it is not. Chomsky (1986) suggests the explanation for this apparent exception to the SSC as follows. He notes that in (46c), we assume that the stories are theirs, whereas in (47c), stories are assumed to be someone else's. Based on this observation, Chomsky (1986) postulates the following structures for (46c) and (47c).

(48) a. *they$_i$ told [PRO$_i$ stories about them$_i$]
b. they$_i$ heard [PRO$_j$ stories about them$_i$]
(Chomsky (1986:167))

Given the structures in (48), which includes an implicit argument with the properties of PRO (cf. Chomsky (1986)), the SSC works properly. In (48a), binding of them by they (via PRO) is barred by the Binding Theory (B), whereas (48b) does not violate the Binding Theory (B), since PRO is not coindexed with (and hence does not bind) them in a noun phrase.[19] Chomsky (1986) suggests that the determiner position can include an implicit argument (hence identified as a PRO-like element). However, judgments of the relevant cases remain basically the same

even if there is an overt determiner the (putting aside the "specificity" effect invoked by the definite article), as shown by the following paradigm (cf. (47)).

(49) a. they$_i$ heard [the stories about each other$_i$]
b. *they$_i$ told [my stories about each other$_i$]
c. they$_i$ heard [the stories about them$_i$]
d. they$_i$ told [my stories about them$_i$]

This seems to indicate that the existence of PRO in a noun phrase correlates not with the existence of an overt determiner, but with the existence of an overt (external) argument of a noun (whose existence, in turn, depends on the noun's thematic structure as well as on the matrix verb's lexical property). This point is further clarified by the following examples.

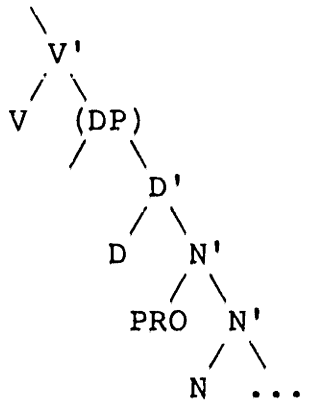
(50) a. they$_i$ heard [yesterday's stories about each other$_i$]
b. *they$_i$ told [my stories about each other$_i$]
c. they$_i$ heard [yesterday's stories about them$_i$]
d. they$_i$ told [my stories about them$_i$]

In (50c), the alleged determiner (or "specifier") position is occupied by an adjunct yesterday, but still PRO must be

present within a noun phrase to make the SSC work properly. Thus, it seems that the determiner (or "specifier") position itself has nothing to do with the (optional) occurrence of PRO within a noun phrase, contrary to Chomsky's (1986) suggestion.

In the system proposed in this study, the facts I have been discussing can be given a natural and straightforward explanation. Recall an "external" argument of a lexical head appears within a projection of that Lexical head in our system. Thus, given our system of category projection, Chomsky's (1986) "PRO-like implicit argument" can show up within a projection of a noun, exactly parallel to an external argument of a verbal head in a clausal case.

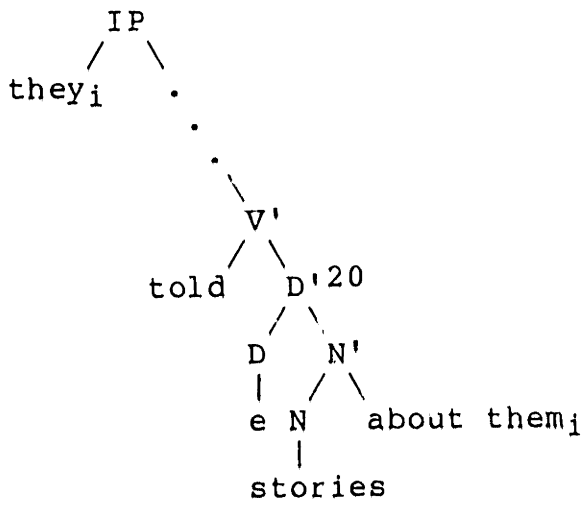
(51)



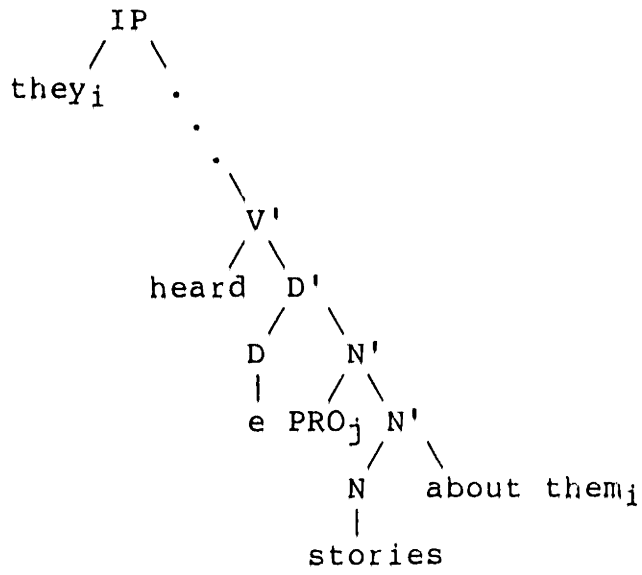
The possibility of the occurrence of PRO in (51) is

determined by the lexical property of the nominal head (and perhaps by the matrix verb's lexical property as well), and has nothing directly to do with the position of the determiner D, which is, in our system, a head of DP (see Brame (1981, 1982)). Thus, the S-structure representations for the relevant cases under discussion should look like the following in our system (irrelevant details omitted).

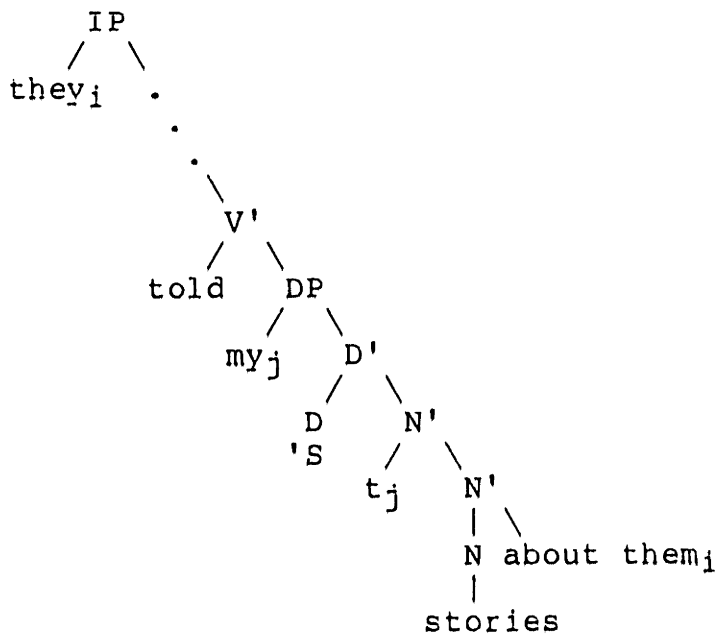
(52) (=(46c))



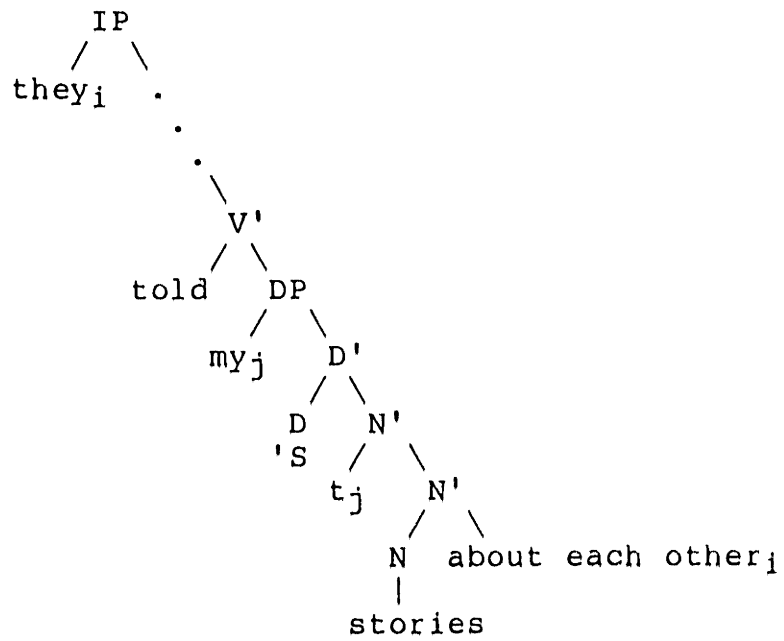
(53) (=(47c))


IP
they$_i$
V'
heard
D'
D
e
N'
PRO$_j$
N'
N
stories
about them$_i$

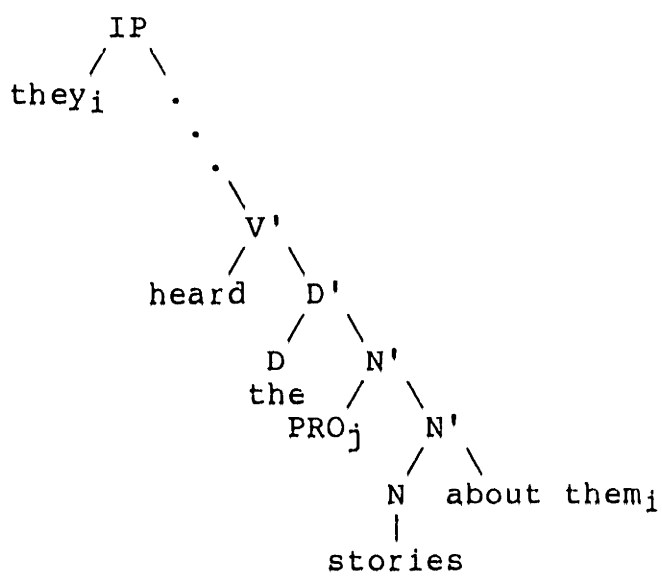
(54) (=(47d))


IP
they$_i$
V'
told
DP
my$_j$
D'
D
'S
N'
t$_j$
N'
N
stories
about them$_i$

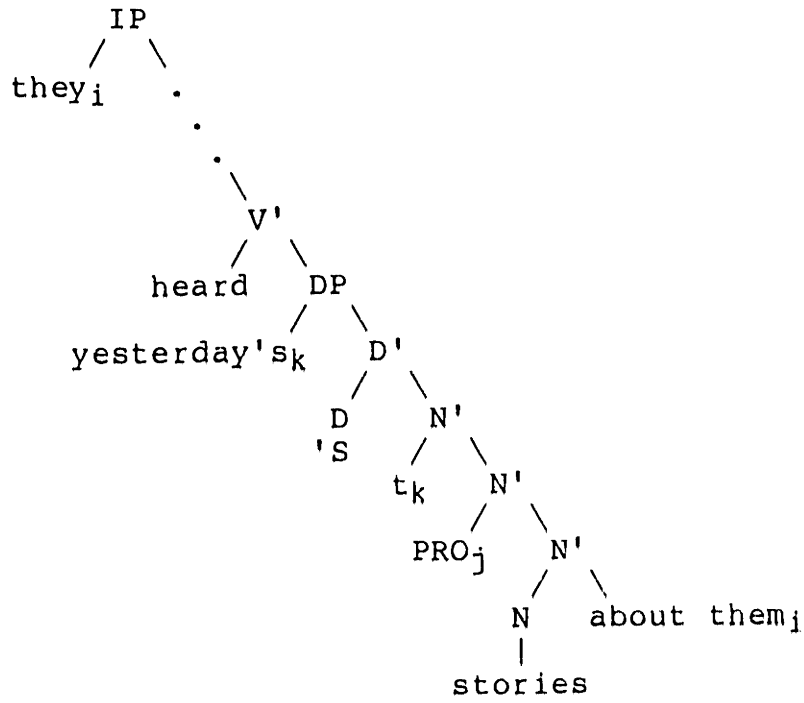
(55) (=(47b))


IP
they$_i$
V'
told
DP
my$_j$
D'
D
'S
N'
t$_j$
N'
N
about each other$_i$
stories

(56) (=(49c))


IP
they$_i$
V'
heard
D'
D
the
N'
PRO$_j$
N'
N
about them$_i$
stories

(57) (=(50c))



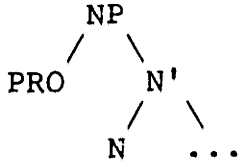
In (52), there is no subject (an external argument of the nominal head) in a noun phrase, so the Binding Theory (B) prohibits the binding of them by they. In (53), on the other hand, there is an external argument, namely PRO, in a noun phrase; therefore, the binding of them by they is allowed by the Binding Theory (B) under the designated indexing of PRO. Also, in (54), there is an external argument of a noun, in this case it is an overt element my, which starts out in a position indicated by the trace t_j at D-structure, and later moves into the specifier of DP position to get Kase. And due to the existence of this

external argument, binding of them by they does not violate the Binding Theory (B). Compare this with (55), in which an anaphor each other appears in place of a pronominal them. The structure is ill-formed, because of the presence of the external argument my, an impossible binder for each other, resulting in a violation of the Binding Theory (A). Note that postulation of PRO (binding each other) is impossible since the existence of such an element will violate the θ-criterion (if we assume, as seems plausible, that the noun stories has only one external θ-role). In (56) and (57), a postulated "implicit argument" PRO acts as a specified subject with a different index from that of them, satisfying the Binding Theory (B). In these cases, a determiner head the is base-generated under D in (56), and an adjunct yesterday is base-generated within a projection of N and is moved into a specifier of DP position to receive Kase; both of these are quite independent of the existence of an "implicit argument" PRO.

Summarizing so far, in all of the relevant cases I have been considering, the existence of an overt determiner the or the presence of an adjunct element in the "determiner" (specifier) position has nothing directly to do with the existence of an "implicit argument" PRO, the (optional) existence of the latter being determined solely by the lexical property of a nominal head, as shown by the

parallelism between (47), (49), and (50). And structural configurations given to the relevant examples in our system of projection ((52), (53), (54), (55), and (57)) explicitly represent this lack of correlation between the "determiner" (specifier) position and the existence of PRO. Also, the structure (55), coupled with the θ-criterion, correctly predicts the impossibility of binding of <u>each other</u> by <u>they</u>, since <u>my</u>, which happens to be in a specifier of DP position at S-structure for Case reasons, acts as a specified subject. Thus, our system of projection makes explicit the fact that the apparent correlation between the determiner position and the existence of PRO in noun phrases is superfluous. Note in passing that the position assigned to PRO in our system will avoid the problem connected with the "PRO theorem," which potentially arises in the standard configuration. That is, if PRO appears in the determiner position of the standard noun phrase structure such as follows, as suggested by Chomsky (1986):

(58)



it is (lexically) governed by the head noun, under the assumption that government is defined in terms of "m-command," since there is no maximal projection dominating the head noun but not dominating PRO (Recall that N' is not a maximal projection under the standard assumptions. See Aoun and Sportiche (1983) for the discussion of the position of PRO in a configuration like (58) with respect to government.). This is in direct conflict with what the "PRO theorem" dictates, i.e., "PRO is ungoverned."[21] On the other hand, in the system I am proposing, the problem about the "PRO theorem" does not arise, since, as I have argued above, the position of PRO in a configuration like (51) is not (at least lexically) governed for the reasons already discussed concerning clausal cases.[22]

Another piece of evidence in favor of the position of PRO within a projection N (rather than in the determiner position) can be obtained from Japanese. In Japanese, as well as in may other languages (cf. Hale (1981), Huang (1982)), certain classes of nouns, generally those indicating "inalienable possession" (e.g. ude 'arm,' asi 'leg,' etc.) or the "kinship" relation (e.g. okaasan 'mother,' tuma 'wife,' etc.), show an "obligatory control-like" property. For example, in the following sentences,

(59) a. John-wa tuma-o nagutta
-Top wife-Acc hit
'John$_i$ hit his$_i$ wife'

b. John-wa ude-o otta
arm broke
'John$_i$ broke his$_i$ arm'

tuma 'wife' in (59a) and ude 'arm' in (59b) must necessarily mean "John's wife" and "John's arm," respectively, and they can never mean someone else's wife or arm. It is not clear at this point whether the "obligatory control-like" property of these classes of nouns should be accounted for by grammatical principles (cf. Washio (1983)), but to the extent that these phenomena are to be handled by grammar (in the narrow sense), the most plausible way of explaining the obligatory coreference between John and the "possessor" of the nouns in (59) would be to postulate PRO in a noun phrase, and to stipulate that this PRO is obligatorily controlled by John. (Recall that the existence of PRO is determined by a noun's lexical property. Thus, only a certain class of nouns allow the existence of PRO).

(60) a. John$_i$-wa [PRO$_i$ tuma ŧo nagutta
b. John$_i$-wa [PRO$_i$ ude ŧo otta

However, it is arguably clear that Japanese lacks determiners and, furthermore, specifiers of noun phrases (the latter is a consequence of the former in our system), as we will see in detail in Chapter 4. This fact does not affect the possibility of PRO in noun phrases in our system, since PRO is postulated within a projection of N. On the other other hand, if PRO is assumed to be in a determiner position, it is predicted that there should be no PRO in noun phrases in a language which lacks determiners, for example, Japanese, which makes it impossible to give an account such as the one just suggested for the obligatory coreference relation observed in Japanese examples (59). Therefore, if the above-mentioned Japanese phenomenon turns out to be the one which should be accounted for by some grammatical principle, it constitutes evidence for the position of PRO postulated in our system of projection.

In short, we have seen in the preceding discussion that the possibility of the existence of external argument within a projection of Lexical category can naturally extend to the analysis of noun phrase internal structure with respect to some binding facts. Let us finally take a brief look at how the hypothesis of "external argument within a Lexical projection" advanced in the previous discussion works for the "wanna-contraction" case.

A phenomenon called "wanna-contraction" has been discussed extensively in various works (see, for example, Bresnan (1971), Selkirk (1972), Lightfoot (1976), Chomsky (1977a, 1981, 1986), Chomsky and Lasnik (1977, 1978), Jaeggli (1980), Bouchard (1982), Pesetsky (1982), Postal and Pullum (1982), Milsark and Safir (1983), Aoun and Lightfoot (1984), Lasnik and Saito (1984), and many others.). One of the important problems connected with wanna-contraction is how to distinguish PRO and trace with respect to their "blocking ability" for the contraction rule, whose rough formulation is as follows.

(61) want + to --> wanna

(Chomsky (1986:162))

Now consider the standard paradigm of wanna-contraction.

(62) a. who_i do you want [t_i^2 [PRO to visit t_i^1]]

b. who do you wanna visit?

(63) a. who_i do you want [t_i^2 [t_i^1 to visit Bill]]

b. *who do you wanna visit Bill?

The ill-formedness of (63b) can be explained by claiming that the intervening trace in the subject position blocks the contraction rule. Two problems arise as to the well-formedness of (62b), however (cf. Lasnik and Saito (1984:273)).

(62) a. why does PRO not block the contraction rule?

b. why does t_i^2 (in the specifier of CP under the approach we are assuming, or in COMP under other approaches) not block the contraction rule?

There have been two approaches proposed toward the solution to these problems. One approach takes the distinction between Case-marked empty categories and those that are non-Case-marked as crucial for blocking contraction: only Case-marked elements blocks (or is "visible" for) the contraction rule (cf. Jaeggli (1980), Chomsky (1981, 1986)). Thus, PRO does not block contraction because it is not Case-marked. Also, the trace in the specifier of CP does not block the contraction rule for the same reason (assuming want to be a non-Caseassigner). Only the subject trace, which is Case-marked presumably by a complementizer for (see Chomsky (1977a, 1981)), is "visible" for the contraction rule, thereby blocking its application

in (63). Another approach is the one proposed by Pesetsky (1982:Chapter 3). Pesetsky points out that the linear order of constituents, in particular the subject, is exclusively determined by general principles of grammar such as Case Theory in the rule-free system advocated by, for example, Chomsky (1981, 1982, etc.) and Stowell (1981). Thus, the lexical subject and the subject trace left by wh movement must appear in the sentence-initial position due to the Case adjacency condition of Stowell (1981). On the other hand, PRO (and trace of NP movement) is not subject to the Case adjacency condition because no Case is assigned to it. Therefore, there is an option for PRO to show up in a position other than the sentence-initial position. For example, (65) is a possible S-structure representation for (62).

(65) who_i do you want $[t_i^2$ [[to visit t_i^1] PRO]]

In (65), PRO no longer intervenes between want and to and hence does not block contraction. As for the problem (64b), i.e., the problem of the intermediate trace, a similar argument applies: Since no principle of grammar demands its presence,[23] it can be absent. Thus, in (62), nothing intervenes between want and to, and the contraction

rule applies under the maximally simple requirement on adjacency between want and to, whereas in (63), the subject trace is forced to be in a position between want and to, due to the Case adjacency condition, and thus blocks the application of the contraction rule.

As noted by Lasnik and Saito (1984), however, one problem still remains under this account. Consider the following example, in which an adjunct how is extracted.

(66) how do you wanna solve the problem?

Contraction is possible in (66). Thus, under Pesetsky's account, nothing can intervene between want and to at S-structure. However, (66) is crucially different from (62) in that an adjunct is extracted in the former, whereas what is extracted in the latter is an argument, i.e., a complement of a verb. Therefore, even if no principle (except for subjacency. see footnote 23) requires the existence of the intermediate trace in (62), (some version of) the ECP requires the intermediate trace to be present as a proper governor in the case of (66) at least at the level of LF, or the original trace of how (or the one in a VP-adjoined position, depending on the analysis of the ECP) will not be properly governed in violation of

the ECP. Note that we have been assuming, following Chomsky (1981, 1986), that the contraction rule applies in PF, to which S-structure is an input. Now we have an apparent paradox. At S-structure, (66) cannot have an intermediate trace so that contraction is possible, but at LF there must be such an intermediate trace to satisfy the ECP.[24]

Our system of projection makes it possible to give an account of wanna-contraction which is close to Pesetsky's (1982) very natural account,[25] but without the problem just mentioned. Recall that in the projection system we are assuming, an external argument, including PRO, appears within a Lexical projection at D-structure. PRO, unlike lexical subjects, need not move into the specifier of IP position to receive Kase. Thus, the S-structure representations for relevant examples should be as follows (irrelevant details omitted), assuming here, following Chomsky (1981), etc., that want takes a maximal projection of C, rather than that of I, as its complement.

(67) (for (62))

who$_i$ do you want [$_{C'}$ [$_{I'}$ to [$_{V'}$ PRO [$_{V'}$ visit t_i]]]]

(68) (for (63))

who$_i$ do you want [$_{C'}$ [$_{IP}$ $\underline{t}_i^2$ [$_{I'}$ to [$_{V'}$ $\underline{t}_i^1$ [$_{V'}$ visit Bill]]]]]

(69) (for (66))

how$_i$ do you want [$_{C'}$ [$_{I'}$ to [$_{V'}$ $\underline{t}_i$ [$_{V'}$ PRO [$_{V'}$ solve the problem]]]]]

In (67) and (69), want and to are (string) adjacent to each other, with PRO in the position structurally "lower" than to (i.e., within the projection of V). Contraction is thus possible under the straightforward adjacency condition. In (68), on the other hand, the embedded subject who must "land" in the specifier of IP position to get Kase (otherwise, the chain headed by who would contain no Kase position, resulting in a violation of general chain condition (cf. Chomsky (1981, 1986)), leaving behind the trace $\underline{t}_i^1$, which breaks the adjacency between want and to. Therefore, the contraction rule cannot apply to (68), and hence (63b) is ruled out.

To sum up, our system of projection provides, as one of its consequences, a maximally simple account of the wanna-contraction phenomenon, which is very close to Pesetsky's (1982) analysis. Conceptually, these two accounts are different (aside from the difference in

treatment of the intermediate trace in the specifier of CP) in that in Pesetsky's account, PRO may or may not intervene between want and to, since no principle forces it to be in a particular position in a relevant structure, whereas, in our system, PRO can never intervene between want and to. Whether or not this conceptual difference yields any empirically relevant difference is not clear at this point.

3.3 The Notion of Maximal Projection and the Status of the X-bar Schema

In the system of projection that I have proposed in Chapter 2, the notion "maximal projection" should have a quite different content than the one generally assumed in the literature. In particular, it is impossible in our system to define the notion of maximal projection in terms of the number of bars of a given node. Also, the status of the X-bar schema as a well-formedness condition on D-structure is not entirely clear in our system of projection. In this section, I will briefly discuss these problems (For a more thorough discussion of these matters,

the reader is referred to Speas (forthcoming)).

As I just mentioned, "maximal projection" cannot be defined in terms of the number of bars of a given node in our projection system. There are two factors in our system which makes this familiar definition (definition in terms of the number of bars) impossible. First, I explicitly took a "non-uniform" view of the category projection: Functional categories can project up to a double-bar level with the presence of a specifier, whereas Lexical categories can project only to a single-bar level. But, if this is the only factor that differentiates our system of projection from the standard one, we could still define the notion of maximal projection in terms of the number of bars by distinguishing these two types of categories, namely, we could stipulate that the maximal projection of X is X" if X is a Functional category; it is X' if X is a Lexical category. However, I have also stated as the Functional Projection Theorem (46) that a Functional head projects up either to a single-bar level or to a double-bar level depending on the presence/absence of a Kase to be discharged to its specifier position. This amounts to saying that, given a Functional head, either a double-bar projection or a single-bar projection can be a "maximal projection" of that head, if we want to maintain a rather plausible assumption that only maximal projections may appear as

non-head terms within a phrase (cf. Stowell (1981), Chomsky (1981, 1985)). For examples, in (70a), IP should be the maximal projection of I, whereas in (70b), the maximal projection of I should be I', both of which are "selected" by <u>believe</u> and <u>seem</u>, respectively.

(70) a. John believes $[_{IP}$ Bill$_i$ $[_{I'}$ to $[_{V'}$ have t_i gone home]]]

b. John$_i$ seems $[_{I'}$ to $[_{V'}$ have t_i finished his thesis]]

These considerations lead us to a definition of maximal projection based on something other than the number of bars. Here, let us try to define the notion of maximal projection in terms of "projection path" defined as follows.

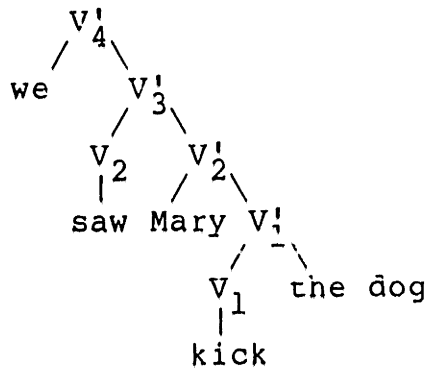
(71) π is a <u>projection path</u> iff

π is a sequence of nodes N=(n_1, ..., n_n) such that

(i) $\forall_i$, n_i immediately dominates n_{i+1}
(ii) all n_i have the same set of FEATURES, and
(iii) the bar level of n_i is equal to or greater than the bar level of n_{i+1}

The term "FEATURES" in (71ii) is used here in a somewhat extended sense. That is, "FEATURES" include not only the syntactic features in the standard sense, but also the set of θ-grids associated with a Lexical head. This extention is necessary because of the so-called "small clause" construction. For example, in the following structure, we have to distinguish V_2' and V_3' in order to say that the projection path starting with the verb V_1 "stops" at V_2'.

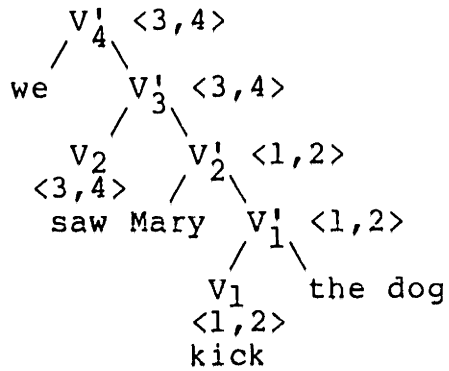
(72)



What we want to say concerning (72) is that V_1, V_1', and V_2' constitute a projection path, and that V_2, V_3', and V_4' constitute another projection path. However, it is impossible to distinguish between V_2' and V_3' in terms of "syntactic features" in the standard sense. Thus, we have to look at each verb's θ-grid in addition to its syntactic features, in order to correctly distinguish V_2' and V_3' in (72). Suppose now that the transitive verbs kick and see

have θ-grids <1,2> and <3,4>, respectively, and that these θ-grids, being one of verb's "FEATURES," project along with other features (FEATURES) of a Lexical head (cf. Higginbotham (1985)). Then, a more appropriate structure for (72) should be such as follows.

(73)



In (73), we can distinguish between V_2' and V_3' by looking at the θ-grids associated with them. V_2' and V_3' cannot form a single projection path because they do not share the set of FEATURES. Thus, there are two projection paths in (73): $\pi_1 = (V_2', V_1', V_1)$, $\pi_2 = (V_4', V_3', V_2)$. Each "path" satisfies the conditions stated in (71).

We now define the "maximal projection node" as follows.

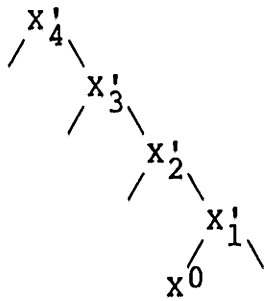
(74) n_i is the <u>maximal projection node</u> of a projection path $\pi = (n_1, \ldots, n_n)$ iff $i = 1$.

That is, the "top" node, the node that dominates every other node and is not dominated by any node, of a given projection path is the maximal projection node of that projection path. Based on this notion of maximal projection node, we can define the "maximal projection category."

(75) α is the <u>maximal projection category</u> iff
α is a projection path $\pi = (\beta_1, \ldots, \beta_n)$
such that
(i) β_1 is the maximal projection node, and
(ii) all β_i have the same number of bars

To illustrate, consider the following hypothetical configuration, where every node shares the set of FEATURES.

(76)



The projection path in (76) is $\pi = (X'_4, X'_3, X'_2, X'_1, X^0)$. The maximal projection node of the projection path π is X'_4. And the maximal projection category is $\alpha = (X'_4, X'_3,$

X'_2, X'_1).

Note finally that under these conceptions, the role of the X-bar schema is a quite limited one. Specifically, the X-bar schema is now reduced to the condition (ii) and (iii) in the definition of a projection path (71), i.e., the condition that all FEATURES of a head project up through a projection path, and the condition that the number of bars cannot "decrease" in the course of projection. The former condition seems to be dispensed with, at least in the case of a projection of a Lexical head, by means of Higginbotham's (1985) mechanism of constructing a constituent structure.

Thus, in our projection system, the status of X-bar schema is reduced to a convention on FEATURE percolation (projection), especially in the case of a projection of a Functional head to which Higginbotham's (1985) mechanism of projection does not apply, plus a prohibition on "bar reduction." If the latter condition can be derived from something more general, as seems possible, then the minimum content of the X-bar schema in our system of projection will be simplest possible statement that FEATURES of a lexical item must project. For further discussion, see Speas (forthcoming).

Notes to Chapter 3

1. I am indebted to Howard Lasnik for these examples.

2. There is a technical problem here which has to do with the interpretation of the base-generated "recursion" (or "adjunction") structure generally allowed for Lexical categories in my system, in particular those in (11). If we assume that May's (1985) distinction between "categories" and "segments" (cf. Chapter 1) holds at D-structure, i.e., that it holds for base-generated structures such as those in (11), and consequently assume that the definition of "dominance" based on this distinction also holds for base-generated structures like (11), then in (11a), for example, a category V' which consists of segments V'_1 and V'_2, and the segment V'_2 are both sisters of the subject NP. Thus, we can say either a category V' directly θ-marks the subject NP or a segment V'_2 directly θ-marks the subject NP, under the assumption that sisterhood is the only condition for direct θ-marking. Although it is not clear that these two options make different empirical predictions, I will tentatively assume in the following discussion that May's distinction between categories and segments with respect to the definition of dominance holds only for "true" adjunction cases (presumably limited to the rules of LF, in particular QR, putting aside for the moment rules like Heavy NP Shift, etc.), i.e., adjunction to XPs, creating a structure that is not base-generable (cf. also Chapter 4), and that in the structures like

(11), the notion of "dominance" is defined in terms of "nodes" rather than categories (Note incidentally that the notion of sisterhood given in (2) is defined in terms of "nodes"). Thus, in (11a), it is the node V'_2 that directly θ-marks the subject NP. This tentative conclusion is quite consistent with the "bottom-up" manner of θ-marking mechanism to be proposed below.

3. In order to determine the "closeness" of a given θ-role to the predicate, we will probably have to look at the type of the θ-role, for example, "Theme" is generally the closest θ-role, "Agent" is generally the least close θ-role, etc. Or, the property of each θ-role with respect to its "closeness" might be derived from the general properties of the Lexical Conceptual Structure.

4. Another possibility is that all non-arguments are absent at D-structure, being introduced in a different dimension, and will be later (at S-structure, perhaps) "hooked up" to the skeletal structure which consists of only heads and its arguments. Although this approach seems plausible, I will not pursue this possibility here.

5. I put aside here the problem of PRO in gerund, although my analysis can have various interesting consequences for the analysis of gerund. See Reuland (1984) and Abney (1985) for the discussion of gerund.

6. A characterization of "adjunction" in the projection system I am proposing will be suggested in Chapter 4 below.

7. The notion of "maximal projection" in our system of projection will be defined in 3.3.

8. Note in passing that considerations on the barrier for PRO is irrelevant here as long as V, the alleged governor of PRO, does not m-command it, since, as I mentioned above (cf. also Chapter 1), m-command is the necessary condition for government. Thus, the fact that the alleged governor does <u>not</u> m-command its governee entails the lack of government of the alleged governee by its "governor," which is quite independent from considerations on barriers.

9. Chomsky (1986), as well as Roeper (1984), presents a number of properties distinguishing "syntactically present" PRO and "lexically present" implicit argument. Chomsky (1986) gives, among others, the following contrasts between PRO and "implicit argument."

(i) a. they expected [PRO to give damaging testimony]

b. *they expected [damaging testimony to be given]

(ii) a. it is impossible [PRO to visit together]
b. *it is impossible [for me to be visited together]
(Chomsky (1986:119-120))

The contrast in (i) shows that only PRO, a syntactically present element, can be controlled by an antecedent. In (ii), an adjunct together can be predicated of PRO in (iia), but it cannot be predicated of an "implicit argument" as shown by the ungrammaticality of (iib).

The contrast in (i) can be accounted for even under the assumption that the "implicit argument" in passive is "syntactically present" (presumably PRO), if we assume Manzini's (1983) theory of control. In (ia), the matrix subject they is, roughly speaking, the closest accessible subject for PRO; thus the control of PRO by they is allowed. On the other hand, after the application of Move-α, it is a noun phrase damaging testimony, and not they, that is the closest accessible subject for an "implicit argument" in (ib). Therefore, the control of an "implicit argument" by they in (ib) is impossible, assuming that the principle of control applies at S-structure (Note that "control" is in fact an instance of "binding" in Manzini's (1983) system. See Manzini (1983) for details). As for (ii), there seems to be an alternative account of the observed contrast, too, if we assume, along the lines suggested by Epstein (1984), that an "arbitrary" PRO like the one in (iia) is actually bound by an empty category licensed by a benefactive for (Epstein (1984) identifies this element as pro which functions as a kind of universal

quantifier), whereas the "implicit argument" in (iib) does not have such a binder; it is a purely "free" PRO. For an adjunct like <u>together</u> to be predicated of <u>x</u>, it is necessary that <u>x</u> has "enough pluralistic meaning." The PRO in (iia) gets this pluralistic meaning from its binder (universal quantifier <u>pro</u>), but the "implicit argument" in (iib) cannot obtain the necessary pluralistic meaning due to the lack of binder which bears such meaning. Hence the impossibility of predication in (iib). This approach is further supported by the following example.

(iii) *John wondered how PRO to visit together

In (iii), even under the "arbitrary" interpretation (Notice that the "arbitrary" PRO is possible (for many speakers) in the configurations similar to (iii), <u>John wondered how PRO$_{arb}$ to fix the sink</u>), the modification of PRO by <u>together</u> is impossible. This seems to indicate that what determines the possibility of predication by adjuncts like <u>together</u> is not whether or not the subject is the "syntactically present" PRO or "lexically present" implicit argument, but rather whether or not the subject bears "enough pluralistic meaning." PRO in (iii) does not get such pluralistic meaning in the absence of "benefactive" <u>for</u>, just like the implicit arguments do not obtain such "enough pluralistic meaning" due to the lack of "benefactive" <u>for</u> (and hence the lack of "universal quantifier" <u>pro</u>.) Chomsky (1986) presents various other arguments that PRO and "implicit argument" are different in nature, of which I have no alternative account at this point.

The issue of "implicit argument" has been a focus of much recent discussion and is far from being settled. As discussed so far, Chomsky (1986), as well as Roeper (1984), considers "implicit argument" as distinct from "syntactically present" argument (such as PRO), whereas Roberts (1985), Jaeggli (1986), among others, regard it as syntactically present. The system I have proposed in Chapter 2 provides a structural position for the "implicit argument" if it is syntactically present, and thus leading us to the latter approach, namely the approach under which "implicit argument" is considered to be "syntactically present." However, the system is neutral as to the determination of exact characterization of this "syntactically present" element (e.g., whether it is "PRO," "pro" (in the latter case it is probably "locally determined" by the "agentive" by, just as benefactive "pro" is locally determined by for, as suggested by Epstein (1984)) or "EN" (cf. Roberts (1985) and Baker, Johnson and Roberts (in progress)), etc.). In what follows, I will tentatively assume that the "implicit argument" in passive is identified as PRO, and explore various implications of this assumption in my system. This does not mean, however, that the following discussion is intended to solve the problem of "implicit argument." For relevant discussions, see, in addition to the above-mentioned works, Lasnik (1984), Williams (1985), Kayne (1986), among others.

10. This implies that (at least in the case of passive) the so-called "Burzio's generalization" (see Burzio (1986))

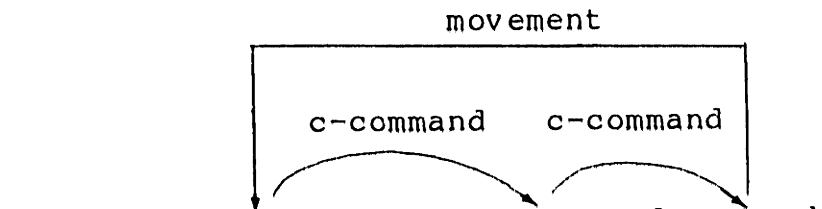
is not bi-directional. I will not explore the implications of my proposal for the proper characterization of "Burzio's generalization." See Massam (1985), Burzio (1986), and Chomsky (1986) for relevant discussions.

11. Chomsky's (1981:333) formulation makes use of the notion "BIND" rather than "bind," where "BIND" means, roughly, "bind" plus the case of "co-superscripting." This distinction is not relevant to our present concern, so I will use the term "bind" here and in what follows. The notions "X-bind" and "locally X-bind" are defined as follows.

> (i) α is <u>X-bound</u> by β if and only if α and β are coindexed, β c-commands α, and β is in an X-position
>
> (ii) α is <u>locally bound</u> by β if and only if α is X-bound by β, and if γ Y-binds α then either γ Y-binds β or $\gamma = \beta$
>
> (iii) α is <u>locally X-bound</u> by β if and only if α is locally bound and X-bound by β
>
> (Chomsky (1981:184-185))

12. There is an apparent problem with Lasnik's (1985) generalization. As Lasnik himself notes (attributed to an anonymous <u>LI</u> reviewer), the derivation of "clitic climbing" in Italian has the abstract property represented in (25), but does not invoke the "crossover" effect. The following example is

well-formed under the designated coindexing.



"Gianni wants to see himself'

(Lasnik (1985:489) with adaptations)

However, Lasnik also notes (Lasnik (1985:488)) that "this observation is somewhat mitigated by the fact that such "clitic climbing" is limited to a small set of so-called restructuring verbs." See Lasnik (1985) for further discussion.

13. This amounts to saying that while Nominative Case assignment, an instance of "SPEC-head" agreement in my system, is subject to the "directionality" parameter, namely it is uniformly to the left, the relation of government itself is not subject to such a directionality requirement.

14. One might say that the trace of John$_i$ in (35a) is actually governed by the inflected verb killed in I position (Recall that "exceptional" government by a lexical head is possible, and the verb raised into I position is lexical). This could be true in this particular case, but such an account cannot be extended to the cases like the following where no V-raising takes place in a complement sentence but the "subject

raising" nevertheless applies in that clause due to the "exceptional Case-marking" by the matrix verb.

(i) John believes $[_{IP}$ Bill$_i$ $[_{I'}$ $[_I$ to] $[_V$ t_i have kissed Mary]]]

15. However, if the analysis proposed by Belleti and Rizzi (1985) on "psych" construction is right, it might be possible to impose a c-command condition on control.

16. One might claim that this account makes it unnecessary to resort to Lasnik's (1985) "generalized strong crossover constraint" invoked above to account for the obligatory disjointness between the S-structure subject and the postulated PRO, in particular, or possibly supercedes it entirely. Note incidentally that the analysis just proposed in the text correctly predicts the grammaticality of the Italian examples (clitic climbing), which is problematic to Lasnik's constraint. For the head of the chain in the Italian data is an anaphor (or at least an element of anaphoric nature) si, and thus coindexing PRO with the trace of si should not violate the Binding Theory (C). This line of research could be on the right track, although I will not pursue this possibility here because there are a set of counterexamples. Namely, in the examples such as follows, the head of the chain in question is an anaphor (each other, himself), hence the status of chain with respect to the Binding Theory should be an anaphor but PRO must nevertheless be obligatorily disjoint from the chain (headed by the S-structure

subject).

(i) they$_i$ believed [each other$_i$ to have been PRO arrested t_i]

(ii) John$_i$ believed [himself$_i$ to have been PRO treated t_i badly]

It could be possible to handle these cases by somehow extending the chain to include they/John as its head, so that the terminal elements of the chains, i.e., the traces, assuming their status from the head of the "extended" chain, act like a pronominal in (i) and an R-expression in (ii), with respect to the binding relation to PRO, a "chain-external" element. I leave this possibility open here.

17. I am assuming, following the standard analysis, that try takes CP (in our terms, the "maximal projection" of C, which is not necessarily CP. See 3.3 for the discussion on this matter.), headed by an empty complementizer.

18. The landing site for PRO_2 might be "lower" than be, namely, right above the position of PRO_1, yielding the structure: ... [$_{I'}$ to [$_{V'}$ be [$_{V'}$ PRO_2 [$_{V'}$ PRO_1 [$_{V'}$ arrested t_2]]]]].

19. The presence of PRO as an implicit argument is a noun phrase should be optional, since the example (47a), we also assume, as in the case of (47c), that the stories are someone else's, but the occurrence of each other

is still allowed. Thus, Chomsky (1986) concludes that "presence of the implicit argument as subject is optional: If present, the interpretation is fixed depending on the indexing; if absent, the interpretation is free." (Chomsky (1986:167)). For discussion, as well as other pieces of evidence for the postulation of optional PRO within a noun phrase, see Chomsky (1986). For some counterarguments, see Williams (1985b.)

20. I am tentatively assuming here that a "Determiner Phrase" without a determiner, in particular plural noun phrases, is headed by an empty determiner, parallel to the cases of "empty complementizers" in clauses. Although this assumption does not have any direct bearing on the present discussion, and will not be discussed in what follows, it has various non-trivial implications for the proper characterization of the distribution of PRO in noun phrases. Also, the postulation of an empty determiner has direct bearings on the problem of "θ-binding (in the sense of Higginbotham (1985)) in nominals. These matters are not at all trivial, but I will not pursue these issues here, pending further research.

21. There appears to be a couple of possible ways of handling this problem. One possibility is to postulate a node such as DET dominating PRO in (58), and to assume it to be a maximal projection, protecting PRO from government by the head noun. This seems to be the position Chomsky takes in Chomsky (1986), where he

suggests "Perhaps, then, the DET position can include an implicit argument with the properties of PRO, ..." (Chomsky (1986:167)), although it is not entirely clear to me how "the DET position" is characterized with respect to government relation. Another possibility is to identify the implicit argument in question as pro, rather than PRO. This solves the problem concerning the "PRO theorem," simply because pro, being a pure pronominal, is not subject to the "PRO theorem." Problem still remains, however, as to how this pro can be "locally determined" (cf. Chomsky (1982)).

22. Government from the verb when D projects up to the D' level is blocked perhaps by the "minimality" condition (D is the "closer" nonlexical governor for PRO) of Chomsky (1985). See the discussion above.

23. Aside form subjacency, under the standard conception of this principle. If subjacency is a condition on representation, then it does require the existence of the intermediate trace. Thus, in Pesetsky's (1982) account, subjacency is crucially assumed to be a condition on movement.

24. Lasnik and Saito (1984) solve this apparent paradox by creating a necessary intermediate trace through the lowering (and subsequent raising) of <u>how</u> in LF, a legitimate operation in their system.

25. Note however that our account is not incompatible with the first approach discussed above.

CHAPTER 4

PHRASE STRUCTURE OF JAPANESE

Japanese has a somewhat peculiar status in the study of generative grammar in that while a tremendous amount of descriptive work has been accumulated, it is still not known what its configurational structure looks like. This is in contrast to the situation of languages like English where the basic phrase structure is, to a reasonable extent, well established and where syntactic arguments for or against some proposed analysis can be constructed on the basis of a reasonably uncontroversial phrase structural configuration. In some cases, for example, different analyses of a single phenomenon in Japanese assume totally different, sometimes even contradictory, configurational structures of the construction in question. Recently,

however, some very interesting works on Japanese phrase structure have emerged (Hoji (1982, 1985), Kuroda (1980, 1983), Saito (1982b,1983,1985), Whitman (1982)), most of which have been inspired by the stimulating research program proposed by Ken Hale on 'configurationality' parameters (cf., among others, Hale (1980, 1983)). It seems to me that we are now in a position to propose an overall picture of Japanese phrase structure, incorporating various observations and insights presented in the above-mentioned works on this issue, so that an analysis of a particular syntactic phenomenon in this language can be tested based on the phrase structural configuration, and, of course, the proposed phrase structure itself can be modified in light of such analyses of various syntactic phenomena in the language. This chapter is intended to be an attempt to lay down the foundation on which such a fruitful interaction can be made. In what follows, I will try to give a general picture of Japanese phrase structure in the light of the system of projection I have proposed in the preceding chapter. In particular, I will concentrate on the basic structure of clauses and noun phrases in Japanese. As the result of this orientation of this chapter, the discussion of each syntactic phenomenon will be somewhat sketchy. A detailed examination of various syntactic phenomena in Japanese based on the general idea

on the phrase structure of this language to be proposed below must follow in the future research.

The organization of this chapter is as follows: In Section 1, I will review the arguments concerning Japanese phrase structure which have been proposed in the literature so far, all of which have to do with the existence of the "VP" node in this language. Section 2 examines the status of Functional categories in Japanese. There, I will also discuss the properties of Japanese "specifiers," paying special attention to noun phrases and sentences, and will show that this language lacks specifiers in the sense defined in Chapter 2. In Section 3, a new phrase structural configuration for Japanese will be introduced based on the observations made in Section 2. And it will be shown that given this phrase structural configuration, we will readily account for the facts both for and against the existence of the "VP" node in Japanese summarized in Section 1. Futhermore, in Section 4, it will be pointed out that the phrase structure for Japanese proposed in Section 3, combined with several rather plausible assumptions, makes it possible to capture some typological characteristics of this language which have been noted in the literature but hitherto have been totally unaccounted for.

4.1 Facts to be Accounted for

In this section, I will discuss a number of facts in Japanese to which any proposed phrase structural configuration of this language must give some explanation. I will first briefly review some of the arguments for the existence of the "VP" node in Japanese and will then go over some of the "classical" arguments against such a node.

4.1.1 Evidence for the "VP" node

4.1.1 Evidence from the Binding Theory

As Saito (1985:56) notes, the existence of the VP node is indicated by various subject/object asymmetries. One of those asymmetries is found in pronominal coreference. Consider the following well-known paradigm in English.

(1) a. John$_i$ [$_{VP}$ loves [$_{NP}$ his$_i$ mother]]

b. *he$_i$ [$_{VP}$ loves [$_{NP}$ John$_i$'s mother]]

c. [$_{NP}$ John$_i$'s mother][$_{VP}$ loves him$_i$]

d. [$_{NP}$ his$_i$ mother][$_{VP}$ loves John$_i$]

The contrast in grammaticality in (1) can be straightforwardly accounted for given the following condition on pronominal coreference, which is essentially (a part of) the clause C of the Binding Theory.

(2) A pronoun cannot c-command its antecedent.

It is easy to see that among the examples in (1), only (1b) violates the condition (2), i.e., only in (1b) does the pronoun he c-command its antecedent John. The crucial evidence for the existence of VP node in English is (1c). If there were no VP node in English, the pronoun him in (1c) would c-command its antecedent John, and the sentence would be incorrectly ruled out as a violation of the principle (2). Thus, the grammaticality of (1c) under the intended coreference reading indicates that English has the VP node.

This argument can be extended to Japanese. If Japanese lacks VP node, as claimed by proponents of what Saito (1985:34) calls 'extreme non-configurational analysis of Japanese,' the corresponding Japanese sentence to (1c) should be ungrammatical with the intended coreference reading, since in that case a pronoun would c-command its antecedent due to the lack of a VP node. As Whitman (1982)

points out, however, this is not the case. Consider the following Japanese paradigm which corresponds to the English one in (1).

(3) a. John$_i$-ga [$_{NP}$ kare$_i$-no okaasan-o aisite-iru
-Nom he -Gen mother-Acc loves

'John$_i$ loves his$_i$ mother'

b.*kare$_i$-ga [$_{NP}$ John$_i$-no okaasan-o aisite-iru
he

'he$_i$ loves John$_i$'s mother'

c. [$_{NP}$ John$_i$-no okaasan-ga kare$_i$-o aisite-iru

'John$_i$'s mother loves him$_i$'

d. [$_{NP}$ kare$_i$-no okaasan-ga John$_i$-o aisite-iru

'his$_i$ mother loves John$_i$'

As we can see in (3), Japanese exhibits exactly the same pattern as English with respect to the coreference possibilities in the relevant examples. In particular, the grammaticality of (3c), like (1c), provides us with a piece of evidence that this language, as well as English, does have a VP node. (cf. Whitman(1982), Saito(1983, 1985) for more detailed discussion).

4.1.1.2 Evidence from Crossover

Another piece of evidence for the existence of a VP node in Japanese can be obtained from some weak crossover facts in this language. It is well known that a movement of some element to an A'-position, "crossing over" the pronoun with which a moved element is coindexed, makes the sentence less acceptable (cf. Postal (1971), Wasow (1972, 1979), Chomsky (1976), Higginbotham (1980), etc.).[1] If the pronoun c-commands the element to be moved before such a movement takes place, we get the "strong crossover effect," whereas if there is no c-command relationship holding between the two elements, the resulting structure is to show the "weak crossover effect," whose acceptability is generally higher than the strong crossover cases. The following examples illustrate the weak crossover effect.[2]

(4) a. ?*who$_i$ does [$_{NP}$ his$_i$ mother][$_{VP}$ love t_i]

b. ?*[$_S$ everyone$_i$ [$_S$ [$_{NP}$ his$_i$ mother][$_{VP}$ loves t_i]]]

In (4a), who$_i$ is moved from its original position marked by t_i to a sentence initial A'-position by a "syntactic" Move-α, crossing over the coindexed pronoun his$_i$. In (4b), a quantifier phrase everyone is moved by an LF rule

QR and is adjoined to S (cf. May (1977)), again crossing over the coindexed pronoun $\underline{his}_i$. Thus, the following LF configuration schematically represents the weak crossover cases.

(5) [Operator$_i$ [... pronoun$_i$ ... $\underline{t}_i$ ...]] (order irrelevance), where neither the pronoun nor the variable (=$\underline{t}_i$) c-commands the other.

(cf. Saito (1985:91))

Both (4a) and (4b) fit in the configuration (5) at LF, and, as expected, the corresponding sentences "who$_i$ does his$_i$ mother love?" and "his$_i$ mother loves everyone$_i$" show marginal acceptability under the intended coreference reading. Note crucially that in (5), neither a pronoun nor the variable should c-command the other, otherwise we would not obtain the weak crossover effect.[3]

Based on this weak crossover effect, there have been two arguments proposed for the existence of a VP node in Japanese; one using an "anaphor" zibun 'self' (cf. Saito and Hoji (1983)), and the other using a null pronominal (Saito (1985)). Here, I will briefly go over the latter argument.[4] Consider first the following examples taken from Saito (1985:103) with minor modifications, (Judgments

are Saito's).

(6) a. [$_{S'}$ John-wa [$_{PP}$ Mary-ga pro$_i$ yomu mae-ni]
-Top -Nom read before

[$_{VP}$ sono hon$_i$-o yonda]]
that book-Acc read

'John read that book$_i$ before Mary read it$_i$'

b.?*[$_{S'}$ John-wa [$_{PP}$ Mary-ga pro$_i$ yomu mae-ni]

[$_{VP}$ dono hon$_i$-o yonda] no]
which book Q

'which book$_i$ did John read before Mary read it$_i$'

In (6a), an empty pronominal pro (cf. Chomsky (1982) for relevant discussion of this element), which is assumed to have the feature [+pronominal, -anaphor] and hence to obey the Binding Theory (B), is "free" (i.e., not bound), simply because its possible antecedent sono hon 'that book' does not c-command it. Or, even if there were no VP node, pro would be free in its "governing category" (see Chomsky (1981, 1986)), which presumably is an adverbial clause or postpositional phrase dominating the adverbial clause. In any event, the grammaticality of (6a) is reasonably without problem either for the structure with a VP node or for the one without it. What is problematic is the marginal status of (6b), which is minimally different from

(6a) in that a wh element dono hon 'which book' appears in place of a non-wh element sono hon 'that book.' However, this contrast between (6a) and (6b) can be straightforwardly accounted for as a case of weak crossover which is schematically represented in (5), on the assumption that wh elements in situ are moved in LF.[5] Consider the following LF representation for (6b).

(7) [S' John-wa [PP Mary-ga pro_i yomu mae-ni]
[VP t_i yonda] no dono hon$_i$]]

The LF representation (7) clearly fits in the weak crossover configuration: an operator dono hon$_i$ 'which book,' its trace t_i, and pro_i are all coindexed, and neither pro_i nor t_i c-commands the other. On the other hand, weak crossover is irrelevant for (6a), because sono hon 'that book,' being a non-wh element, does not move in LF. Thus, we can account for the contrast in (6) based on the weak crossover configuration (5).

It should be noted that this account of the contrast in (6) crucially assumes the existence of a VP node. If there is no VP node in (7), then the variable t_i would c-command pro_i. However, we know independently that there is no weak crossover effect in such cases. Saito

(1985:104) cites the following example to illustrate this point.

(8) dare$_i$-ga [$_S$' pro$_i$ Mary-ni kirawarete iru to]
who-Nom -by be-disliked that

omoikonde iru no?
be-convinced Q

'who$_i$ is convinced that he$_i$ is disliked by Mary'

After the application of an LF wh movement, we get the following LF representation:

(9) [$_S$' [$_S$ t$_i$ [$_S$' pro$_i$ Mary-ni kirawarete iru to]
omoikonde iru] no dare$_i$]

In (9), the variable t$_i$ c-commands pro$_i$, and the corresponding sentence (8) is in fact grammatical under the intended interpretation.

Therefore, in order to rule (6b) out as an instance of weak crossover, there must be no c-command relationship between pro$_i$ and t$_i$ in (7), which requires us to assume a VP node. Thus, the contrast between (6a) and (6b) provides additional evidence for the VP node in Japanese.

4.1.1.3 Evidence from the Distribution of PRO_{arb}

Kuroda (1983) cites another piece of evidence for the existence of the "VP" node in Japanese from the distribution of PRO_{arb} in this language (See also Saito (1982a)). As is well-known, the distribution of PRO is constrained by the condition usually called the "PRO theorem," which dictates that PRO be ungoverned (or "not lexically governed," see Chapter 3). This condition accounts for the subject/object asymmetry in English with respect to the distribution of PRO, i.e., PRO can only appear in subject position and can never appear in object position. For, in English, object position is always (lexically) governed whereas subject position may not be. Kuroda (1983) points out that this subject/object asymmetry found in English can also be observed in Japanese. To show the existence of the subject/object asymmetry in Japanese, he takes up the distribution of PRO_{arb} in the language,[6] and gives the following contrast.

(10) a. [PRO_{arb} sensei-ni au koto wa muzukasii
PRO_{arb} teacher with meet -Top difficult

'it is difficult to meet teachers'

b. *[gakusei-ga PRO_{arb} au koto wa muzukasii

Lit. 'it is difficult for students to meet'

(Kuroda (1983:154) with adaptations)

The contrast in (10) shows that in Japanese, as well as in English, PRO_{arb} may appear in subject position and, more crucially, it may not appear in object position. This asymmetry with respect to the distribution of PRO_{arb} in Japanese can straightforwardly accounted for, given the "PRO theorem," if we assume that Japanese, like English, has a VP node. For, in that case, the verb will govern object position, but not subject position due to the existence of VP, as desired.

In this subsection, I have reviewed some of the arguments for the VP node in Japanese.[7] In the next subsection, I will briefly summarize the "classical" arguments against such a node in this language.

4.1.2 Evidence against the "VP node"

In this subsection I will briefly go over some of the "classical" arguments against the VP node in Japanese presented in Hinds(1973) and another piece of such evidence pointed out in Whitman(1982, 1984).

4.1.2.1 Evidence from VP Movement Rules

It is well known that if a transformation applies to an element x, then x is a constituent. More specifically, if a transformation (Move-α) can apply to x, then x must be a phrasal category (X^{max}) or a head.[8] Thus, if there is a transformation which moves a category in a language, then that constitutes evidence for such a category in the language. As for the VP node, English clearly has such a movement rule which specifically refers to VP. This is illustrated by the following example (Whitman (1984:13)).

(11) Heather promised to come at 10, and come at 10 she did.

As Whitman describes, this "VP fronting" rule involves detachment of the VP from the subject and INFL (auxiliary and tense), and moves the VP to the front. Such a detachment is simply impossible in Japanese.

(12) *Susan-wa zyuuzi-ni kuru to yakusokusita, sosite
-Top ten -at come that promised and
o'clock

(zizitu) zyuuzi-ni ki/ku kanozyo-wa ta
in fact come she past
(stem)

Lit. 'Susan promised to come at 10, and (in fact) come at 10 she Past'

The absence of the corresponding "VP fronting" rule in Japanese suggests that this language does not have the category corresponding to English VP node (cf. Hinds (1973)).

4.1.2.2 Evidence from the Distribution of Adverbial Elements

Hinds (1973) notes that another piece of evidence against VP node in Japanese can be obtained from the distribution of adverbial elements in this language. His

that in languages like English, adverbials may not normally occur intervening between a verb and its direct object. He states, citing Schwartz (1972):

> He (=Schwartz (1972);N.F.) states (p.126) that "a feeling of interruptability (more accurately, a resistance toward interruptabilty) can be used as evidence for constituency relations." He further points out (p.126) that "such evidence has been used in the past ... in support of major constituency breaks." In order to use this criterion, it is necessary to show that an element which is otherwise freely positioned cannot be placed in a specific context. His example involves Indonesian, in which time adverbials can usually be placed anywhere in a sentence. However, they cannot occur in between an object and a verb. Schwartz(p.217) claims that "the intuitive sense of a bond or valence between these predicates and complements is substantiated by a probing of constituency breaks; the stronger the bond, the worse the violation."
>
> (Hinds (1973:46))

Thus, for example, in English, adverbial elements, which otherwise can generally be placed anywhere in a sentence (cf. Keyser (1968)), cannot occur between a verb and its direct object. Consider the following contrast cited

from Whitman (1982:23).

(13) a. John now is reading that book

b. *John is reading now that book

In (13a), an adverbial element now intervenes the subject and the rest of the sentence, but the sentence is still grammatical, whereas in (13b), the occurrence of such an adverbial element between the verb and its direct object makes the sentence ungrammatical. On the other hand, there is no such contrast in Japanese.

(14) a. John-ga ima sono hon-o yonde iru
-Nom now that book-Acc be reading

Lit. 'John now is reading that book

b. John-ga sono hon-o ima yonde iru

Lit. 'John is reading now that book'

(Whitman (1982:24) with adaptation)

Although this argument is, as Hinds himself notes (Hinds (1973:46-47)), not at all a strong argument against VP node, the contrast in (13) and the lack of such contrast in (14) seem to indicate that the relationship between a verb and

its direct object in Japanese is somewhat different from that in English.

4.1.2.3 Evidence from Empty Pronominal

Whitman (1982) observes another fact in Japanese which appears to indicate the absence of the VP node in the language. This fact has to do with the behavior of empty pronominals. Empty pronominals in Japanese show quite different properties than overt pronominals like kare 'he' (cf. (3)) with respect to the relation to their antecedents. Consider the following examples from Whitman (1982:26).

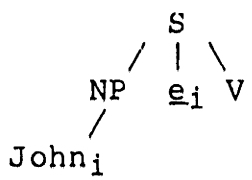
(15) a. *[$_{NP}$ John$_i$-no okaasan-ga e_i aisite-iru
-Gen mother-Nom loves
Lit. 'John$_i$'s mother loves e_i'

b. *e_i [$_{NP}$ John-no okaasan-o aisite-iru
Lit. 'e_i loves John$_i$'s mother'

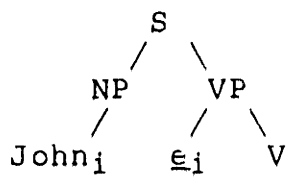
The crucial fact is the ungrammaticality of (15a) under the intended coreference reading. Unlike the overt pronominal kare 'he' in (3c), the empty pronominal in (15a) cannot

take John as its antecedent. Given the condition stated in (2), the ungrammaticality of (15a) would directly follow if we assume that Japanese lacks VP node. That is, if we assume that (15a) has the configuration (16a), rather than (16b) (irrelevant details omitted),

(16) a.



b.



the ungrammaticality of (15a) would follow as a violation of the condition (2), since in (16a) an empty pronominal e_i does c-command its antecedent John. On the other hand, if Japanese has VP node, i.e., the structure of (15a) is the one depicted in (16b), we cannot readily account for the ungrammaticality of (15a), since in (16b), the empty pronominal does not c-command its antecedent John, due to the existence of a VP node. Thus, the ungrammaticality of (15a) with the intended interpretation provides a piece of evidence against the existence of the VP node in Japanese, just as the grammaticality of (3c) provides us with evidence for the existence of such a node, a quite contradictory situation.

4.2 Japanese as a SPECless Language

In this section, I will argue that Japanese lacks "specifiers" in the sense defined in Chapter 2. That is, this language does not have elements that "close off" category projections. Since the existence of specifiers is closely connected with the existence and the nature of Functional categories as discussed in detail in Chapter 2, I will first examine the nature of Functional categories in Japanese, and then go on to discuss the properties of elements in Japanese which have been called "specifiers" in the literature. It will be shown there that none of these elements has the property of closing off the category projection.

4.2.1 Functional Categories in Japanese

4.2.1.1 D

It is a well-established fact that Japanese does not have articles corresponding to the or a in English. thus, noun phrases in this language can freely occur without any

articles with them.

(17) a. John-ga hon-o yonda
-Nom book-Acc read
cf. *John read book

b. John-ga ronbun-o kai-ta
article wrote
cf. *John wrote article

c. inu-ga heya-ni haitte-kita
dog room-to in came
cf. *dog came into room

This fact lends initial support for the claim that Japanese lacks a Functional category D.

Another candidate for a Functional category D is a class of demonstratives such as this and that. Unlike the case of articles, Japanese does have elements which roughly correspond to English this and that, namely, ko-no 'this,' a-no 'that,' and so-no 'that, the.'[10]

(18) a. ko-no hon 'this book'
this book

b. a-no ronbun 'that article'
that article

c. so-no onnanoko 'that/the girl'
that/the girl

The use of these demonstratives is constrained by very interesting functional factors (cf. Sakuma (1936), Hattori (1968), and Kuno (1973), among others). But such considerations are immaterial to our present concern. What is important to our purposes here is whether or not these Japanese demonstratives bear the properties of Functional categories. Recall that I have argued in the preceding chapter that one of the characteristic properties of Functional categories is that only Functional categories project up to XP level, a structurally closed level. Thus, nothing can show up outside the c-command domain of a Functional head, if the Functional head contains no Kase; and if the Functional head is a Kase-assigner, only a single element, the specifier of the Functional head, may appear.

(19) D = the (non-Kase-assigner)

a. the book
b. *John the book

(20) D = 'S (Kase-assigner)

a. John's lecture
b. *yesterday's John's lecture

On the other hand, Lexical heads do not have the property of closing off the projection, and there is no structural limit on iterating modifiers as long as all of the modifiers are appropriately interpreted and licensed.

(21) a. a red car
b. an expensive red car
c. a big expensive red car
etc.

Demonstratives in English clearly pattern with Functional heads in this regard, namely, they have a property of closing off the projection.

(22) a. this book
b. *John's this book

(23) a. that lecture
b. *yesterday's that lecture

The ungrammaticality of (22b) and (23b) is straightforwardly accounted for, if we assume that English demonstratives are just like <u>the</u> in that they are Functional heads without Kase-grid.

Japanese demonstratives, on the other hand, do not have such a property of closing off the category projection, as shown by the following examples.

(24) a. ko-no hon 'this book'
b. John-no ko-no hon Lit. 'John's this book'
c. akai John-no ko-no hon Lit.'red John's this book'
red

(25) a. a-no kuruma
car
'that car'
b. John-no a-no kuruma
Lit. 'John's that car'
c. ookina John-no a-no kuruma
big
Lit. 'big John's that car'

(26) a. so-no koogi
lecture
b. Yamada-sensei-no so-no koogi
teacher
Lit. 'Prof. Yamada's that/the lecture'
c. kyonen-no Yamada-sensei-no so no koogi
last year
Lit. 'last year's Prof. Yamada's that/the lecture'

Some of the examples in (24) - (26) are a little odd due to the semantic conditions imposed on the ordering among

prenominal elements (cf. Chapter 2). But it is unquestionable that all of these examples are significantly better than corresponding English expressions which are clearly ungrammatical, and should be judged to be fully grammatical.

These facts indicate that Japanese demonstratives behave like English prenominal modifiers and never have the property of closing off the category projection as the corresponding English demonstratives do. In the absence of any other plausible candidates for a Functional head D in Japanese,[11] I conclude that this language lacks the Functional category D. Note finally that given the total lack of Functional category D, it immediately follows that Japanese noun phrases are projections of N, namely N', and therefore are never closed. This prediction has already been partially attested by the grammaticality of (24c), (25c), and (26c), which indicates that not only the demonstratives but also the genitive phrases such as John-no 'John's' and Yamada-sensei-no 'Prof. Yamada's' do not close off the projection of N; they are exactly like English prenominal modifiers in this respect. This point will be further strengthened in 4.2.2.1, where I will discuss the 'open' character of Japanese noun phrases.

4.2.1.2 I

Aside from the actual occurrence of modals, the justification of the Functional head I is highly theory internal even in English. The exitence of the Functional category in a language like English is signaled by a number of syntactic phenomena in relation to various principles of grammar. Subject-verb agreement is best described as an instance of "SPEC-head" agreement between a head I (containing AGR) and its specifier. Nominative Case assignment can be treated in a similar way. The so-called "subject-Aux inversion" can be described as a head movement from I to C only if a Functional head I is postulated. The "Nominative island" effect and the distributional property of PRO can both be derived from the Binding Theory if we posit the syntactic category I. The scopal ambiguity of such elements as even and only (Jackendoff (1972)) can receive a natural account if we separate I from the maximal projection of a verb.

It has been pointed out by various linguists that Japanese lacks all of these properties which indicate the existence of a Functional category I. Subject-verb agreement is simply lacking in Japanese, and so is "subject-Aux inversion." Nominative Case assignment (ga marking) takes place quite independent from whether the sentence is

tensed or not (cf. Saito (1982b), Kuroda (1983)).[12] The "Nominative island" effect is also nonexistent in this language, and the possibility of PRO in subject position is not affected by tensedness of that clause (Kuroda (1983)). And the domain of scope-bearing elements in Japanese is always restricted to the verb (Kuno (1980), Whitman (1982)).

In view of this total lack of the cluster of the properties indicating the existence of I in Japanese, a possibility that immediately comes to mind is to claim that there is no such category in the language, i.e., Japanese lacks I, as suggested by Whitman (1982, 1984). Thus, "tense morphemes" such as -ta (Past) and -ru (Present/Non-past) in Japanese do not form a syntactic category I, but are part of a verbal head; and Japanese sentences are basically projections of V, rather than these of I (cf. Whitman (1982)), the choice of the head of S being parametrized (Taraldsen (1983), Chomsky (1986)). This approach may very well be true, or at least the basic insights behind this approach seems to have real content.

Maintaining the core insight of this "$S=V^{max}$" approach toward Japanese, there is another possibility to capture the fact that this language lacks above-mentioned set of properties that signals the existence of I in a language like English. That is, we can assume, as a null hypothesis,

that Japanese has a Functional category I, but this category contains no features, in particular agreement features, at all. Under this approach, the Functional category I exists without any content (grid) and functions just as a place holder for the "tense morphemes" such as -ta (Past) and -ru (Present/Non-past). Since the posited I does not have features, Nominative Case assignment must take place in a way independent of this category, and in fact Japanese Nominative Case assignment (ga marking) takes place in a way quite different from that in English, as we will see later in this chapter. Also, this Functional head I is totally "transparent" with respect to the "Nominative island" effect because it does not contain, above all, agreement features. Verb does not "raise" into the I position since the latter has no "trigger" for the rule of Verb-raising, i.e., agreement feature. Tense particles are attached to the verbal stem perhaps by a PF rule, under the strict string adjacency. Aside from its role as a place holder for the tense particles, the function of this postulated I may be, if any, to "bind" the event position of a verb's θ-grid, if we assume Higginbotham's (1985) story of θ-binding.

Positing this very defective I in Japanese seems to have several descriptive advantages over the "$S=V^{max}$" approach. First, tense morphemes always show up at the

end of a verbal complex in Japanese.[13] For example, tense morphemes can never intervene between a causative verb/morpheme and a passive verb/morpheme, but must appear at the end of the complex verb.

(27) a. tabe - sase - rare - ta
eat Cause passsive past

'was caused to eat'

b. *tabe - sase - ta - rare

c. *tabe - ta - sase - rare

This fact can readily be accounted for if we assume the position outside the projection of V, in which tense morphemes are specified to appear, namely I.

In relation to this distributional requirement of tense morphemes, let us look at the so-called "soo su-" pro form (cf. Nakau (1973), Hinds (1973)). The exact nature of this pro form is not clear at this point and also is not relevant to our present concern. The point is that in a soo su- construction, just like in English do so construction, tense morphemes can freely appear outside the part of the sentence replaced by soo su-. Consider, for example, the following.

(28) John-wa kinoo tosyokan-de benkyoo-si-ta
-Top yesterday library-in study-do-past

'John studied in the library yesterday'

a. Bill-wa ototoi soo si-ta
the day so do
before
yesterday

'Bill did so the day before yesterday'

b. Bill-wa asita soo su-ru-daroo
tomorrow Pres. seems

'Bill seems to do so tomorrow'

c. Bill-mo asita soo si-tagatte-i-ru
also want to

'Bill wants to do so tomorrow, too'

As the examples (28b) and (28c) clearly show, tense morphemes (which are underscored) stand outside the domain replaced by a pro form soo su-. If we assume that tense morphemes occupy the I position which is outside of the projection V, this fact can be straightforwardly accounted for by saying that soo su- replaces the projection of V.

Description of topicalization in Japanese (at least the one which involves movement. See Saito (1985) and Hoji (1985) for much relevant discussions) can also be simplified under the "very defective I" approach. Two distinctions must be made with respect to topicalization in Japanese: First, we must account for the difference

between ga (Nominative) and wa (Topic) in regard to their assignment. And second, we have to explain the difference between scrambling and topicalization, particularly the iterability of the former and the noniterability of the latter. We can give a unified account of these phenomena by positing a defective I. As we will see directly, assignment of ga in Japanese, unlike English Nominative Case assignment, is purely structural. Roughly speaking, once every position of the verb's Case-grid has been discharged, ga marking takes place as a default process, assigning ga to any noun phrase which is a sister of V', hence the possibility of "multiple ga" (See 4.3 for discussion). On the other hand, wa is attached to a phrase which is adjoined to a projection of I, namely I', by topicalization, assuming that topicalization is an adjunction operation in Japanese, along the lines suggested for English topicalization by Baltin (1982) and by Howard Lasnik in his forthcoming work. (See also Whitney (1984), Saito (1985)). As for the difference between scrambling and topicalization, I will argue later in this chapter that scrambling takes place within a projection of V, substituting a constituent into a base-generated 'adjoined' position, which is a fundamentally different operation from topicalization, a true adjunction. Thus, the iterability of scrambling follows from the iterability of

"specifiers" within a Lexical category as I have discussed in Chapter 2, and the non-iterability of topicalization might be accounted for by a general ban on double adjunction in syntax. Also, this analysis of topicalization gives a natural account of the fact that a wh element cannot be topicalized, on the assumption that a wh element cannot be adjoined in syntax (see 4.4).

(29) a. John-ga sono hon-o katta
-Nom that/the book-ACC bought

'John bought the book'

b. *dare-wa sono hon-o katta ka
who -Top Q

'who bought the book'

c. { sono hon / sore / it/that }-wa John-ga t katta

'{ that/the book / that }, John bought'

d. { *dono hon / which } -wa John-ga t katta ka
{ *nani / what }

Lit. ' which book, did John buy'
what

By contrast, scrambling of a wh element corresponding to these examples is possible.

(30) a. John-ga sono hon-o katta

b. sono hon-o John-ga t katta

c. { dono hon / nani } -o John-ga t katta ka

This difference between topicalization and scrambling can, again, be straightforwardly accounted for if we assume that scrambling takes place within a projection of V, substituting elements into a base-generated position, while topicalization is a process of detaching an element from a basic clausal structure, i.e., it is an operation of picking out an element and adjoining it to a projection of I. Thus, assuming the analysis of English topicalization put forth by Baltin and Lasnik, English and Japanese are basically the same with respect to topicalization (involving movement). In fact, as Howard Lasnik pointed out to me, topicalization of a wh element is also prohibited in English, as exemplified by the following contrast.

(31) who thinks that { *who / Bill } John saw t

And I speculate that the lack in English of Japanese-type scrambling, which allows fronting of a wh element, is

probably due to the fact that the subject in English must move to the specifier of IP position to receive Case, and therefore any operation placing some element, say object, in a position preceding the subject must necessarily involve movement of an element to somewhere outside of the projection of a Lexical category, in particular, adjunction to IP.

It is not clear how these differences between topicalization and scrambling in Japanese, as well as the differences between ga and wa, can be given a natural account under the approach which does not posit a syntactic category I.

From these considerations, I will in what follows tentatively assume a very defective I in Japanese, which does not contain any relevant syntactic feature. It should be noted, however, that this does not mean any strong commitment to the existence of a Functional category I in the language. Among the three problems with the configuration lacking the category I pointed out above, the first two could be solved if we come up with some appropriate morphological explanations for the distributional property of tense suffixes in Japanese; for the third problem, no satisfactory account seems to be available at this point, in the absence of sufficient understanding of the nature of topic in Japanese and other languages.

Note also that even under the "very defective I" approach, Japanese sentences are, in a sense, projections of V. That is, aside from topicalization, every grammatical process, including "Nominative Case" assignment (ga marking), takes place within a projection of V in Japanese, which is in sharp contrast to English where Nominative Case is assigned to the specifier of IP position. Notice in this connection that the specifier position of I in Japanese can never be licensed since this very defective I has no F-Feature to discharge.

In this subsection, I have pointed out that there is no indication in Japanese that the Functional category I plays a vital role in this language. In view of the total lack of the set of facts signaling the existence of I, the immediate possibility is to claim that there is simply no such element. Although this is an attractive claim and seems to be correct in its essentials, there are some problems with this approach. Having been unable to solve these problems, I tentatively concluded that Japanese has very defective I which does not have any features (grids) and thus never projects up to the XP level. This view still keeps intact the core insight behind the "$S=V^{max}$" idea, namely, even under this view of Japanese sentential structure, every grammatical process, except, of course, topicalization, takes place within the projection

of V. In the following discussion, I will assume this "very defective I" view for the sake of exposition, although the choice between this view and the "$S=V^{max}$" view does not seem to have crucial bearings on the relevant phenomena to be discussed below.

4.2.1.3 C

Let us now turn our attention to another Functional category, C. I will take up two representative elements which have been assumed so far to be complementizers in Japanese, the so-called question morpheme ka and the subordinate clause marker to.

Consider first ka. It is well-known that in Japanese, any sentence containing a wh element must end in the "particle" ka.[14] This is true irrespective of whether the wh element appears in a matrix sentence or in an embedded one.

(32) a. dare-ga sore-o kaimasi-ta ka
who -Nom it-Acc buy -Past Q

'who bought it'

b. *dare-ga sore-o kaimasi-ta

c. John-wa [dare-ga sore-o katta ka] siranai
-Top bought not know

'John does not know who bought it'

d. *John-wa [dare-ga sore-o katta] siranai

Also, the occurrence of ka is necessary not only for wh questions but for the so-called yes/no questions.

(33) a. John-wa sore-o kaimasi-ta ka

'Did John but it'

b. John-wa sore-o kaimasi-ta

"John bought it'

c. Bill-wa [John-ga sore-o katta ka] siranai
not know

'Bill does not know whether John bought it or not'

d. *Bill-wa [John-ga sore-o katta] siranai

Example (33b) does not have the interrogative meaning in the absence of ka, but rather it is an ordinary declarative sentence. As for (33d), it not only lacks the interrogative meaning, but also is ungrammatical. It is clear from these

examples that ka functions as the "Q-morpheme" (Baker (1970)) having the feature [+Q]. What is relevant to our present concern is, however, the categorial status of this "Q-morpheme": Is ka an instance of C? Or is it something else? I will argue in the following discussion that this element is a noun which bears the feature [+Q].[15].

Notice that the ungrammaticality of (33d) already suggests that the nominal nature of ka, since it is known that the factive verb sir- 'know' requires a noun phrase complement and the only possible reason for the ungrammaticality of (33d), which is minimally different from the grammatical (33c) with respect to the presence/absence of ka, is that the embedded clause lacking ka does not satisfy this requirement by the verb sir- 'know.' In fact, (33d) becomes a grammatical (declarative) sentence if we attach a nominal head koto 'fact,' with the assigned Case particle -o (Acc), to the embedded clause.

(34) Bill-wa [John-ga sore-o katta koto]o siranai
'Bill does not know the fact that John bought it'

The same point is further strengthened by the fact that Case particles such as -ga and -o can be attached to a clause accompanied by ka. As is well known, these Case

particles can only be attached to a noun phrase[16] and can never be attached to other categories, as shown by the following paradigm.

(35) a. [$_{NP}$ John+ga kita
came

'John came'

b. [$_{NP}$[$_{S}$ John-ga Mary-o nagutta] koto+ga akiraka da
hit (fact) obvious is

'It is obvious that John hit Mary'

c. *[$_{S}$ John-ga Mary-o nagutta+ga akiraka da

'It is obvious that John hit Mary'

d. *[$_{PP}$ John-kara+ga tegami-ga kita
from letter came

'From John, a letter came'

(36) a. John-ga [$_{NP}$ Mary+o sitte-iru
knows

'John knows Mary'

b. John-ga [$_{NP}$ [$_{S}$ Bill-ga Mary-o nagutta]

koto+o sitte-iru

'John knows (the fact) that Bill hit Mary'

c. *John-ga [$_{S}$ Bill-ga Mary-o nagutta+o sitte-iru

d. *John-ga [[$_{S}$ Bill-ga Mary-o nagutta] to+o
that

sitte-iru

e. *John-ga Bill-{e / to / kara / from / ni / to / . . .}-o {nagutta / okutta / sent / kaita / wrote / . . .}

In general, any combination x-o, where x is not a noun phrase, is disallowed in Japanese. Consider now the clauses marked by ka. As shown below, both -ga and -o can in fact be attached to clauses accompanied by ka.

(37) a. [[S John-ga nani-o katta] ka ↓ga mondai da
what bought problem is

'The problem is what John bought'

b. [[S John-ga sore-o katta] ka (dooka)↓ga
mondai da

'The problem is whether John bought it'

c. John-wa [[S Bill-ga nani-o kau]ka↓o siritagatte-iru
-Top buy want-to-know

'John wants to know what Bill is going to buy'

d. Boku-wa [[John-ga nani-o katta] ka↓o siritai
I

'I want to know what John bought'

e. Boku-wa [[John-ga sono hon-o katta]
that/the

ka (dooka) to siritai

'I want to know whether John bought that/the book'

The grammaticality of these examples constitutes strong evidence for the nominal nature of ka.

Let us now turn to the categorial status of to 'that.' I would like to argue in what follows that to is a postposition. The fact that to has an independent use as a postposition lends initial support for this hypothesis.

(38) John-wa Mary-to kaimono-ni itta
-with shopping-to went

'John went shopping with Mary'

A stronger piece of evidence that to is a postposition even when it is used as a "clause marker," can be obtained from the attachability of the topic particle -wa. Consider the following examples.

(39) a. [NP John↓wa Bill-o nagutta
hit
'John hit Bill'

b. [PP John-kara↓wa nagaikoto tegami-ga konai
from for a long time letter not come
'It is from John that letters have not come for a long time'

c. [PP Tokyo-e ↓wa Bill-ga itta
to went
'As for Tokyo, Bill went there'

d. *[S John-ga Bill-o nagutta ↓wa mondai da
problem is
'It is a problem that John hi. Bill'

These examples show that the topic marker -wa can be attached to a noun phrase or to a postpositional phrase, but can never be attached to a sentence. However, clauses accompanied by to can freely occur with the topic marker -wa.

(40) a. [[S John-ga Mary-o nagutta] to ↓wa odoroki da
suprising is
'It is surprising that John hit Mary'

b. [[S John-ga sono mondai-o toita] to↓wa totemo
that/the problem solved never
omoenai
not conceivable
'It is inconceivable that John solved that problem'

The examples in (40) clearly indicate that the clauses with to must constitute either a noun phrase or a post-positional phrase, since only to these phrases a topic marker -wa can be attached. The ungrammaticality of the following examples, in which Case particles -ga or -o is attached to a clause with to, shows that clauses with to cannot be noun phrases, because, as we have seen above, -ga and -o can only be attached to noun phrases.

(41) a. *[[S John-ga Mary-o nagutta] to+ga odorokida

'It is surprising that John hit Mary'

b. *[[S John-ga sono mondai-o toita] to+ga

totemo omoenai

'It is inconceivable that John solved that problem'

c. *John-wa [[S Bill-ga Mary-o nagutta] to+o sitteiru

'John knows that Bill hit Mary'
(cf. John-wa [Bill-ga Mary-o nagutta koto+o sitteiru)

d. *John-wa [[S zibun-ga mukasi issyookenmei
self in the past hard

benkyoo sinakatta] to+o kookai site-iru
study did not do regret do

'John regrets that he did not study hard in the past'

(cf. John-wa [zibun-ga mukasi issyookenmei benkyoo sinakatta koto+o kookai site-iru)

From these considerations, we should conclude that clauses with to are postpositional phrases and hence to is a postposition.

Our discussion so far has shown that a "Q-morpheme" ka is a noun and that a "subordinate clause marker" to is a postposition. A stronger argument, namely, an argument that these two elements cannot constitute a single syntactic category, say, C, can be made on the basis of the following fact. That is, it is possible to put these elements together and attach the topic marker -wa to them. For example,

(42) a. [[[[gengogaku-o yaru koto+ni imi-ga aru]
linguistics do in meaning exists

ka] to+wa ii situmon da
good question is

'Whether or not there is meaning in doing linguistis is a good question'

b. [[[John-ga dare-o korosita]ka] to+wa
who killed

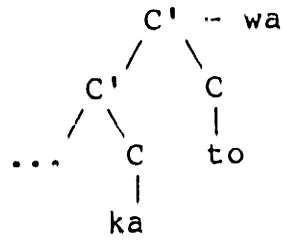
ii pointo da
point

'It is a good point as to who John killed'

If both ka and to belong to a Functional category C, then

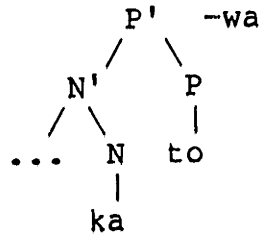
the structure of the phrases marked by -wa in these examples should be as follows.

(43)



The structures like (43), to the best of my knowledge, have not been attested in any language so far, and also do not fit in any version of the X-bar theory, including our system of projection. On the other hand, if our arguments presented above are correct, and the categorial status of these elements ka and to are a noun and a postposition, respectively, then the structure of the phrase in question would be:

(44)



The structure (44), unlike (43), is a quite regular

structure, i.e, one of the typical internal structures of postpositional phrase, and hence no problem arises as to the special treatment of the examples such as (42). Thus, the possibility of the successive occurrences of ka and to, as exemplified by the examples in (42), seems to constitute evidence that these two elements do not form a single Functional category C, but rather, categorially, belong to different syntactic categories, namely, noun and postposition.

In this subsection, I have taken up two elements in Japanese that have widely been assumed to be complementizers in this language, and have argued that there are good reasons to believe that these elements belong to different syntactic categories, viz., ka is a noun and to is a postposition. If this is true, it is very likely, in the absence of plausible candidates, that there is no syntactic category C in Japanese. It should be stressed here that the "function" of the elements such as ka and to is exactly like that of a Functional category C, i.e, ka clearly has an F-Feature (it is a "Q-morpheme"), and it is extremely implausible to attribute any θ-grid to these elements. The fact peculiar to Japanese, which distinguishes this language from, say, English is that these purely functional elements still retain their categorial status as "Lexical categories," and do not form a single Functional category

C. It is not clear at this point why this should be so, but our analysis in this subsection clearly shows that this is actually the case.[17]

4.2.2 "Specifiers" in Japanese

In the preceding subsection, I have argued that Japanese lacks Functional categories D and C. As for I, syntactic evidence for this Functional category in Japanese is very scarce, which suggests the nonexistence of this category in the language. Although this may well turn out to be eventually on the right track, as I pointed out above, I nevertheless tentatively concluded that Japanese should have "very defective I" because of several problematic cases for the approach under which no I is posited. Thus, my conclusion in the preceding discussion is that Japanese lacks D and C, but this language has very defective I which does not have any F-Feature. What, then, is the prediction from the system of projection introduced in Chapter 2 for Japanese with respect to the existence of specifiers in this language?

Recall that in the system of projection I have

proposed in Chapter 2, only Functional categories can have specifiers within these projections. Thus, if there is no Functional category, the absence of specifier is an automatic consequence. Furthermore, even if the head is a Functional head, its specifier position is not licensed and hence nonexistent unless some Kase is discharged to that position. Therefore, given our system of projection, it is predicted that there is no specifier in Japanese. In the cases of D and C, there are simply no such Functional heads in the language. In the case of I, even if there may be such Functional head, as I suggested above, it is very defective in that it does not contain any feature, in particular agreement feature. Hence, the specifier position of the Functional head I in Japanese can never be licensed and thus nonexistent.[18]

I thus conclude that our system of projection predicts that there is no specifier (in the sense defined in Chapter 2) in Japanese. In what follows, I will argue that this prediction is indeed true, by showing that none of the elements which have been treated as specifiers in the literature exhibits a property of "closing off" the category projection, a characteristic property of specifiers.

4.2.2.1 Noun Phrases

We have already seen in Section 4.2.1.1 that genitive phrases, as well as demonstratives, do not close off the projection of N, so that the following Japanese examples are all grammatical in contrast to the corresponding English phrases in the quotes, which are all ungrammatical.

(45) a. Yamada-sensei-no so-no koogi
teacher-Gen that/the lecture
Lit. 'Prof. Yamada's that/the lecture'

b. sensyuu-no Yamada-sensei-no so-no koogi
last week
Lit. 'last week's Prof. Yamada's that/the lecture'

c. Tokyo-daigaku-(de)-no sensyuu-no
university(at)
Yamada-sensei-no so-no koogi
Lit. 'Tokyo University's last week's Prof. Yamada's that/the lecture'

Notice that to say that there is no specifier in noun phrases in Japanese amounts to saying that a projection of N in Japanese is never closed off, i.e., Japanese noun phrases are always "open" in the sense that, given a noun phrase, it is always possible to add something else to it from outside, as long as licensing conditions on the

interpretation of prenominal elements are satisfied. From this angle, we can present further evidence that there is no noun phrase specifier in Japanese, namely, modifiability of pro forms. It is well known that in English, pro forms such as it, he, himself, etc. do not allow further modification, whereas a pro form one does.

(46) a. *big it
b. *short he
c. *yesterday's himself
d. an expensive one

The contrast in (46) can most naturally be accounted for if we assume that pro forms like it, he, etc. are "NP" pro forms, while one is an N' pro form. NP is a closed category, so that it does not allow further modification, whereas N' is at an "open" level and allows iteration of modifiers, as argued in Chapter 2.[19]

The striking difference between English and Japanese in this regard is that Japanese does not have "NP" pro forms, which do not allow further modifications. In other words, Japanese pro forms always allow further modifications as long as the semantic conditions are met. Consider, for example, pro forms sore 'it,' kare 'he,' zibun 'self.' Semantically, pro forms like sore 'it' are most resistant

to further modifications, due to their definiteness. However, given an appropriate context, even these pro forms can be freely modified in Japanese. Thus, the following examples, in which the pro form and its modifier are underlined, are perfectly grammatical.

(47) a. sore 'it'

Tokyo-no biru-no okuzyoo kara mita
-Gen building-Gen top from (I)saw

Haree-suisei-wa smog-no tame bonyarito
Halley's Comet-Top smog-Gen due to faintly

nigotte ita ga, Okinawa-no Naha-de mita
blurrred was but -Gen -in (I)saw

sore-wa yozora-ni kukkirito kagayaite-ita
it-Top night sky-in vividly shining was

Lit. 'Halley's Comet that (I) saw from the top of a building in Tokyo was blurred by the smog, but it that (I) saw in Naha City in Okinawa was vividly shining in the night sky'

b. kare 'he'

kinoo Taroo-ni atta ka-i?
yesterday Taro-with met Q

Lit. 'Did you meet with Taro yesterday?'

un, demo kinoo-no kare-wa sukosi yoosu-ga
Yes, but yesterday-Gen he-Top somewhat state-Nom

hendat-ta
be strange-Past

Lit. 'Yes, but yesterday's he was somewhat strange'

c. zibun 'self'

kukyoo-ni tatasare-ta Saburoo-wa nanno
hardship-in forced to face-Past Saburo-Top not any

kuroo-mo siranakat-ta mukasi-no zibun-ni
sufferings-even not-know-Past old days-Gen self-to

modoritai-to omotta
wanted to go back-that thought

Lit. 'Saburo, who was stranded in hardships, wanted to go back to old day's himself who did not know any sufferings'

The well-formedness of the examples in (47) clearly shows that these pro forms, unlike the corresponding English ones,[20] are N' forms. In fact, there is no "NP" pro forms, i.e., pro forms which do not allow further modifications, in Japanese. This fact strongly supports the claim that Japanese noun phrases are never "closed off,"

and hence that there is no noun phrase specifier in this language, which is a direct consequence of the hypothesis put forth in the preceding subsection that there is no Functional category D in Japanese.

A similar observation can be made for the so-called "stacked" relative clauses. As is well-known, restrictive relatives can stack in English, whereas appositive (non-restrictive) relatives cannot (cf. Chomsky (1977a), Jackendoff (1977)).

(48) a. people who go to MIT who like math will get jobs
b. *John, who goes to MIT, who likes math, will get a job

(Chomsky (1977a:66))

It is not our present concern how we should derive this and other differences between restrictive and appositive relatives (see Jackendoff (1977) for relevant discussions). What is important for our present purposes is that Japanese lacks such a contrast between restrictive and appositive relatives with respect to their stackability. Both of them can stack. Thus, both of the following examples are acceptable.

(49) a. restrictive

[NP [S Osaka-(de)-no kokusai-kaigi-ni
in -Gen international conference at

sanka-suru koto-ni-natte-iru][S America-kara
attend is supposed to from

kaette-kita bakari-no gakusya-tati-wa ima
came back just scholar-plural-Top now

Tokyo-no hotel-ni tomatte-imasu]
at are staying

Lit. 'The scholars who are supposed to attend the international conference in Osaka who just returned from America are now staying at a hotel in Tokyo'

b. appositive

[NP [S Osaka-(de)-no kokusai-kaigi-ni sanka-suru

koto-ni-natte-iru][S America-kara kaette-ki-ta

bakari-no John-wa ima Tokyo-no hotel-ni

tomatte-imasu

Lit. 'John, who is supposed to attend the international conference in Osaka, who just returned from America, is now staying at a hotel in Tokyo'

The lack of asymmetry between Japanese restrictive and appositive relative clauses with respect to their stackability exemplified by the above examples shows that these two types of relatives are syntactically indistinguishable[21] in Japanese; neither of them close off the projection of

N. That is, both of the two types of relatives pattern with prenominal modifiers in, say, English, and neither of them has a property of specifiers.

We have seen in the previous discussion that what have been regarded as specifiers in Japanese, in particular, genitive phrases, as well as demonstratives, do not have the characteristic closing property of specifiers. We have also argued that there are independent evidence that Japanese noun phrases are never closed (never reach the "closing" level, XP), by showing that even what have been treated as "NP" pro forms in this language can freely be modified, and that not only restrictive relatives but also appositive ones can stack in Japanese, which is, again, in sharp contrast with English. It should be now clear that there are no specifiers that close off the projection of N in this language, as predicted by our system of projection.

4.2.2.2 Sentences

In the general framework of grammar I am assuming in this thesis, there are two kinds of specifiers in the clausal case, one is the specifier of C (or CP), and the other, the specifier of I (or IP). The former type of the

specifier is instantiated by a moved wh element (cf. Chomsky (1986))[22] and the latter type by the subject. If my arguments presented in Section 4.2.1 above are valid, the specifier of C does not exist in Japanese simply because there is no Functional category C in the language. And in fact wh elements do not move (in syntax) in Japanese, as is well known. As for the specifier of I, our system of projection predicts, as I argued before, that there is no specifier of I in Japanese, because, first of all, the existence of the Functional category I in Japanese itself is questionable, and, even if there is such a Functional category, it does not have agreement feature to discharge. I thus suggested there that subject in Japanese is within a projection of V. Note in passing that this is also true for English in our system. That is, not only Japanese but also English has a subject within a projection of V at D-structure. English differs from Japanese in that the Functional category I in English has agreement features (in the tensed case) so that subject must move to the position of specifier of I to discharge I's agreement features, or put it differently, subject must move, if it is lexical, to receive Case from I in order to avoid a Case Filter violation. By contrast, Japanese has a way of Nominative Case assignment, which is quite independent from the Functional head I, as we will see later in 4.4. Also, due

to the lack of agreement features in I, there is no position to which subject can be moved to get Case. Therefore, there is no necessity nor possibility for subject to move into the specifier position of I in Japanese. In short, what distinguishes between English and Japanese in this respect is the fact that English has I with agreement features but it lacks the structural Nominative Case assignment mechanism, while Japanese lacks I with agreement features but it has the structural Nominative Case assignment (ga marking). It is of course desirable if we can derive one from the other, i.e., if we can derive the existence of structural Nominative Case assignment from the lack of agreement features in I, or vice versa. At this point it is not clear to which direction the derivation goes, although I speculate, in view of the significant role the agreement features play in various other places in a grammar, that the presence/absence of agreement features is the fundamental parametric property of a language from which other properties, e.g., the existence of purely structural Nominative Case assignment mechanism, must follow.

Returning to the discussion of the position of subject in Japanese, if, as I have argued, subject in this language stays within a projection of V and hence is not the specifier of I, no principle of grammar requires

its uniqueness,[23] namely it can be iterated like other adjuncts/modifiers within a projection V, as long as all occurrences of subject are appropriately interpreted. This is indeed the case as shown by the well known "multiple subject" construction exemplified below.

(50) a. heikin-zyumyoo-ga mizikai
average-lifespan-Nom is short

'The average lifespan is short'

b. dansei-ga heikin-zyumyoo-ga mizikai
male

'It is men that their average lifespan is short'

c. bunmeikoku-ga dansei-ga heikin-zyumyoo-ga
civilized countries

mizikai

'It is civilized countries that men, their average lifespan is short in.'

(cf. Kuno (1973:Ch.3))

Thus, the existence of the so-called "multiple subject" construction in Japanese supports our claim that subject in this language is not the specifier (of I), but is within a projection of V.

The facts about scrambling provide further evidence

for our claim that there is no specifier in a clausal system in Japanese. As is well-known, "multiple" scrambling is freely allowed within a single sentence in Japanese.

(51) a. John-ga Mary-ni so-no-hon-o watasita
-Nom -to that/the book-Acc handed/gave

'John gave the book to Mary'

b. Mary-ni John-ga so-no-hon-o watasita

c. so-no-hon-o Mary-ni John-ga watasita

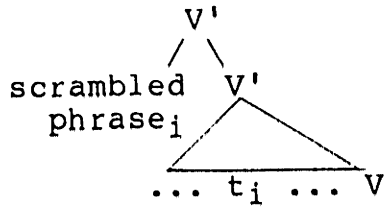
The "multiple" scrambling such as the one exemplified by (51) should not be allowed if scrambling is a movement into a specifier position, because in our system of projection, specifier, if any, must be unique and cannot be iterated (cf. Chapter 2). The "multiple" scrambling will also be prohibited under the assumption that it is an adjunction operation (cf. Saito (1985)), if we assume, following Gueron and May (1984), that only a single element can be adjoined to each category, i.e., that "multiple" adjunction is generally banned.[24] On the other hand, if we assume that scrambling takes place within a projection of V and that it is a movement operation distinct from adjunction, then the perfect acceptability of "multiple" scrambling sentences is successfully accounted

for. What, then, is the status of scrambling operation? Under the general theoretical framework I am assuming here, all movement operations are either adjunction or substitution. I have just argued that characterizing scrambling as an adjunction creates undesirable results. Therefore, it must be a substitution operation. In our system of projection, however, it is not clear how the "landing site" for scrambling could be "licensed" at D-structure. If, then, there is no base-generated landing site (empty node) available to scrambling, then it is impossible to characterize this rule as a "substitution" in the standard sense. Hence, I would propose the following definition of "adjunction" to reconcile this apparent dilemma.

(52) A movement is an adjunction iff the structure created by that movement is non-base-generable (otherwise, the movement is a substitution).[25]

Given the definition (52), scrambling is an instance of substitution, as desired, since the resulting structure after the application of scrambling is obviously "base-generable."

(53)



Recall that in our system of projection, there is no structural limit on the "recursion" of Lexical categories. This amounts to saying that, given the definition (52) of adjunction, there is no "adjunction" to a projection of a Lexical category, due to the possibility of "free recursion" within a projection of a Lexical category.[26] Thus, for example, scrambling can be regarded as an operation that takes place "inside" the projection of V, due to the "open" nature of the Lexical category. In contrast, "adjunction" to a projection of a Functional category is always an adjunction according to the definition (52), since Functional categories do not allow "recursion" and hence the resulting structure after the application of "adjunction" is always non-base-generable.

I have argued in this subsection that Japanese clauses, as well as noun phrases (see 4.2.2.1), do not have elements that close off their projections and thus are always "open" as exemplified by the possibility of "multiple subject" constsruction and that of "multiple" scrambling.

4.2.3 Summary

In this section, I have examined a potential class of Functional categories in Japanese and have concluded that there are no Functional categories D and C in this language. As for the Functional category I, I have observed that syntactic evidence for the existence of this Functional category in Japanese is very scarce. This suggests that Japanese sentences should be analyzed as a projection of V, rather than that of I. We have seen that this is basically correct, but not entirely, due to some problems with this "$S=V^{max}$" approach. I thus posited a "very defective I" in this language in order to handle such problematic cases, noting that even under this "very defective I" approach, Japanese sentences are, in a sense, projections of V in that every major grammatical process, including Nominative Case assignment (ga marking) takes place within the projection of V.

Given the total lack of Functional categories with F-Features in Japanese, our system of projection proposed in Chapter 2 predicts that there is no specifier (in the sense defined there) in the language. In the subsection 4.2.2, I have examined the elements in Japanese that have hitherto been assumed to be specifiers, and have argued that none of those elements has the characteristic property

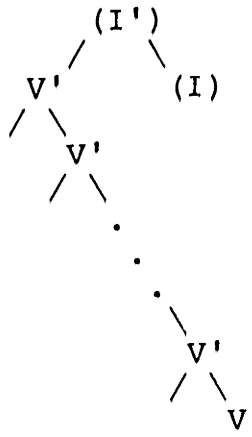
of specifiers. I have also suggested, in this connection, that scrambling in Japanese should be analyzed as a substitution operation, rather than as an adjunction operation. My conclusion was, then, that there is no specifier in Japanese, exactly as predicted by our system of projection introduced in the previous chapter.

4.3 Phrase Structure of Japanese

4.3.1 Phrase Structure of Japanese : A Proposal

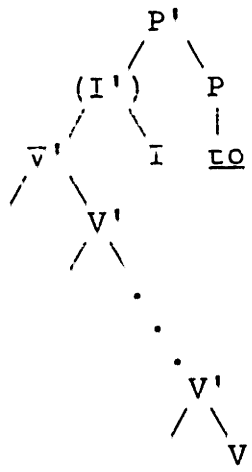
Our conclusions in the preceding section naturally lead us to the following phrase structural configurations for Japanese, taking sentences and noun phrases as representative examples.

(54) basic clausal structure

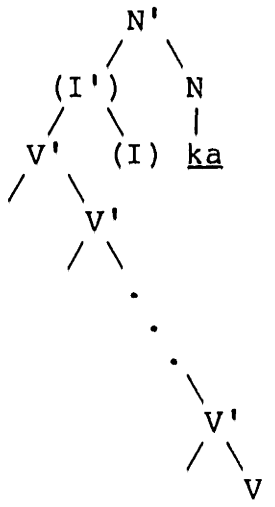


(55) clauses with "COMP"

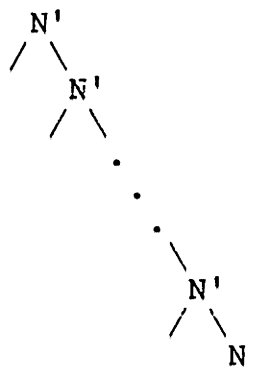
a. declaratives



b. interrogatives



(56) noun phrases



Notice that aside from the categorial retention of Japanese "complementizers," these configurations are almost identical to the structures for Lexical categories in English (except, of course, different choices of the position of head). Thus, the overwhelming superficial differences

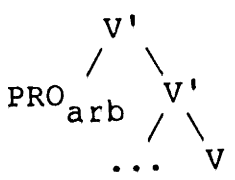
between English and Japanese can basically be reduced to the fact that English has a rich set of Functional categories with agreement features, whereas Japanese lacks such syntactic categories; Japanese either totally lacks Functional categories (if the existence of "very defective I" can somehow be eliminated), or, even if it has one of them, namely, I, this category does not have any agreement features, unlike the corresponding Functional category in English. Various typological differences between English and Japanese might naturally follow from this minimal difference between the two languages, namely, English has agreement phenomenon but Japanese does not, which is very closely connected with the existence and/or the nature of Functional categories in the two languages. I have already discussed some of the typological features distinguishing English and Japanese, for example, the "multiple subject" construction, Nominative Case assignment, scrambling, etc. Some other consequences of our view for the typological differences will be discussed, with further clarifications of those that I have already mentioned, in Section 4.4.

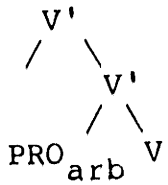
Before going into such discussions, however, let us briefly see in this section if our phrase structural configurations are compatible with the set of data summarized in 4.1 before.

4.3.2 Facts Recaitulated

4.3.2.1 Evidence for the "VP" node

Let us first consider the set of evidence for a "VP" node in Japanese summarized in 4.1.1. Notice crucially that all of the facts discussed there have to do with the Binding Theory, and it is known that the Binding Theory does not refer to a particular node label (e.g., VP), but rather, what is crucial for the Binding Theory is the structural relation 'c-command" (cf. Chomsky (1985, 1986), among others). In other words, the facts discussed in 4.1.1 show that there is a node that dominates the object but not the subject, but they do not say anything about the actual node label of the node in question. It may or may not be VP. The configurations I proposed above do satisfy this factual requirement. That is, there is a node that dominates the object but not the subject, namely, V', in the proposed configuration. As for the distributional property of PRO_{arb} in Japanese, exactly the same explanation as the one for the distrubutional property of PRO in English is possible. PRO_{arb} in the subject position and PRO_{arb} in the object position appear in the schematic representations (57a) and (57b), respectively.

(57) a. 

b. 

PRO_{arb} in the object position, namely, the one in (57b), is clearly (lexically) governed by the verb, and thus will be excluded as a violation of the "PRO theorem." PRO_{arb} in the subject position appears in the configuration (57a). And it has already been argued in Chapter 3 that this position is not governed by the verb (see 3.2). Thus, the distribution of PRO_{arb} can be straightforwardly accounted for by our proposed configurational structures for Japanese without any problem.[27]

I have thus argued that all the facts summarized in 4.1.1 are compatible with the phrase structural configuration proposed above, since what is indicated by the set of facts is the existence of a node dominating object but not subject, and our configuration does indeed posit such a node, namely, V'. Let us now turn our attention to the facts summarized in 4.1.2, evidence against the "VP" node.

4.3.2.2 VP Movement

As we discussed in 4.1.2.1, Japanese lacks VP movement in contrast to English. This difference between the two languages is illustrated by the contrast between (11) and (12), repeated here for convenience as (58) and (59), respectively.

(58) Heather promised to come at 10, and come at 10 she did.

(59) *Susan-wa zyuuzi-ni kuru to yakusokusita,
-Top ten -at come that promised
o'clock

sosite (zizitu) zyuuzi-ni ki/ku kanozyo-wa ta
and in fact come she Past
(stem)

Lit. 'Susan promised to come at 10, and (in fact) come at 10 she Past'

This fact has been taken as evidence against "VP" node in Japanese. However, it seems to me that the basic difference between English and Japanese in this regard is not the categorial status of the element to be moved, but different status of the inflectional element in the two languages. In fact, Saito (1985:235-244) proposes a plausible account of the lack of VP movement in Japanese in these terms. He cites Kuno's (1978b) observation that Japanese does not

have an auxiliary verb that can be used independently such as English do and proposes that the lack of VP movement in Japanese can be accounted for on the basis of this fact and the existence of the general condition such as follows.

(60) INFL must be realized.

(Saito (1985:238))

That is, in English, the auxiliary verb do is inserted into the position of I after the application of VP movement, realizing the inflectional feature and thus satisfying the condition (60). Japanese, on the other hand, lacks such elements as do in English. Therefore, if VP is moved to the front, the inflectional element cannot be realized in violation of (60), since Japanese does not have do and affix hopping, a way of realizing inflectional features, requires adjacency, but in a structure after VP movement, verb and the inflectional elements are not adjacent to each other. Thus, the presence of VP movement in English and the lack of such movement in Japanese is not due to the difference in categorial status of the category in question, but due to the fact that English has elements like do as INFL realizers, whereas Japanese does not.

This analysis seems to me to be on the right track.

Furthermore, in our system, the difference between English and Japanese observed by Kuno and Saito can be reduced to a more basic difference between the two languages. Namely, English I has inherent agreement features, so that it is possible that some element (e.g., do) other than a verb bears these features to realize them. By contrast, Japanese I does not have any inherent feature and functions only as a "place holder" for the tense morphemes. Thus, there is no possibility that the "inflectional features" are realized, simply because there are no such features in I, hence the lack of do in Japanese. However, the tense morphemes in Japanse must be "suffixed" to the verbal stem by some PF rule. And this PF rule requires a strict adjancency between the verbal stem and the tense morphemes. Therefore, "VP movement" in Japanese always results in an ill-formed sequence, rendering this PF rule inapplicable.

If the above account of the lack of "VP movement" rule in Japanese is right, our phrase structure of Japanese is quite consistent with the facts summarized in 4.1.2.1.

4.3.2.3 The Distribution of Adverbial Elements

The fact discussed in 4.1.2.2 concerning the distri-

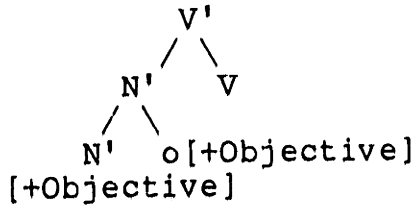
bution of adverbial elements in Japanese, in particular, the possibility of the occurrence of an adverbial element in between a verb and its direct object, can be restated in current terms as the lack of "adjacency requirement" (Stowell 1981)) on Case-assignment in Japanese. The explanation for this phenomenon, however, can be given quite independently from the issue of the existence of the "VP node" in Japanese. The essential content of Stowell's idea on Case-assignment can be stated as follows (cf. Stowell (1981:113), irrelevant part omitted).

(61) In the configuration [α β ...] or [... β α],
α Case-marks β, where
(i) α governs β, and
(ii) α is adjacent to β

As can be seen from the formulation (61), the "adjacency condition" is regarded as a part of Case-assignment mechanism itself. However, this is not the only conclusion we can draw from the impossibility of the occurrence of an adverbial element between a verb and its direct object in English. Suppose that the "adjacency condition" (61ii) is detached from the Case-assignment itself, so that Case-assignment takes place under the requirement of government between a Case-assigner and the Case-assignee. We assume

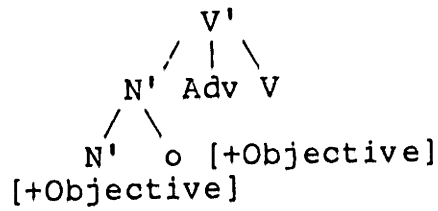
that the Case-assignment takes place at S-structure. Under this assumption, Case is assigned under government uniformly across languages. Suppose further that there is an additional mechanism in UG that checks if the Case which is assigned abstractly to a noun phrase[28] through the Case-assignment process is appropriately "realized." Let us call this mechanism a "Casechecking" mechanism and assume that it applies in PF. The Case-checking mechanism works in different ways in different languages. In languages like Japanese, where there are overt Case-markers, the process of Case-checking takes place <u>within</u> a noun phrase. Consider the following hypothetical configuration in Japanese where a Case-marker <u>-o</u> (an accusative/objective particle) is assumed to be adjoined as a Case-realizer to a noun phrase which has already been assigned an abstract Case [+Objective] by a verb at S-structure under government.

(62)



The Case-checking mechanism applies, checking if the feature assigned to a noun phrase and the feature of the Case-particle match. In (62), these features match, both of them are [+Objective], and the structure will be marked well-formed with respect to the Case theory. Note that this process in (62) takes place within a noun phrase in question. Thus, even if an adverbial element intervenes between a noun phrase and a verb such as follows,

(63)

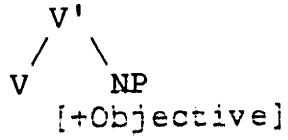


the structure will also be well-formed.

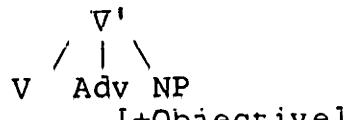
On the other hand, in languages like English where there is no overt Case-markers comparable to the ones in Japanese, the Case-checking mechanism would have to look at the Case-assigner to see if an appropriate Case is assigned to a noun phrase in question, because there is no information available within the noun phrase itself as to the appropriateness of the feature, e.g. [+Objective], assigned to that noun phrase. And if we assume that the Case-checking mechanism can only look at adjacent elements,

as seems plausible in view of the fact that it takes place in PF, then we can account for the impossibility of the intervening adverbials in a language which lacks overt Case-markers. In the following hypothetical configurations in English, for example, (64a) is well-formed but (64b) is ruled out, since in the latter case the Case-checking mechanism cannot work properly; it can only look at two adjacent elements, but the verb which is a Case-assigner, and the noun phrase, a Case-assignee, are not adjacent to each other.

(64) a.



b.



In this way, the difference between English and Japanese, i.e., the existence of the requirement that a verb and its direct object be adjacent to each other in the former and the lack of such requirement in the latter, can be accounted for, quite independent of the "VP" issue, by postulating the Case-checking mechanism in PF, whose character is

strictly local in the sense that it can only look at two adjacent elements. Note incidentally that by detaching this mechanism from the Case-assignment process itself and placing the former in PF, it becomes possible to keep the main body of the abstract Case-assignment system uniform across languages, i.e., it takes place under government crosslinguistically, attributing the apparent difference between, say, Japanese-type and English-type languages, to the fact that the former type of languages have overt Case-markers attached to noun phrases, whereas the latter type of languages do not.

4.3.2.4 Empty Pronominals

Let us finally discuss the problem posed by the ungrammaticality of (15a), reproduced here as (65).

(65) *[$_{NP}$ John$_i$-no okaasan ɨga e$_i$ aisite-iru
-Gen mother -Nom loves

As discussed in 4.1.2.3, the explanation for the ungrammaticality of (65) would not be straightforward if we assume

an "intervening" node which dominates the object of a verb but not the subject, whatever the label of that node might be. (Recall that the relevant notion for the Binding Theory is "c-command," not "m-command," cf. Chomsky (1985)). If there is such an "intervening" node between subject and object, there will be no c-commanding relationship between $\underline{\text{John}}_i$ and $\underline{e}_i$ in (65). Therefore, principles such as (2) cannot be invoked to account for the impossibility of the coreference reading.

However, the ungrammatical status of the examples like (65) under the coreference reading is not as entirely clear as the ones with overt pronominals. From this we might claim that the ill-formedness of (65) is due to some extragrammatical factors and therefore should not be ruled out in terms of grammatical principles such as the Binding Theory. This position is in fact taken by Hoji (1985, Appendix A). In support of his claim that the alleged impossibility of the coreferent interpretation in sentences like (65) should not be caused by syntactic conditions such as (2), Hoji points out that the change in pragmatic control or slight change of the relevant structure (e.g., the use of a different verb/noun, the addition of the intensifiers such as -<u>sae</u> 'even,' etc) makes the sentence significantly better (Hoji (1985:382)). The following examples illustrates this point.

(66) [NP John$_i$-no teki-sae ‡ga $\underline{e}_i$ aisite-iru
-Gen enemy even loves
'Even John$_i$'s enemies love him$_i$'

(66) is certainly much better than (65) under the coreference reading. However, if the impossibility of the intended coreference reading in (65) (even for (65), Hoji reports, some speakers, including himself, do get the coreference reading) is to be handled by some syntactic condition such as (2), this should not be the case, because a violation of syntactic condition, the Binding Theory, cannot be readily circumvented by pragmatic control or slight change of the relevant structure. For example, the grammatical status of the sentences in (67) under the intended coreference reading, which violates the condition (2), a part of the clause (C) of the Binding Theory, cannot be improved by the addition of -$\underline{sae}$ 'even' as shown by the total ungrammaticality of the sentences in (68) (from Hoji (1985:382) with slight adaptations).

(67) a. *kare$_i$-ga [S John$_i$-ga Mary-o butta‡to omotta
he -Nom -Nom -Acc hit that thought
Lit. 'He$_i$ thought that John$_i$ hit Mary'

b. *e_i [S John$_i$-ga Mary-o butta‡to omotta
Lit. '$\underline{e}_i$ (=he$_i$) thought that John$_i$ hit Mary'

(68) a. *kare$_i$-ga [$_S$ John$_i$-sae-ga Mary-o butta↑to omotta
even
Lit. 'He$_i$ thought that even Johni hit Mary'

b. *e$_i$ [$_S$ John$_i$-sae-ga Mary-o butta↑to omotta
Lit. '$\underline{e}_i$(=he$_i$) thought even John$_i$ hit Mary'

From these considerations, Hoji (1985) concludes that the impossibility of the coreference reading in sentences like (65) is not due to syntactic constraints such as (2), but rather, it should be handled by some non-syntactic (pragmatic) constraint. Thus, according to him, syntactic constraints should rule (65) in, with the explanation for the apparent impossibility (or "difficulty," since there are some speakers who allow the coreferent reading) of the coreferent reading attributed to pragmatic constraints.

This conclusion seems reasonable in view of the sharp contrast between (65) and (66) on the one hand, and (67) and (68) on the other. And if Hoji's (1985) approach is right, then (65) is no longer problematic for the phrase structural configuration proposed above, since what is required to make the coreferent reading possible is the node that dominates $\underline{e}$ but not the subject noun phrase. In our analysis, there is indeed such a node, namely V', as required.

4.4 Some Consequences

This section briefly discusses several consequences of the phrase structural configuration for Japanese proposed in this chapter.

An initial consequence is that we can now explain why Japanese does not have syntactic wh movement. Recall that in the theory of grammar we are assuming throughout this study, wh element is assumed to be moved into the specifier (of CP) position (cf. Chomsky (1985)). This mode of movement is impossible in Japanese, since, as I have argued above, this language does not have specifiers. Therefore, the only possible way of moving a wh element in syntax is to adjoin it to some category. However, an adjunction of a wh element is generally prohibited as indicated by the impossibility of topicalizing a wh element (cf. (31)). Notice that "VP-adjunction" proposed by Chomsky (1985) is not an instance of adjunction in our system. As we saw above, given the definition of adjunction (52), an "adjunction" to VP, or to a projection of a Lexical category in general, is always a substitution operation. Thus, "VP-adjunction" of a wh element in English is allowed, if not required, just like the scrambling of a wh element is possible in Japanese.

If these considerations are right, the reason for the

lack of syntactic wh movement in Japanese is obvious: wh elements in Japanese cannot move (aside from scrambling) in syntax due to the lack of possible landing site.

Needless to say, the ban on the adjunction of wh elements does not apply in LF, even in Japanese, wh elements must move in LF to get a scope over a proposition. This scope assignment presumably takes place under adjoining a wh element to I'.

Another consequence of the proposed phrase structure for Japanese is that it now becomes possible to unify assignments of ga 'Nom' and no 'Gen.' It has been noted (Saito (1982a), Kuroda (1983)) that ga marking in Japanese is independent of government (and θ-marking), and that it takes place in a purely structural manner. However, to the best of my knowledge, the fundamental similarity between ga marking and no marking has never hitherto been stated explicitly, although it seems clear that assignment of no is also independent of government (and θ-marking). Given the phrase structure for Japanese proposed in this chapter, ga marking and no marking can be collapsed into the following very simple schema.

(69) In the environment $\{ \begin{smallmatrix} N' \\ P' \end{smallmatrix} \}$ ___ X',

(i) insert ga if X = V
(ii) insert no if X = N

We thus capture the fundamental similarity between ga marking and no marking, unifying them as two instances of basically the same process formulated as an insertion rule (69).

Various other consequences might follow from the Japanese phrase structure proposed in this chapter, including the significant simplication of Case-marking, θ-marking, etc. These consequences should be further explored in future research.

Notes to Chapter 4

1. In what follows, I will describe the weak crossover facts in a somewhat "standard" way, namely, the way in which the A/A' distinction is used. This is just for the sake of exposition. As we discussed in Chapter 2, in the system of projection proposed there, a different explanation is needed. See Chapter 2 for some possible accounts of the weak crossover phenomena in our system of projection. Also, I will assign rather traditional structures to the examples, especially when I cite from someone else's work, as long as the argument is not directly affected by postulating such structures. Recapitulation of the facts presented in this section under our conception of phrase structure in general and of Japanese phrase structure in particular, will be done in Section 3 below.

2. I will not discuss the strong crossover cases in the following. See, among others, Postal (1971), Chomsky (1976), Reinhart (1979), Higginbotham (1980), Saito (1982b), and Lasnik (1985) for relevant discussion on strong crossover. Also see Chapter 2.

3. For example, (i) does not exhibit the weak crossover effect, since in its LF representation (ii), the trace left by QR c-commands the pronoun.

(i) everyone$_i$ loves his$_i$ mother
(ii) [$_S$ everyone$_i$ [$_S$ t_i [$_{VP}$ loves his$_i$ mother]]]

4. For the argument in terms of the weak crossover effect created by the scrambling of zibun, see Saito and Hoji (1983). See also Farmer and Tsujimura (1984) for the criticism of Saito and Hoji's argument.

5. See Huang (1982) and Lasnik and Saito (1984) for much relevant discussion on LF wh movement.

6. Kuroda takes up PRO_{arb}, rather than 'controlled PRO,' because there are complicated factors in the latter case that obscure the line of arguments. He writes (Kuroda (1982:162)):

> Data concerning controlled PRO is more analysis-dependent than that concerning PRO_{arb} and cannot serve as evidence without argument as directly as data with controlled PRO (sic. should be "PRO_{arb}"; N.F.), but one can expect that it does not provide any counterevidence.

I agree with him in that data concerning controlled PRO is more analysis dependent especially in view of the rather free occurrence of pro in Japanese, and will concentrate on the distribution of PRO_{arb} in the following discussion.

7. Kuroda (1983) also presents additional evidence for the "VP" node in Japanese in terms of "Quantifier Float" phenomenon in this language (cf. Okutsu (1969), Kamio (1977b), Shibatani (1977), Inoue (1978), and Kuno (1978a) for the relevant discussion of this phenomenon). It seems to me that the argument presented by Kuroda (1983), as well as the one by Haig (1980), is essentially for the existence of the "basic word order" in Japanese, or in other words, for the existence of "scrambling rule" in the language. Although the issue of scrambling rule and the existence of the VP node are closely connected (See Farmer (1980), Hale (1980, 1982, 1983), and Saito (1985), among others), they are nevertheless logically independent.

8. See, for instance, Koopman (1984), Travis (1984), Baker (1985), and Chomsky (1985), for the discussion of the latter ("head movement") case.

9. See Stowell (1981) for an analysis of this phenomenon in terms of his 'adjacency condition on Case assignment.' See also the discussion in Section 3.

10. I assume that the second part of these demonstratives, i.e., <u>-no</u>, is an instance of structurally assigned genitive Case, since <u>ko</u>-, <u>a</u>-, and <u>so</u>- have different forms when they appear in environments other than prenominal position, e.g., <u>ko-re</u>, 'this,' <u>a-re</u> 'that,' and <u>so-re</u> 'it,' etc. See Section 4 for

discussion of Case-assignment in Japanese.

11. It might be possible to consider Japanese Case particles as Functional heads comparable to D in English, thus forming a Functional projection KP ("Kase Phrase," following Ken Hale's terminology). I will not pursue this possiblity here, although the "KP" idea seems to provide a refreshing crosslinguistic perspective and is definitely worth pursuing in the future research. See Lamontagne and Travis (1986) for a similar approach.

12. For our treatment of ga marking in Japanese, see 4.4.

13. Except for some modal-like elements and various sentence-final particles. The former set of elements can perhaps be treated as sort of verbs that take clausal complement. Some of these elements inflect and other do not, and if they take tense morphemes, these morphemes again appear at the end of the verb. As for the latter set of elements, it is not clear how to treat them, although some of them might be reanalyzed as a lexical element taking a clausal complement (see below). These class of elements seem to be, by and large, outside the scope of the X-bar theory. See Inoue (1976b), Kamio (1981) for relevant discussion.

14. I put aside here the possibility of no, which can best be described, I believe, as a structurally inserted "genitive Case" before a nominal ka, and then bears the function of optionally replacing ka in some environments especially in the matrix sentence, a similar process to the case of the so-called "pronominal" no (e.g., John-no 'John's' < John-no x (=nominal), with x being deleted.)

15. A similar characterization of this element has been suggested by Saito (1985:273).

16. There are some exceptions with -ga, namely, -ga can sometimes be attached to a postpositional phrase, a property similar to that of -wa, as we will discuss shortly.

Tokyo-kara - ga New York-ni iki-yasui
from wa -to go -easy.
'It is from Tokyo that one can go to New York easily'

This does not affect our argument, however, since our present purpose is to deny the status of ka as a complementizer. Also, there is no comparable exception to -o. It is always attached to a noun phrase, and , crucially, it can be attached to ka-marked clause as well. Incidentally, this difference between -ga (and -wa) and -o probably stems from the fact that these particles are assigned in different ways, i.e, -ga, as well as -wa, is assigned in a purely structural way, whereas -o is assigned under government by the

verb. See 4.4 for discussion on this matter.

17. The only problem with the "ka = noun" analysis proposed in this subsection is that root interrogative sentences such as (32a) and (33a) should also be noun phrases under this analysis. Semantically, however, the problem is not as serious as it seems, since ka does not have "meaning" (referentiality, θ-grid) and the whole expression (i.e., a clause with ka,) is still propositional, rather than referential, even though its categorial status is a projection of N.

18. The "exceptional" Kase discharge from outside of the projection of I to its specifier position is impossible in Japanese, since a projection of I is always (except when it occurs as an independent sentence) taken as as complement by a postposition (e.g. to) or a noun (e.g. ka) which can arguably be assumed to lack Kase-grid. Unlike English, a projection of I cannot be directly taken by V as its complement.

19. Of course, "NP" should be a projection of D, namely, D' or DP, in our system of projection.

20. Even in English, there are a small member of marked cases where a pronoun is modified, e.g., the real you, my former self, he who casts the first stone, etc. The existence of such marginal N' pro forms in

English does not affect my argument, however. The crucial fact for my argument here is that there are no non-modifiable pro forms in Japanese.

21. See Inoue (1976b) and Chomsky (1977a) for relevant discussion. Cf. also Kamio (1977a) for some counter-arguments. Our analysis of Japanese phrase structure suggests the direction in which the semantic difference between "restrictive" and "appositive" relatives in this language is captured in terms of interpretive devices, just like other prenominal elements. See Inoue (1976b), Whitman (1981) for relevant discussion.

22. Topic could be another candidate that fills in the specifier position of C, if we extend the analysis of topicalization put forth in Chomsky (1977b) and reformulate it as a movement into the specifier of C.

23. Of course the θ-criterion, or the principle governing predication (Williams (1980)), requires the uniqueness of the external argument. What is under discussion here is not "subject," in the sense of external argument, but is "subject" meaning, roughly, "an element which is marked by Nominative Case." The uniqueness of the external argument, as we expect, holds even for the "multiple subject" construction such as (50) below. Only one of the ga-marked phrases, usually the lowest one, again as expected, is θ-marked by the predicate; other ga-marked phrases are licensed by being interpreted as, say, Focus (or

Topic) under the "aboutness" relation (cf. Chomsky (1982)).

24. This condition is called "adjunction constraint" by May (1985). See Guéron and May (1984) and May (1985) for details.

25. Here, I depart from May's (1985) position that there is no difference between adjunction and substitution with respect to their "structure-preservingness." Note that this departure is independent of the issue of how we define "dominance" in the case of adjunction. My claim in (52) is that structure-preserving property of a movement should be defined in terms of "node," rather than "category" in May's (1985) sense, which does not directly deny the general validity of of the distinction between nodes and categories proposed by May (1985).

26. Note incidentally that the adjunction of PRO in the passitve case that I suggested in Chapter 3 should now be regarded as an instance of substitution.

27. Lexical government of I from outside the projection of V is impossible, since we are assuming that Japanese does not have a process of V-raising into I. See above.

28. We restrict our attention to noun phrases, here.

REFERENCES

Abney, S. (1985) "Functor Theory and Licensing : Toward an Elimination of the Base Component," unpublished manuscript, MIT, Cambridge, Massachusetts.

Anderson, M. (1984) "Prenominal Genitive NPs," The Linguistic Review 3:1.

Aoun, J. (1982) "The Formal Nature of Anaphoric Relations," Doctoral dissertation, MIT, Cambridge, Massachusetts.

Aoun, J. and D. Sportiche (1984) "On the Formal Theory of Government," The Linguistic Review 2:3.

Aoun, J. and D. Lightfoot (1984) "Government and Contraction," Linguistic Inquiry 15:3.

Baker, C.L. (1970) "Notes on the Description of English Questions : The Role of an Abstract Question Morpheme," Foundations of Language 6:2.

_____ (1978) Introduction to Generative-Transformational Syntax, Prentice Hall, Englewood Cliffs, New Jersey.

Baker, M. (1985) Incorporation : A Theory of Grammatical Function Changing, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Baker, M., K. Johnson, and I. Roberts (in progress) unpublished paper on passives.

Baltin, M. (1982) "A Landing Site Theory of Movement Rules," Linguistic Inquiry 13:1.

Barss, A. (1984) "Chain Binding," unpublished manuscript, MIT, Cambridge, Massachusetts.

_____ (forthcoming) Chains and Anaphoric Dependencies, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Barss, A. and H. Lasnik (1986) "A Note on Anaphora and Double Objects," Linguistic Inquiry 17:2.

Belletti, A. and L. Rizzi (1985) "Psych-verbs and θ-Theory," talk presented at MIT Lexicon Seminar, MIT, Cambridge, Massachusetts.

Bouchard, D. (1982) On the Content of Empty Categories, Doctoral dissertaion, MIT, Cambridge, Massachusetts.

Brame, M. (1981) "The General Theory of Binding and Fusion," Linguistic Analysis 7:3.

_____ (1982) "The Head-Selector Theory of Lexical Specificaitons and the Nonexistence of Coarse Categories," Linguistic Analysis 10:4.

Bresnan, J. (1971) "Contraction and the Transformational Cycle in English," unpublished manuscript, MIT, Cambridge, Massachusetts.

Burzio, L. (1986) Italian Syntax : A Government-Binding Approach, Reidel, Dordrecht.

Chomsky, N. (1970) "Remarks on Nominalization," in R. Jacobs and P. Rosenbaum, eds., Readings in English Transformational Grammar, Ginn, Waltham, Massachusetts.

_____ (1972) Studies on Semantics in Generative Grammar, Mouton, The Hague.

_____ (1973) "Conditions on Transformations," in S.R. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle, Holt, Rinehart and Winston, New York, New York.

_____ (1976) "Conditions on Rules of Grammar," Linguistic Analysis 2:4.

_____ (1977a) Essays on Form and Interpretation, North-Holland, Amsterdam.

_____ (1977b) "On Wh-movement," in P. Culicover, T. Wasow and A. Akmajian, eds., Formal Syntax, Academic Press, New York, New York.

_____ (1981) Lectures on Government and Binding, Foris, Dordrecht.

_____ (1982) Some Concepts and Consequences of the Theory of Government and Binding, The MIT Press, Cambridge, Massachusetts.

_____ (1985) Barriers, unpublished manuscript, MIT, Cambridge, Massachusetts (To be published by the MIT Press).

_____ (1986) Knowledge of Language : Its Nature, Origin, and Use, Praeger, New York, New York.

Chomsky, N. and H. Lasnik (1977) "Filters and Control," Linguistic Inquiry 8:3.

Chomsky, N. and H. Lasnik (1978) "A Remark on Contraction," Linguistic Inquiry 9:2.

Czaykowska-Higgins, E. (1986) "On Adverbial Modifications," unpublished manuscript, MIT, Cambridge, Massachusetts.

diSchiullo, A-M. (1980) "N and X-bar Theory," in NELS 10, University of Ottawa, Ottawa.

Emonds, J. (1985) A Unified Theory of Syntactic Categories, Foris, Dordrecht.

Epstein, S. (1984) "Quantifier-pro and the LF Representation of PRO_{ARB}, " Linguistic Inquiry 15:3.

Farmer, A. (1980) On the Interaction of Morphology and Syntax, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Farmer, A. and N. Tsujimura (1984) "Variables, Pronouns and Anaphors in Japanese : A Reply to Saito and Hoji," unpublished manuscript, University of Arizona, Tucson, Arizona.

Fillmore, C.J. (1968) "The Case for Case," in E. Bach and R.T. Harms, eds., Universals in Linguistic Theory, Holt, Rinehart and Winston, New York, New York.

Fukui, N. (1984) "Studies on Japanese Anaphora I : The Adjunct Subject Hypothesis and Zibun," unpublished manuscript, MIT, Cambridge, Massachusetts.

Fukui, N. and M. Speas (forthcoming) "Specifiers and Projection," to appear in N. Fukui, T. Rapoport, and E. Sagey, eds., MIT Working Papers in Linguistics 8, MIT, Cambridge, Massachusetts.

George, L. (1980) Analogical Generalization in Natural Language Syntax, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Guéron, J. and R. May (1984) "Extraposition and Logical Form," Linguistic Inquiry 15:1.

Goodall, G. (1984) Parallel Structures in Syntax, Doctoral dissertaion, University of California, San Diego.

Haig, J.H. (1980) "Some Observations on Quantifier Floating in Japanese," Linguistics 18.

Hale, K. (1978) "The Structure of English Sentences," unpublished class notes, MIT, Cambridge, Massachusetts.

_____ (1980) "Remarks on Japanese Phrase Structure : Comments on the Papers on Japanese Syntax," in Y. Otsu and A. Farmer, eds. MIT Working Papers in Linguistics 2, MIT, Cambridge, Massachusetts.

_____ (1981) "Preliminary Remarks on the Grammar of Part-Whole Relations in Warlpiri," in Hollyman and Pawley, eds., Studies in Pacific Languages and Cultures In Honor of Bruce Briggs, Auckland.

_____ (1982) "Preliminary Remarks on Configurationality," in NELS 12.

_____ (1983) "Warlpiri and the Grammar of Non-configurational Languages," Natural Language and Linguistic Theory 1:1.

Harris, Z. (1946) "From Morpheme to Utterance," Language 22.

_____ (1951) Methods in Structural Linguistics, The University of Chicago Press, Chicago, Illinois.

Hattori, S. (1968) Eigo Kisogoi-no Kenkyuu, Sanseido, Tokyo.

Hawkins, J.A. (1982) "Cross-category Harmony, X-bar and the Predictions of Markedness," Journal of Linguistics 18:1.

Higginbotham, J. (1980) "Pronouns and Bound Variables," Linguistic Inquiry 11:4.

_____ (1981) "Anaphora and GB : Some Preliminary Remarks," in NELS 10, University of Ottawa, Ottawa.

_____ (1985) "On Semantics," Linguistic Inquiry 16:4.

Hinds, J. (1973) "On the Status of VP Node in Japanese," Language Research 9.

Hoji, H. (1982) "X-bar Schema in Japanese and the * Parameter," unpublished manuscript, University of Washington and MIT.

_____ (1985) Logical Form Constraints and Configurational Structures in Japanese, Doctoral dissertation, University of Washington, Seattle, Washington.

Hopcroff, J.E. and J.D. Ullman (1979) Introduction to Automata Theory, Languages and Computation, Addison-Wesley, Reading, Massachusetts.

Hornstein, N. (1984) Logic as Grammar, The MIT Press, Cambridge, Massachusetts.

Hornstein, N. and D. Lightfoot (1981) "Introduction," in N. Hornstein and D. Lightfoot, eds., Explanation in Linguistics : The Logical Problem of Language Acquisition, Longman, London.

Huang, C.-T.J. (1982) Logical Relations in Chinese and the Theory of Grammar, Doctoral dissertation, MIT, Cambridge, Massachusetts.

_____ (1984) "On the Distribution and Reference of Empty Pronouns," Linguistic Inquiry 15:4.

Inoue, K. (1976a) "Reflexivization : An Interpretive Approach," in M. Shibatani, ed., Syntax and Semantics 5 : Japanese Generative Grammar, Academic Press, New York, New York.

_____ (1976b) Henkei-Bunpoo to Nihongo, Taishukan, Tokyo.

_____ (1978) Nihongo-no Bunpoo-Kisoku, Taishukan, Tokyo.

Jackendoff, R. (1972) Semantic Interpretation in Generative Grammar, The MIT Press, Cambridge, Massachusetts.

_____ (1977) X' Syntax : A Study of Phrase Structure, The MIT Press, Cambridge, Massachusetts.

Jaeggli, O. (1980) "Remarks on To Contraction," Linguistic Inquiry 11:2.

_____ (1982) Topics in Romance Syntax, Foris, Dordrecht.

_____ (1986) "Passives," unpublished manuscript, University of Southern California, Los Angeles, California (To appear in Linguistic Inquiry).

Johnson, K. (1985) A Case for Movement, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Kamio, A. (1977a) "Restrictive and Non-Restrictive Relative Clauses in Japanese," in Descriptive and Applied Linguistics 10, ICU, Tokyo.

_____ (1977b) "Suuryoosi-no sintakkusu," Gengo, 6:9.

_____ (1981) "Da-no Hantyuu-ni kansuru Oboegaki," in Gendai-no Eigogaku edited by Yasui Minoru-hakase Kanreki Kinen Ronbun-shu Hensyuu-iinkai, Kenkyusha, Tokyo.

Kayne, R. (1986) "Participles, Agreement, Auxiliaries, se/si and PRO," unpublished manuscript, MIT, Cambridge, Massachusetts.

Keyser, S.J. (1968) "Review of Sven Jacobson, Adverbial Position in English," Language 44.

Kitagawa, Y. (1984) "Clauses," unpublished manuscript, University of Massachusetts, Amherst.

Koopman, H. (1984) The Syntax of Verbs, Foris, Dordrecht.

Koopman, H. and D. Sportiche (1985) "θ-Theory and Extraction," paper presented at GLOW meeting, Brussels.

_____ (1986) "ECP and Subjects," talk given at MIT, Cambridge, Massachusetts.

Kuno, S. (1973) The Structure of Japanese Language, The MIT Press, Cambridge, Massachusetts.

_____ (1978a) "Theoretical Perspectives on Japanese Linguistics," in J. Hinds and I. Howard, eds., Problems in Japanese Syntax and Semantics, Kaitakusha, Tokyo.

_____ (1978b) "Japanese : A Characteristic OV Language," in W. Lehman, ed., Syntactic Typology, University of Texas Press, Austin, Texas.

_____ (1980) "The Scope of the Question and Negation in Some Verb-Final Languages," in CLS 16.

_____ (1986) "Anaphora in Japanese," in S.-Y. Kuroda, ed., Working Papers from the First SDF Workshop in Japanese Syntax, University of California, San Diego.

Kuroda, S.-Y. (1980) "Bunkoozoo-no Hikaku," in T. Kunihiro, ed. Niti-Eigo Hikaku-Kooza 2 : Bunpoo, Taishukan, Tokyo.

_____ (1983) "What Can Japanese Say about Government and Binding?," in WCCFL 2.

_____ (1985) "Whether You Agree or Not : Rough Ideas about the Comparative Grammar of English and Japanese," unpublished manuscript, University of California, San Diego.

Lamontagne, G. and L. Travis (1986) "The Case Filter and the ECP," talk given at MIT, Cambridge, Massachusetts.

Larson, R. (1985) "Bare NP Adverbs," Linguistic Inquiry 16:4.

Lasnik, H. (1976) "Remarks on Coreference," Linguistic Analysis 2:1.

_____ (1984) "Random Thoughts on Implicit Arguments," unpublished manuscript, University of Connecticut and MIT.

_____ (1985) "Illicit NP Movement : Locality Conditions on Chains?," Linguistic Inquiry 16:3.

Lasnik, H. and M. Saito (1984) "On the Nature of Proper Government," Linguistic Inquiry 15:2.

Levin, B. (1983) On the Nature of Ergativity, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Lewis, H.R. and C.H. Papadimitriou (1981) Elements of the Theory of Computation, Prentice-Hall, Englewood Cliffs, New Jersey.

Lightfoot, D. (1977) "On Traces and Conditions on Rules," in P. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York, New York.

_____ (1979) Principles of Syntax, Cambridge University Press, Cambridge.

Lumsden, J. (1985) "A Baroque Theory of Subjects," unpublished manuscript, MIT, Cambridge, Massachusetts.

Manzini, M.R. (1983) "On Control and Control Theory," Linguistic Inquiry 14:3.

Marantz, A. (1984) On the Nature of Grammatical Relations, The MIT Press, Cambridge, Massachusetts.

Martin, J.E. (1968) A Study of the Determiners of Preferred Adjective Order in English, Doctoral dissertation, University of Illinois.

Massam, D. (1985) Case Theory and the Projection Principle, Doctoral dissertation, MIT, Cambridge, Massachusetts.

May, R. (1977) The Grammar of Quantification, Doctoral dissertation, MIT, Cambridge, Massachusetts.

_____ (1985) Logical Form : Its Structure and Derivation, The MIT Press, Cambridge, Massachusetts.

McCawley, J. (1970) "English as a VSO Language," Language 46.

McCawley, N.A. (1976) "Reflexivization : A Transformational Approach," in M. Shibatani, ed., Syntax and Semantics 5 : Japanese Generative Grammar, Academic Press, New York, New York.

Milsark, G. and K. Safir (1983) "On Adjacency in Phonological Form," unpublished manuscript, Temple University and University of Pennsylvania.

Nakau, M. (1973) Sentential Complementation in Japanese, Kaitakusha, Tokyo.

Nash, D. (1980) Topics in Warlpiri Grammar, Doctoral dissertation, MIT, Cambridge, Massachusetts.ef

Ostler, N. (1979) Case-Linking : A Theory of Case and Verbal Diathesis Applied to Classical Sanskrit, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Okutsu, K. (1969) Seisei Nihon-Bunpoo Ron : Meisiku-no Koozoo, Taishukan, Tokyo.

Pesetsky, D. (1982) Paths and Categories, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Postal, P. (1971) Crossover Phenomena, Holt, Rinehart and Winston, New York, New York.

Postal, P. and G.K. Pullum (1982) "The Contraction Debate," Linguistic Inquiry 13:1.

Radford, A. (1981) Transformational Syntax : A Student's Guide to Chomsky's Extended Standard Theory, Cambridge University Press, Cambridge.

Reuland, E. (1984) "Governing -ing," Linguistic Inquiry 14:1.

Reinhart, T. (1979) "Syntactic Domains for Semantic Rules," in F. Guenthner and S.J. Schmidt, eds., Formal Semantics and Pragmatics for Natural Languages, Reidel, Dordrecht.

Riemsdijk, H.van, and E. Williams (1981) "NP Structure," The Linguistic Review 1:3.

Rizzi, L. (1982) "On Chain Formation," unpublished manuscript, Universita della Calabria, Cosenza.

Roberts, I. (1985) The Representation of Implicit and Dethematized Subjects, Doctoral dissertation, University of Southern California, Los Angeles, California.

Roeper, T. (1983) "Implicit Thematic Roles in the Lexicon and Syntax," unpublished manuscript, University of Massachusetts, Amherst, Massachusetts.

_____ (1984) "Implicit Arguments and the Projection Principle," unpublished manuscript, University of Massachusetts, Amherst, Massachusetts.

Rosenbaum, P.S. (1970) "A Principle Governing Deletion in English Sentential Complementation," in R.A. Jacobs and P.S. Rosenbaum, eds., Readings in English Transformational Grammar, Ginn, Waltham, Massachusetts.

Rothstein, S. (1983) The Syntactic Forms of Predication, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Saito, M. (1982a) "Case Marking in Japanese : A Preliminary Study," unpublished manuscript, MIT, Cambridge, Massachusetts.

_____ (1982b) "Scrambling, Topicalization, and Strong Crossover," unpublished manuscript, MIT, Cambridge, Massachusetts.

_____ (1983) "Comments on the Papers on Generative Syntax," in Y. Otsu, et al, eds., Studies in Generative Grammar and Language Acquisition, ICU, Tokyo.

_____ (1985) Some Asymmetries in Japanese and Their Theoretical Implications, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Saito, M. and H. Hoji (1983) "Weak Crossover and Move-α in Japanese," Natural Language and Linguistic Theory 1:2.

Sakuma, K. (1936) Gendai Nihongo-no Hyoogen-to Gohoo, Kooseikaku, Tokyo.

Schwartz, A. (1972) "The VP-constituent of SVO Languages," in J. Kimball, ed., Syntax and Semantics 1, Academic Press, New York, New York.

Selkirk, E.O. (1972) The Phrase Phonology of English and French, Doctoral dissertation, MIT, Cambridge, Massachussets.

Shibatani, M. (1977) "Grammatical Relations and Surface Cases," Language 53.

Simpson, J. (1983) Aspects of Warlbiri Morphology and Syntax, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Speas, M. (forthcoming) Adjunctions and Projections in Syntax, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Stowell, T. (1981) Origins of Phrase Structure, Doctoral dissertation, MIT, Cambridge, Massachusetts.

_____ (1982) "Subjects Across Categories," The Linguistic Review 2:3.

Taraldsen, K.T. (1983) Parametric Variation in Phrase Structure : A Case Study, Doctoral dissertation, University of Transø.

Travis, L. (1984) Parameters and Effects of Word Order Variation, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Wahl, A. (1985) "Two Types of Locality," unpublished manuscript, University of Maryland, College Park, Maryland.

Washio, R. (1983) "Tensedness and Related Matters," in K. Inoue, E. Kobayashi, and R. Linde, eds., Issues in Syntax and Semantics : Festschrift for Masatake Muraki, Sansyusya, Tokyo.

Wasow, T. (1972) Anaphoric Relations in English, Doctoral dissertation, MIT, Cambridge, Massachusetts.

_____ (1979) Anaphora in Generative Grammar, E. Story-Scientia Gent.

Whitman, J. (1981) "The Internal Structure of NP in Verb Final Languages," in CLS 17.

_____ (1982) "Configurationality Parameters," unpublished manuscript, Harvard University, Cambridge, Massachusetts.

_____ (1984) "Configurationality Parameters," unpublished manuscript, Harvard University, Cambridge, Massachusetts (A revised version of Whitman (1982)).

Whitney, R. (1984) The Syntax and Interpretation of A'-Adjunctions, Doctoral dissertation, University of Washington, Seattle, Washington.

Williams, E. (1980) "Predication," Linguistic Inquiry 11:1.

_____ (1985a) "NP Trace in θ-Theory," unpublished manuscript, University of Massachusetts, Amherst, Massachusetts.

_____ (1985b) "PRO and Subject of NP," Natural Language and Linguistic Theory 3:3.

Ziff, P. (1960) Semantic Analysis, Cornell University Press, New York.

Zubizarreta, M.L. (1985) "The Relation between Morphophonology and Morphosyntax : The Case of Romance Causatives," Linguistic Inquiry 16:2.