INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*

4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

***For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.**

UMI University Microfilms International

8603940

Clark, Robin Lee

BOUNDARIES AND THE TREATMENT OF CONTROL

University of California, Los Angeles PH.D. 1985

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

UNIVERSITY OF CALIFORNIA

Los Angeles

Boundaries and the Treatment of Control

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Linguistics

by

Robin Lee Clark

1985

The dissertation of Robin Lee Clark is approved.

Edward L. Keenan

Kenneth Wexler

Timothy A. Stowell, Committee Chair

University of California, Los Angeles

1985

TABLE OF CONTENTS

To my wife, Maryellen.

Symmetry makes me sick.

--Luis Buñuel
The Phantom of Liberty

Acknowledgements

I would like to begin with a story: According to Aristophanes (Clouds), Socrates and his friend Chaerephon became interested in how many times its own length a flea could jump. It is, as one might imagine, no simple task to measure the length of a flea. Socrates and Chaerephon, however, were quite learned individuals and they were aware of a technique, developed by the great geometer, Pythagoras, called ex pede Herculem. Literally translated, this means "from the foot, Hercules." At the time of Pythagoras, a stadium measured six hundred feet in length, but the Hercules' stadium at Olympia was much longer. Hence, reasoned Pythagoras, the foot of Hercules was proportionately longer than the normal human foot. Now, the foot bears a certain ratio to the height, so the height of Hercules can be easily derived from the length of his foot. Socrates and Chaerephon

applied Pythagoras' technique to the flea. By taking an impression in wax of a flea's foot, they were able, ex pede Herculem, to derive the flea's height. The solution to their question was now trivial. They compelled the flea to jump from Socrates' hand to Chaerephon's hand, measured the distance and derived the answer by means of simple multiplication. Having told my story I would like to thank the following people for their help with various measurements over the years.

First, I would like to thank Tim Stowell, the chairman of my committee, for his sound advice and criticisms. I have derived immense profit from our association, more than I can repay. Over the years, he has proved himself to be the best of all possible chairmen.

Next, many thanks are due to Ed Keenan, who taught me the true use of formalism. I have placed great value on our conversations and I look forward to many more in the years to come. Ed is among the most reasonable (and humane) of individuals on top of being a damn good linguist.

I owe thanks to the other members of my committee: Hagit Borer, Carlos Otero, Eric Wehrli and Ken Wexler.

Among my many debts is the one I owe to Joseph Aoun, without whom much of what follows would have been inconceivable. Our work together has been, to say the least, gratifying and, again, I look forward to a continued association.

I would also like to thank Mario Montalbetti who has influenced my thinking in many subtle ways. In particular, he

has persuaded me that there is no single shining path into Domain D. I value his friendship highly.

Among other people whose friendship I value highly are Osvaldo Jaeggli and Nina Hyams. Again, the debt is greater than my repayment may ever be.

Turning to the good old days, I would like to thank Mike Hammond, John V. Singler, Tracy Thomas-Flinders and Laurie Tuller. Together, they conspired to make my years at UCLA much richer than they would otherwise have been. I can only hope that Mike will not take my attack on trees personally.

Gratitude is also in order for Vicki Carstens, Noam Chomsky, Sandy Chung, Susanna Cumming, Victoria Fromkin, Allen Halsell, Susan Halsell, Bruce Hayes, Jim Higginbotham, Norbert Hornstein, Kyle Johnson, Juliette Levin, Geoff Lindsay, Richard Kayne, Wilson Manik, Jack Martin, Robert May, Pam Munro, Fritz Newmeyer, David Pesetsky, Carlos Quicoli, Ken Safir, Mamoru Saito, Paul Schachter, and Robert Stockwell. I hope that I haven't left anyone out. In addition, I want to thank my parents, Ralph and Rosemond, for all that they have given me. I also must thank Russ Abbott and the friendly folks at Silogic, Inc.

Most importantly, I must thank my wife, Maryellen MacDonald, for the years of love and patience she has given me (not to mention the countless grammaticality judgements). She is a rare combination of brilliance and tenderness. It is to her that I dedicate this dissertation.

VITA

March 8, 1957	Born, Tucson, Arizona
1980	B.A., University of Texas, Austin
1983	M.A., University of California, Los Angeles
1981-1984	Teaching Assistant, Department of Linguistics University of California, Los Angeles

PUBLICATIONS

Clark, R. (1983). "Clitics as $\bar{A}$-binders", in Cornell Working Papers in Linguistics, #4.

Clark, R. (1984). "The Syntactic Nature of Logical Form: Evidence from Toba Batak" in (P. Schachter, ed.) Studies in the Structure of Toba Batak, UCLA Occasional Papers in Linguistics, Number 5, June 1984

Clark, R. (1984). "Control into NP", Proceedings of the Third West Coast Conference on Formal Linguistics, 1984

Clark, R. (1985). "PRO in NP", Proceedings of the Fifteenth meeting of the North Eastern Linguistics Society.

Clark, R. and Aoun, J. (1985). "On Non-Overt Operators",in (G. Gilligan, M. Mohammad & I. Roberts, eds.) Studies in Syntax, Southern California Occasional Papers in Linguistics, Number 10, May 1985

Clark, R. (1985). "The Syntactic Nature of Logical Form: Evidence from Toba Batak", Linguistic Inquiry, 16.4, 1985, pp. 663-669

ABSTRACT OF THE DISSERTATION

Boundaries and the Treatment of Control

by

Robin Lee Clark
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 1985
Professor Timothy A. Stowell, Committee Chair

The unifying theme of the dissertation is that properties of lexical argument structure "drive" the syntax in a number of interesting ways. First, lexical argument structure plays an important role in the determination of extraction possibilities in the syntax. Second, lexical properties are important in determining a number of phenomena at Logical Form; in particular, lexical semantics plays an important role in determining the interpretation of structures of "arbitrary" control.

Chapters two and three of the dissertation deals with boundaries to extraction, particularly the phenomena subsumed under the Subject Condition and the Constraint on Extraction Domains. Chapter two focuses on a restricted class of nominals in English. The main puzzle addressed is the ability to strand prepositions in these nominals but not in other sorts

of nominals. In chapter three, the ability to extract from a constituent is related to the thematic relations which the constituent in question enters into. It is further demonstrated that the Boundary Condition allows us to abstract away from details of tree configuration in providing an account of these island phenomena.

Chapters four and five develop an account of control based on recent research on Non-overt Operators. Particular attention is paid to so-called arbitrary control and it shows that arbitrary control differs from obligatory control only insofar as the former is a property of Logical Form while the latter is an S Structure property. Particular attention is given to the nature of Logical Form, how implicit arguments are realized at that level and how adverbs of quantification enter into control relations. The treatment of control is shown to bear a strong relationship to such divers structures as purposive clauses, parasitic gaps, infinitival relatives, "tough" movement constructions and certain sentential predicates.

Chapter 1
Introduction

In this work we will consider two distinct topics within the framework of Government-Binding theory: First, we will examine the nature and formulation a boundary condition that governs syntactic representations; second, we will examine the status of a particular phonologically empty noun phrase, PRO, and the ways in which this element may enter into coreference relations with other noun phrases. In general, particular theoretical principles will be introduced and discussed as needed, so no general overview of the particular principles of Government-Binding theory will be given in this chapter. Instead, we will focus on some more general properties of the functioning of Government-Binding theory, particularly the fact that the theory assumes distinct levels of mental representation and the modular character of the theory.

To say that the theory assumes distinct levels of representation is roughly equivalent to saying that, for each well-formed string in a language, the grammar assigns an ordered n-tuple of labelled bracketings, minimally consisting of a D Structure representation, an S Structure representation, a PF (Phonetic Form) representation and an LF (Logical Form) representation. The labelled bracketings are related by means of a general rule component ("move alpha" and possibly some language-particular rules involving the deletion of some specified element(s) and "stylized" movement), and each level of representation must obey certain constraints on representation placed upon it by the grammar. We can represent this visually by means of the following diagram:

(1)

```
                    D Structure
                          |
                          |  "Move alpha"
                          |
                    S Structure
                        /\
"Stylized" Rules  /      \  "Move alpha"
                  /        \
                 PF         LF
```

We will return to particular details of the above diagram shortly. For the moment, let us restate the rule "move alpha" as "move any element in a labelled bracketing anywhere in that labelled bracketing, the output being subject to the constraints placed upon representations by the grammar."

To say that the theory has a modular character means that the theory of grammar is divided into a number of sub-theories

(i.e., "modules") each of which operates independently of the others. Chomsky (1982) gives the following set of modules:

(2)a. $\overline{X}$ theory
b. θ-theory
c. Case theory
d. Binding theory
e. Bounding theory
f. Control theory
g. Government theory

Each of the above modules may be in operation at one or more of the levels of representation. We will be most concerned with this last fact in this chapter, since Government-Binding theory derives much of its theoretical interest and explanatory power from the relationship between levels of representation and modules of grammar.

Let us begin by hypothesizing the existence of an infinite set of labelled bracketings (trees). For convenience, we will call this set "T*". The set of possible D Structures for a given language will be just those labelled bracketings in T* that (1) obey $\overline{X}$ theory, (2) obey θ-theory and (3) obey the Extended Projection Principle. Let us consider, briefly, what it means for a labelled bracketing to obey $\overline{X}$ theory. $\overline{X}$ theory is, roughly, a filter on labelled bracketings; schematically, it can be stated in the following way (substituting primes for the number of bars above the X):

(3)a. X" --> Spec X´
b. X´ --> X Comp

The "X" in (3) is a variable that ranges over nouns ([+N, -V]), verbs ([-N, +V]), adjectives ([+N, +V]), and prepositions ([-N,

-V]). "Spec" (Specifiers) and "Comp" (Complements) are also variables that range over grammatical categories, determiners and so on. For any given language, it is part of the task of the language learner to determine, on the basis of experience, exactly which items may instantiate Spec and Comp, whether X´ precedes or follows Spec and whether X precedes or follows Comp.

For English, it appears that Spec precedes X´ and X precedes Comp. Thus, taking the category "determiner" to be a possible Spec for N´, we can assign the following labelled bracketing for the NP, <u>the destruction of the city</u>:

(4) $[_{N''}\ [_{Det}\ \text{the}]\ [_{N'}\ [_{N}\ \text{destruction}]\ [_{P''}\ [_{P'}\ [_{P}\ \text{of}]\ [_{N''}\ [_{Det}\ \text{the}]\ [_{N'}\ [_{N}\ \text{city}]]]]]]]$

where the determiner always precedes N´ and the head noun, <u>destruction</u>, precedes its complements. Notice that Spec and Comp may be null. Fortunately, we rarely need to present labelled bracketings in such detail as the example in (4) during the course of a theoretical discussion. We will take the category S (clause) to be a projection of an abstract element Infl (inflection), the Spec of which must be instantiated by an NP.[1] Given what we have said about English, the expansion of S will be as in (5):

(5)a. Infl" --> N" Infl´
   b. Infl´ --> Infl Comp

---

1 Infl is not associated with a lexical item but, rather, contains abstract grammatical information regarding tense, aspect, mood and agreement.

I will assume that the Comp of Infl is generally instantiated by VP, although it may be instantiated by an adjective phrase, a prepositional phrase or a noun phrase in copular sentences.

Consider, now, the following set of labelled bracketings (I have eliminated irrelevant details under the assumption that the internal structure of the NPs and VPs obey $\overline{X}$ theory):

(6)a. [[[ate the quiche] Infl] [the dog]]
Infl" Infl´ VP NP

b. [[Infl [ate the quiche]] [the dog]]
Infl" Infl´ VP NP

c. [[the dog] [[ate the quiche] Infl]]
Infl" NP Infl´ VP

d. [[the dog] [Infl [ate the quiche]]]
Infl" NP Infl´ VP

The examples in (6) show the possible orderings of the constituents which are internal to Infl" and Infl´. From what we said about the way in which English instantiates $\overline{X}$ theory, it follows that only (6d) is a potential candidate for being an element of the set of D Structures for English. The example in (6a) is ruled out because the Spec of Infl´ precedes its head (i.e., Infl´) and the Comp of Infl precedes its head (i.e., Infl). Example (6b) is ruled out because the Spec of Infl´ precedes its head. Finally, example (6c) is ruled out because the Comp of Infl precedes its head. Notice that since Government-Binding theory is modular and contains distinct levels of representation, we could exempt D Structure from obeying $\overline{X}$ theory, in which case all of the structures in (6) would be potential D Structures of English.

To say that D Structures are those elements of T* which obey

θ-theory and the Extended Projection Principle means, roughly, that every element in the labelled bracketing occurs in a position that receives a thematic (θ) role and only one thematic role and, furthermore, that Infl" must have a subject position. Consider a particular lexical item like hit. We know that hit takes an object (an internal argument); thus, the following subsequence of a labelled bracketing is well-formed at D Structure:

(7) $[_{V''}\ [_{V'}\ [_{V}$ hit$]\ [_{N''}$ Bill$]]]$

while the labelled bracketing in (8) is not a possible subsequence of any D Structure:

(8) $[_{V''}\ [_{V'}\ [_{V}$ hit$]]]$

Furthermore, the VP, hit Bill, assigns a thematic role to a subject position (an external argument). From this fact and the Extended Projection Principle, it follows that the labelled bracketing in (9) is a possible D Structure:

(9) $[_{Infl''}\ [_{N''}$ John$]\ [_{Infl'}$ Infl $[_{VP}$ hit Bill$]]]$

Notice that the Extended Projection Principle requires that Infl" have a Spec position even if that position does not receive a thematic role. If we take a VP like seems that John is sick, then the labelled bracketing in (10a) is a possible D Structure for English while the labelled bracketings in (10b) and (10c) are

not (note that e represents the empty string):[2]

(10)a. [[e] [Infl [seems [that John is
Infl" N" Infl´ VP $\bar{S}$
sick]]]]

b. [[Infl [seems [that John is
Infl" Infl´ VP $\bar{S}$
sick]]]]

c. [[Bill] [[seems [that John is
Infl" N" Infl´ VP $\bar{S}$
sick]]]]

The structure in (10b) violates the requirement placed on representations by the Extended Projection Principle, since Infl" lacks a Spec position. The structure in (10c) has a Spec position, but this position is occupied by a referential NP, Bill, which does not receive a thematic role; hence, (10c) violates the restriction placed on representations by θ theory that arguments must receive a thematic role.

Consider, in the light of what we said above, the status of the representation of passive sentences at D Structure. According to the analysis of passives found in Chomsky (1981) and Jaeggli (1985), the thematic role assignment to the subject position (Spec of Infl") is blocked by the passive morphology. Of the following set of labelled bracketings, only one, the bracketing in (11a), is a possible candidate for the D Structure representation of Bill was hit:

---

2 I will assume that the empty string in subject position comes to be covered by the expletive element, it, which does not refer despite having a phonological realization.

(11)a. [[e] [[was] [hit [Bill]]]]
Infl" N" Infl´ Infl VP N"

b. [[[was] [hit [Bill]]]]
Infl" Infl´ Infl VP N"

c. [[Bill] [[was] [hit [e]]]]
Infl" N" Infl´ Infl VP N"

d. [[Bill] [[was] [hit]]]
Infl" N" Infl´ Infl VP

In (11a), Bill occurs in the object position of hit and receives the "theme" role from hit; the Spec of Infl" is present, but is lexically empty, as required by θ theory. In (11b), the Spec of Infl" is not present and, hence, the Extended Projection Principle is violated. In (11c), Bill is present in the Spec of Infl" position, which does not receive a thematic role; hence, θ theory is violated since a referential NP is not incorporated into the argument structure of the predicate. Furthermore, the empty string, e, is in the object position of hit; since the empty string is non-referential, it cannot bear the theme role and θ theory is violated. In the final example, (11d), Bill does not receive a thematic role and there is nothing after the verb, hit, to receive the theme role. Hence, θ theory will rule out (11d) as a possible D Structure representation.

We can see, from the above examples, that $\overline{X}$ theory, θ theory and the Extended Projection Principle work together to limit the labelled bracketings in T* that can count as possible D Structures for any given language. It should be observed, at least in passing, that θ theory is based on the lexical specifications of lexical items, so an important part of the theory of Universal Grammar is to restrict the set of possible

lexical representations. Since the theory is modular, we can imagine that any of the above sub-theories does not apply to D Structure representations. As the reader can easily verify, any such change would significantly alter the character of the set of possible D Structures.

We will say that S Structure is the closure of the set of D Structures under the operation "move alpha" (see above). This means that for any element of the set of D Structures, application of "move alpha" to any subpart of the labelled bracketing results in an S Structure. "Move alpha" may be applied any number of times to a D Structure, including 0 times and the result is an S Structure. The set of possible S Structures for a given language are just those S Structures that obey a number of constraints. For example, we may wish to restrict "move alpha" in such a way that it will only apply to maximal projections (X"). Let us assume, for a moment, that a possible S Structure is an S Structure which obeys the binding theory, Case theory and the Extended Projection Principle, in addition to the previously mentioned constraint. We will discuss the binding theory extensively in chapter 4 of this dissertation, so we will present only an informal view of Case theory.

Case theory requires that NPs which are realized phonologically must bear abstract Case. An NP bears abstract Case only if it stands in a particular structural relationship with some other element in a labelled bracketing. The exact nature of this structural relationship is the subject of

Government theory, so I will put off further discussion of it here. Instead, let us state the environments of abstract Case assignment in the following way:

(12)a. An Infl with the feature [+tense] assigns Nominative Case to the subject.
b. A verb assigns Accusative Case to its object.
c. A preposition assigns Oblique Case to its object.
d. The Spec of N" may receive Genitive Case.

Consider the following representations:

(13)a. [[Bill] [[to] [see Mary]]]
Infl" N" Infl´ Infl VP
b. [for [[Bill] [[to] [see
S Infl" N" Infl´ Infl VP
Mary]]]] would be a mistake

c. John wants [[Bill] [[to] [see
Infl" N" Infl´ Infl VP
Mary]]]

The representation in (13a) is not a possible S Structure representation because there is no tensed Infl to assign Nominative Case to the subject, Bill. Hence, the representation violates the constraint that phonologically realized NPs must bear abstract Case. In (13b), the subject of Infl" will receive Oblique Case from the preposition, for, it will count as a possible S Structure representation since it does not violate Case theory. Finally, the subject of the embedded Infl" in the S Structure in (13c) will receive Accusative Case from the verb want so it too counts as a possible S Structure representation.

Let us return, for the moment, to the issue of passive sentences. Recall that the possible D Structure of a passive sentence must be parallel to:

(14) [[e] [[was] [hit [Bill]]]]
Infl" N" Infl´ Infl VP N"

A fact about passive morphology in English is that it prevents the assignment of Accusative Case by any verb bearing this morphology (see the references cited above). From this fact, it follows that (14) is not a possible S Structure representation although it is a possible D Structure representation. The only way in which we can map (14) onto a possible S Structure representation is if we place the object NP in a position that will receive Case. The grammar provides us with only one mechanism for doing this: We must apply "move alpha" to the NP, Bill, in such a way that it lands in a Case-marked position. Hence, the following is an image of the D Structure representation in (14) that will count as a possible S Structure:

(15) [[Bill] [[was] [hit [e]]]]
Infl" N" i Infl´ Infl VP N" i

(Notice that I have included indices on the moved NP, which is required by the binding theory; see chapter 4.) In (15), Bill will receive Nominative Case from the tensed Infl.[3]

---

[3] It is not possible to give a detailed account of Case theory in this limited space. The following examples appear to be problematic, but have received extensive attention in the literature:

(i) John gave Mary a book.

(ii) Bill believes [John to be innocent]
S

Example (i) is an instance of dative shift and it appears that two NPs must receive Case. For discussion, see Kayne (1984) and the references cited there. Example (ii) is an instance of

Let us pause to make a few observations. The mechanism which we used to rule out (13a) from the set of possible S Structure representations is exactly the same mechanism which we used to rule out (14) from the set of possible S Structure representations and is also the same mechanism which "forced" Bill to move to the subject position in (15). There is no sense in which we had to state an obligatory passive rule to generate the representation in (15). This illustrates one of the advantages of a modular grammar. Suppose that we had no sub-theory concerned with Case to which we could make appeal in deriving the above facts. We would, therefore, be forced to state explicitly that an overt (i.e., phonologically realized) NP may occur in the subject position of an infinitival clause just in case it is immediately preceded by the preposition for or by a verb like want. Furthermore, we would be forced to make a special provision concerning passives; namely, an NP obligatorily moves from the object position of a verb bearing passive morphology to the Spec position of Infl". Notice that these special provisions do not make any generalization about the mechanism underlying the derivation of passive sentences and the distribution of overt NPs in the subject position of infinitival clauses.

Exceptional Case Marking; see Chomsky (1981) and the references cited there.

By exploiting the sub-theory of grammar concerned with Case, we were able to state a generalization concerning the distribution of NPs that is not otherwise available to us. Furthermore, in our imaginary theory which lacked this sub-theory we were forced to state a special obligatory rule which was concerned only with the derivation of passive sentences. Thus, by assuming the existence of Case theory, we may state rules in their most general form--"move alpha"--and, furthermore, we need not assume that rules apply obligatorily since we can use the various modules of grammar to filter out ill-formed representations. The ability to state rules in their most general form is, of course, of incalculable benefit for the field of comparative linguistics and makes it possible to pursue the problem of specifying the possible limits of variation within the class of possible grammars for natural languages. For example, rather than concentrating on the form that the rule of passivization takes in various languages--an impossibility, given that there is no such rule in any natural language--emphasis would be placed on studying the principles of grammar (including, of course, the various modules responsible for the derivation of passive constructions in these languages). Notice that the term "passive" takes on a taxonomic nature in this viewpoint. In short, a modular grammar provides an extremely reasonable research program for a theory that seeks to give an account of what it means to be a possible human language; to my mind, a modular grammar is the null hypothesis.

Returning, now, to the various levels of representation, let us say that LF is the closure of the set of possible S Structures under the rule of Quantifier Raising. Quantifier Raising is a form of "move alpha" which applies to NPs having a quantifier or wh-elements that have been left in place in the syntax and adjoins them to some (maximal) projection (see May, 1985 for some further discussion). Thus an S Structure like that in (16a) will be mapped onto an LF like that in (16b):

```
(16)a. [        [   John] [        Infl [   saw
        Infl"  N"         Infl´        VP
       [   every student]]]]
        N"
    b. [        [   every student]  [        [   John]
        Infl"  N"                 i  Infl"  N"
       [        Infl [   saw  [   e] ]]]]
        Infl´         VP        N"  i
```

Intuitively, LF is a level which represents the scope relationship between various logical operators, where the scope of an operator is defined as the set of nodes which it c-commands.[4] Naturally, we must place a number of constraints in the closure of the set of possible S Structures under "move alpha" before we have the notion of the set of possible LF representations. Let us assume that among these constraints is the Empty Category Principle (ECP), which concerns itself with the distribution of the empty string, e. While we will return to a more precise formulation of the ECP in subsequent chapters, we

---

4 On the definition of c-command, see Reinhart (1983). We will have occasion to discuss the c-command relation below.

can informally state some relevant consequences of this principle here:

(17)a. An empty string in the object position of a verb obeys the ECP.

b. An empty string in the subject position of a tensed S obeys the ECP only if it is (i) immediately preceded by its antecedent or (ii) immediately preceded by an empty category left by its antecedent.

While the "precede" relation mentioned in (17b) is, technically, not the correct relation between an empty category and its antecedent, it will serve our purposes for the moment. As a consequence of (17a) the following (partial) labelled bracketings are well-formed at LF:

(18)a. $[_{\bar{S}}$ who$_i$ $[_{Infl''}$ did John $[_{VP}$ see $\underline{t}_i$]]]

b. $[_{\bar{S}}$ who$_i$ $[_{Infl''}$ Bill think $[_{\bar{S}}$ that $[_{Infl''}$ John $[_{VP}$ saw $\underline{t}_i$]]]]]

The empty string in each of the examples in (18) is in object position and, so, they obey the ECP. Consider, now, empty strings left in the subject of a tensed S:

(19)a. $[_{\bar{S}}$ who$_i$ $[_{Infl''}$ $\underline{t}_i$ saw Mary]]

b. $[_{\bar{S}}$ who$_i$ $[_{Infl''}$ did Bill think $[_{\bar{S}}$ $\underline{t}_i$ ´ $[_{Infl''}$ $\underline{t}_i$ saw Mary]]]]

c. $[_{\bar{S}}$ who$_i$ $[_{Infl''}$ did Bill think $[_{\bar{S}}$ that $[_{Infl''}$ $\underline{t}_i$ saw Mary]]]]

The empty category in the matrix subject position in example (19a) is immediately preceded by its antecedent, <u>who</u>. According to the result in (17b), this empty element obeys the ECP and, all

other things being equal, the labelled bracketing in (19a) should be a well-formed LF representation. In example (19b), the empty category in the subject position of the embedded clause is immediately preceded by the trace left by its antecedent (i.e., t´). Again, the labelled bracketing in (19b) should count as a well-formed LF representation. Finally, consider the labelled bracketing in (19c). Here, the trace left in the subject position of the embedded clause is not preceded either by its antecedent of by the trace of its antecedent. It is, rather, immediately preceded by the complementizer, that (for further discussion, see Pesetsky, 1982). The labelled bracketing in (19c) is, therefore, not a possible LF representation since it violates the ECP; this is, indeed, the correct result given the ungrammaticality of:

(20) *who did Bill think that saw John

Having seen a bit of how the ECP works to constrain LF representations, let us consider the problem of multiple wh-questions in English (for a recent discussion, see Lasnik & Saito, 1984). In particular, consider the following contrast:

(21)a. Who saw what?
b. *what did who see

Recall that, as part of the mapping to LF, the rule of QR must move wh-elements that have been left in situ at S Structure. We may assume that (21a) has the following labelled bracketing at S Structure:

(22) $[_{\bar{S}}$ who$_i$ $[_{Infl''}$ $\underline{t}_i$ $[_{VP}$ saw what]]]

Notice that <u>who</u> has already moved in the mapping from D Structure to S Structure. Given what we have said, we may assume that the S Structure in (22) is mapped onto the following LF:

(23) $[_{\bar{S}}$ what$_j$ who$_i$ $[_{Infl''}$ $\underline{t}_i$ $[_{VP}$ saw $\underline{t}_j$]]]

Recall that an empty category in the object position of a verb always obeys the ECP, so the trace left by applying QR to <u>what</u> is well-formed in (23). Since the empty category in the subject position of the clause is immediately preceded by its antecedent, <u>who</u>, the labelled bracketing in (23) counts as a well-formed LF representation.

Consider, now, the derivation of (21b). We may assume that (21b) has the following labelled bracketing at S Structure:

(24) $[_{\bar{S}}$ what$_j$ $[_{Infl''}$ did who $[_{VP}$ see $\underline{t}_j$]]]

Since <u>who</u> has been left in the subject position, QR must apply to the above representation to move it to Comp where it may be assigned a scope. The only possible LF representation is:

(25) $[_{\bar{S}}$ who$_i$ what$_j$ $[_{Infl''}$ did $\underline{t}_i$ $[_{VP}$ see $\underline{t}_j$]]]

As usual, the empty category following the verb presents no problem since it obeys the ECP. The empty category in the subject position of the labelled bracketing in (25) is a different matter, however. Since this empty category is preceded neither by its antecedent nor by the trace left by its antecedent, it violates the ECP. Hence, the labelled bracketing

in (25) is not an element of the set of possible LF representations.

Let us return to the question of how the assumption that the theory of grammar exploits distinct levels of representation. Suppose, for the sake of argument, that the ECP does not apply to constrain the set of possible LF representation but, rather, is a constraint placed upon the set of possible S Structure representations. As we have already seen, the S Structure representation of (21b) is:

(26) $[_{\bar{S}}$ what$_j$ $[_{\text{Infl}''}$ did who $[_{\text{VP}}$ see $\underline{t}_j]]]$

The labelled bracketing in (26) contains only one empty category. Since we are assuming that the ECP applies at S Structure, we must check to see if this empty category obeys the ECP. Since it occurs in the object position of the verb <u>see</u>; from the result in (17a), we already know that such an empty category obeys the ECP. Hence, the above S Structure is well-formed. Let us return now to the LF representation associated with (26):

(27) $[_{\bar{S}}$ who$_i$ what$_j$ $[_{\text{Infl}''}$ did $\underline{t}_i$ $[_{\text{VP}}$ see $\underline{t}_j]]]$

Although the LF representation in (27) violates the ECP, this will not present a problem since we are assuming that the ECP applies only to S Structure. All of the wh-elements in (27) have been assigned a scope, so the labelled bracketing in (27) will count as a well-formed LF representation. We will therefore incorrectly predict the grammaticality of (21b) under the

hypothesis that the ECP applies only to S Structure representations. In order to save this theory, one would be forced to place an independent constraint on the grammar to the effect that a wh-element left in situ in the subject position of a tense S is ill-formed. But such a constraint will not apply to rule out ill-formed structures like that in (19c) (repeated here as (28)):

(28) [$_{\bar{S}}$ who$_i$ [$_{Infl''}$ did Bill think [$_{\bar{S}}$ that [$_{Infl''}$ $\underline{t}_i$ saw Mary]]]]

The above examples indicate how the assumption that there are several distinct levels of representation involved in the derivation of sentences contributes to the nature of the generalizations we can make about the human language faculty. Because we assumed an abstract level of representation, LF, at which wh-elements are assigned a scope, we could collapse the explanation of the ungrammaticality of structures like that in (28) with the explanation the effects found in wh-in-situ constructions like those in (21). It is not at all obvious how a theory of grammar which assumes only a single level of representation could provide so general an account of these phenomena (for further discussion, see Clark, 1985).

We have seen, in the preceding discussion, how modularity and distinct levels of representation significantly increase the explanatory powers of the theory of grammar. Great care must be taken, however, in explaining the characteristics of particular

constructions. The researcher must take care in attributing these characteristics to particular modules of grammar; for example, it is not a priori obvious whether to attribute the ungrammaticality of the following construction to the ECP or to bounding theory:

(29) $*[_{\bar{S}}$ what$_j$ [Infl" did John wonder $[_{\bar{S}}$ who$_i$ [Infl" $\underline{t}_i$ saw $\underline{t}_j$]]]]

Additional evidence must be found that we can bring to bear upon the question. The construction in (30), for example, is significantly worse than that in (29):

(30) $*[_{\bar{S}}$ who$_i$ [Infl" did John wonder $[_{\bar{S}}$ what$_j$ [Infl" $\underline{t}_i$ saw $\underline{t}_j$]]]]

Since (30) is a clear violation of the ECP, we can provide an account of the relative status of the examples in (29) and (30) by assuming that (29) violates only bounding theory while (30) violates both the ECP and bounding theory. We would then expect that (30) should be worse than (29).

Similarly, great care must be taken in determining exactly which modules of grammar apply at the various levels of representation. We saw above the assumption that the ECP applies only at S Structure had significant effects upon the potential analyses of various constructions. The far-reaching effects of the modularity hypothesis and the hypothesis that various levels of grammatical representation exist has led to a deductive character in grammatical analysis which has great appeal given

the problem of language acquisition. That is, we may assume that a child acquiring a particular language does not induce rules that account for particular grammatical constructions but, rather, deduces the properties of these grammatical constructions from principles of grammar. Again, the problem of the possible variation in natural languages is of central concern. This, in turn, has led to the assumption that Universal Grammar is made up of a set of parameterized principles; in other words, the principles of Universal Grammar contain variables which must be instantiated in deriving the grammar of a particular language (for some discussion, see the papers in Hornstein & Lightfoot, 1981).

This dissertation is structured as follows: In chapter 2, we will consider the properties of a restricted, but quite interesting, set of nominal constructions in English. It will become apparent that the properties of these nominals follow, in large part, from the assumption that an empty nominal element is being moved internally to the NP. In chapter 3 we will approach the question of whether the output of this movement is constrained only be the ECP or by some principle of bounding theory. As a wider range of data is considered, it will become apparent that the output movement is constrained by bounding theory and we will formulate a boundary condition. We will then test this condition across a divers set of constructions to determine the level of representation at which this condition

applies. In chapter 4, we will consider various approaches to the binding theory, in particular those of Chomsky (1981) and Aoun (1985). We will also outline an account of non-overt operators found in Aoun & Clark (1985). In chapter 5, we will apply this account of non-overt operators to the analysis of control structures, as in:

(31)a. John tried [PRO to win]
b. It is rarely easy [PRO to please John]

where it has been assumed that a special empty category, PRO, is involved. We will account for these structures in terms of the binding of non-overt operators. Again, the question of the level of representation that a particular relation holds of and the constraints placed upon these relations will be of central concern. We will also discuss various properties of LF that have received little attention in Government-Binding theory which will in turn shed some light on the problem of control.

Chapter 2
On a Certain Class of Nominals

In this chapter, I will examine the properties of a rather small class of constructions in English:

(1)a. My room needs a thorough picking up.
b. His criticisms don't merit worrying about.

Despite the rather limited distribution of this construction, an examination of its properties has some rather interesting ramifications for the theory of empty categories and the way in which this theory interacts with thematic (θ) theory. I will also defend the hypothesis that syntactic representations include empty categories. This hypothesis provides us with a beneficial methodology insofar as it forces us to investigate constructions in a way that furnishes great explanatory power and allows us to formulate research questions that may otherwise be unavailable. Furthermore, the theory is so constrained as to give us a

relatively small hypothesis space; given a theory of empty categories similar to the one outlined below, a theory of movement and a theory of binding, we have very few choices about the possible representations associated with a particular sentence. Turning to examples like (1), the theory makes strong predictions about their representation. The theory of the relationship between argument structure and the syntax force us to hypothesize an empty category inside the nominal. The Binding Theory and θ-theory force the empty category to be a pure anaphor. Since pure anaphors must be A-bound from a $\overline{\theta}$-position, we must assume that these nominals contain a subject. Finally, the pure anaphor inside these nominals must be related to the empty category in the subject position of NP either by movement or a coindexation process that mimics the effect of movement. The representation of (1a) is (1a´):

(1)a´. $[_{S}\ [_{N''}\ \text{My room}]_{i}$ needs $[_{N''}\ \text{PRO}_{i}$ a thorough picking $\underline{t}_{i}$ up]]]

Since the NP-internal movement is similar to passivization, without any overt marking (such as a post-posed <u>by</u> phrase), I will refer to this construction as a retroactive nominal.[1]

---

[1] The term "retroactive" is taken from Jespersen (1940) who applied it to certain infinitives. Hantson (1984) uses the term with reference to gerunds. Hantson arrives at an analysis of a gerundive construction that is parallel to the nominal construction examined here. His analysis, insofar as it involves passivization, is similar to my analysis of the nominal construction.

2.1 Some Preliminaries

The Projection Principle of Chomsky (1981) stands as a fundamental breakthrough in the study of natural language. The principle, itself, has a very simple formulation:

(2) Lexical features must be represented at every syntactic level.

Despite its simplicity, the Projection Principle has far-reaching effects. For example, it provides the foundation of the theory of empty categories and can be taken as a basic methodological guideline that has shaped much of the research carried out in Government Binding theory since its inception.

By way of illustration, consider the problem of how to represent wh-questions in English:

(3) Who did John see?

(4)a. $[_{\bar{S}}$ who$_i$ $[_S$ did John see]]

b. $[_{\bar{S}}$ who$_i$ $[_S$ did John see $[_{NP}$ e$_i$]]

The simplified labelled bracketings in (4) represent two possible S-Structures for the sentence in (3). In (4a), the verb see is not followed by an NP; despite the fact that see is a transitive verb, (4a) represents it as an intransitive. By contrast, (4b) preserves the information that see is a transitive verb, since it is followed by an NP; this NP, however, lacks phonological content. The bracketing in (4a) violates the Projection Principle because a lexical feature of see, namely its transitive

subcategorization feature, is not satisfied by an NP in the syntactic representation. Thus, the Projection Principle unambiguously forces us to select (4b) as the only possible representation of (3) since it is only in (4b) is the subcategorization of see satisfied.

The assumption that the representation in (4b) is basically correct makes it possible to account for a number of (apparently) unrelated facts:

(5)a. *who$_i$ did he$_i$ see e$_i$

b. *who$_i$ did you think that e$_i$ left

In (5a), the pronoun he binds[2] the empty category left by wh-movement; if we assume that the trace left by wh-movement must be locally bound, then we can easily account for the ungrammaticality of (5a). If we assume further that such an empty category must be locally identified and that the complementizer that is not sufficient to identify the empty category in the subject position of (5b), then we can also easily provide an account of the ungrammaticality of (5b).

The representations in (4b) and (5) contain the empty string. Prima facie, allowing the empty string to occur in syntactic representations appears to invite disaster, since it

---

2 An element is bound if it is c-commanded by a coindexed category; A is locally bound by B if A is bound by B and there is no element C such that C binds A but does not bind B. On the definition of c-command and government, see Aoun & Sportiche (1983).

arms us with a very powerful notational device. For example, given that the empty string may satisfy lexical requirements, what prevents us from predicting the following sentences to be grammatical:

(6)a. *[$_{NP}$ e] saw him

b. *he saw [$_{NP}$ e]

How is it possible for a child acquiring English to learn, on the basis of positive evidence, that examples like (6) are ungrammatical?[3] This phenomenon is especially curious when it is observed that examples parallel to (4) are perfectly grammatical in Chinese (Huang, 1984):

(7)a. [$_{NP}$ e] kanjian ta le.

´[He] saw him´

b. ta kanjian [$_{NP}$ e] le.

´He saw [him]´

Since, by hypothesis, Chinese and English both obey the core principles of Universal Grammar and since the child cannot reasonably be assumed to have much in the way of direct evidence

---

3 See is taken, here, in its transitive sense only. Notice that it is insufficient to say that the child fails to construct well-formed representations for structures like (6a) simply because such examples are not included in the child´s primary linguistic data. The child does, in fact, have evidence of phonetically null subjects of tensed sentences in relative clause structures: "the man that John said e saw him". The question that needs to be answered is why the child acquiring English fails to generalize from relative clause structures to tensed sentences in general. This question is of even greater interest given the data from Chinese (below).

about things that he or she cannot hear in the speech signal, it would appear that the assumption that syntactic representations contain empty categories is rather difficult to maintain. It is exactly this tension between basic principles of grammar and apparent counter-examples that has forced the kind of careful research that is a prerequisite for explanatory adequacy.

Let us assume, then, that the theory of empty categories and the Projection Principle are correct in their essential details. Furthermore, let us assume, following Chomsky (1982), that the Binding Theory partitions the class of NPs into four categories. An NP, like <u>himself</u>, is an anaphor which must be bound from an argument position contained within some local domain. A pronoun, like <u>she</u> or <u>her</u>, may not be bound from an argument position contained within some local domain. A name, like <u>John</u>, has inherent reference and, so, may not be bound from an argument position. These differential properties of NPs are illustrated in (8-10):

```
(8)a.  John  saw himself  in the mirror
          i            i
   b.  *John  said [  Mary saw himself  in the mirror]
          i        S                  i
(9)a.  *John  saw him  in the mirror
          i          i
   b.   John  said [  Mary saw him  in the mirror]
          i        S              i
(10)a. *He  saw John  in the mirror
        i          i
    b. *He  said [  Mary saw John  in the mirror]
        i        S              i
```

The differential behavior of NPs with respect to binding suggests the following set of features:

(11) [+anaphor, -pronominal]
[-anaphor, +pronominal]
[+anaphor, +pronominal]
[-anaphor, -pronominal]

The feature system in (11) predicts the presence of four types of NPs, but, examples (8-10) illustrate only three types of overt NPs. Consider the properties of an NP marked [+anaphor, +pronominal]. By virtue of being marked [+anaphor], such an NP must be bound from an argument position in its local domain. By virtue of being marked [+pronominal], it may not be bound in its local domain. Since no element can be both bound and free in any local domain, and since the definition of local domain crucially relies on government, it follows that such an element may not be governed. A minimal requirement on phonetically overt NPs is that they bear abstract Case. Since assignment of abstract Case is contingent on government, we may assume that an ungoverned element will not receive Case. Hence, phonetically overt NPs may not bear the features [+anaphor, +pronominal].[4] Notice that the requirement that an NP bear Case does not hold for non-overt NPs. Hence, the theory allows an empty category to bear the features [+anaphor, +pronominal] just so long as that empty category appears in an ungoverned position. Like an anaphor, this empty

---

4 For discussion on the definition of local domain and government, see Chomsky (1981) and the references cited there. If we assume that genitive Case is assigned configurationally rather than on the basis of government then, a priori, one might expect an overt NP with the features [+anaphor, +pronominal] could occur in this position. Some proposals made in Huang (1982) and Aoun (forthcoming) may circumvent this problem.

category could be obligatorily bound in some configurations and, like a pronominal, this empty category has the option being free in other configurations. Such an empty category is found in the subject position of infinitives and gerunds, as illustrated in (12-13):[5]

(12)a. John wanted [PRO to shave himself]
$_i$ $_S$ $_i$ $_i$

b. *John wanted [PRO to shave oneself]
$_i$ $_S$

(13)a. It is obvious how [PRO to boil water]
$_S$

b. [PRO boiling water] is trivial
$_S$

Again following Chomsky (1982), let us assume that the null hypothesis is that the properties of empty categories are not inherent, but can be determined given the properties of the structures in which the empty categories occur.[6] Specifically, let us assume that there is only one type of empty category, unmarked for the features [± pronominal, ± anaphor]. The values for these features can be filled in given information about whether the empty category is bound or free and, if the empty

---

5 For a rather different approach to PRO, see Bouchard (1984) and Sportiche (1983) where it is argued that PRO may be either a pure anaphor or a pure pronominal but not [+anaphor, +pronominal]. See below for arguments that the distribution of PRO is not limited to the subject position of infinitives and gerunds.

6 Recent research (e.g., Brody (1984) and Safir (1984b) indicates that empty categories do have inherent properties that cannot be determined on the basis of structural properties alone. Nevertheless, since the basic points that interest us here will not be significantly altered, I will assume that the functional definition of empty categories provides a set of diagnostics for determining the inherent features of the empty category in question.

category is bound, information about the nature of the position occupied by the binder. We will say that an empty category bears the features [-pronominal, -anaphor] if it appears in an A-position and if its antecedent appears in an $\bar{A}$-position. Otherwise, an empty category in an A-position is marked [+anaphor]. If the empty category is either free or bound by an element with an independent θ-role, then it is also marked [+pronominal]; otherwise, it bears the feature [-pronominal].[7]

(14) [+anaphor, -pronominal] = NP-trace[8]
[-anaphor, +pronominal] = pro
[+anaphor, +pronominal] = PRO
[-anaphor, -pronominal] = variable

2.2 Some properties of retroactive nominals

We are now in a position to apply the above diagnostics to the following set of examples:

(15)a. John could use [a good talking to]$_{N''}$

b. These ideas merit [some working on]$_{N''}$

c. This problem bears [a good deal of thinking about]$_{N''}$

d. My room needs [a thorough picking up]$_{N''}$

Each of the above sentences contains a derived nominal in object

---

7 An element is in an A(rgument)-position if that position is a subject or receives a θ-role; otherwise, it is in a non-argument (A-bar) position.

8 For the moment, I will exclude pro, the empty pure pronominal, from the discussion.

position and each of the derived nominals contains a stranded preposition. From the Projection Principle, we can derive the fact that the prepositions must have empty nominal complements. Hence, the sentences in (15) could be more accurately represented as in (16):

(16) a. John could use [$_{N''}$ a good talking to [e]]

b. These ideas merit [$_{N''}$ some working on [e]]

c. This problem bears [$_{N''}$ a good deal of thinking about [e]]

d. My room needs [$_{N''}$ a thorough picking up [e]]

Intuitively, the empty categories in the above sentences are A-bound from the subject position of the clause that contains them. By applying the functional determination of empty categories, we derive the marking [+anaphor] for each of the empty categories. Let us assume, for the moment, that the empty category in these structures is a pure anaphor. Minimally, then, the empty category in this structure must comply with condition A of the Binding Theory. In order to test this hypothesis, we need only find a structure in which the subject position of the sentence is outside of the governing category of the putative anaphor. If there is a SUBJECT accessible[9] to the anaphor within

---

9 An element A is the SUBJECT of a projection X if A is the most prominent nominal element within X. A is accessible to B only if coindexing A with B does not violate the i-within-i condition. For discussion, see Chomsky (1981) and the

the NP, then it should be the case that the anaphor will not be bound by the matrix subject:

(17)a. *John$_i$ could use [$_{N''}$ [$_{N''}$ a competent psychiatrist's] talking to [e]$_i$]

b. *These ideas$_i$ merit [$_{N''}$ [$_{N''}$ Bill's] working on [e]$_i$]

c. *This problem$_i$ bears [$_{N''}$ [$_{N''}$ everybody's] thinking about [e]$_i$]

d. *My room$_i$ needs [$_{N''}$ [$_{N''}$ the janitor's] picking up [e]$_i$]

As the examples in (17) show, it is impossible for the matrix subject to act as the antecedent for the empty category inside the nominal if the nominal contains a SUBJECT accessible to the anaphor. The ungrammaticality of the examples cannot be attributed to a semantic restriction on the form of the NP in these constructions, since, although the examples are ungrammatical with an agent in the subject position of the NP, an agent may appear in a postposed <u>by</u> phrase, as illustrated in (18):

(18)a. John$_i$ could use [$_{N''}$ a good talking to [e]$_i$ by a competent psychiatrist]

b. These ideas$_i$ merit [$_{N''}$ some working on [e]$_i$ by Bill]

---

references cited there.

c. This problem$_i$ bears [$_{N''}$ some thinking about [e]$_i$ by everybody in the class]

d. My room$_i$ needs [$_{N''}$ a thorough picking up [e]$_i$ by the janitor]

The constraint that the NP in these constructions not contain a SUBJECT follows from a purely syntactic constraint on binding relations and not from a constraint on the semantic representation of this construction.[10]

Having seen some evidence that the empty category in this construction is anaphoric, it is now necessary to try to determine the value for the feature [±pronominal]. Recall that the empty category is assigned [+pronominal] if it is free or A-bound by an element with an independent θ-role. The empty category in this construction is interpreted as bound by the subject of S, so we can rule out the case of the empty category being free. The question, therefore, reduces to whether or not the subject of S receives a θ-role independently of the empty category inside the nominal; in other words, we must determine whether the subject of S and the empty category form an A-chain.[11] If the subject position of S is not assigned a θ-role independently of the empty category, then it should be the case

---

10 Judging from the Specified Subject Condition effect noted in the text it appears that these nominals appear to be transparent to binding, like picture NPs.

11 On A-chains, see Chomsky (1981) and Safir (1985).

that a pleonastic element can occupy this position:

(19)a. *It could use [a good talking to John]
N"

b. *It merits [some working on these ideas]
N"

c. *It bears [some thinking about this problem]
N"

d. *It needs [a thorough picking up of my room]
N"

Since pleonastic <u>it</u> is barred from the subject position in this construction, the conclusion that this position is assigned a θ-role that is independent from the one assigned to the empty category in the NP appears to be warranted. If this conclusion is correct, then the value for the pronominal feature is positive. The functional determination of empty categories, therefore, leads us to the conclusion that the empty category is [+anaphor, +pronominal], or PRO.

Identifying the empty category with PRO has two important consequences. First, we would expect that the position that it occupies is not governed. Recall that any element marked [+anaphor, +pronominal] must obey both conditions A and B of the binding, which is a contradiction. In order to avoid this contradiction, any element with such a feature specification must occur in an ungoverned position. Second, it should be possible to find environments in which the empty category receives an arbitrary interpretation; namely, there should be structures where there is no possible antecedent for the PRO inside these nominals. If the position inside the NP is, in fact, ungoverned, then placing the NP in the subject position of an intransitive

sentence should be sufficient to induce an arbitrary interpretation:

(20)a. *[$_{N''}$ a good talking to PRO] is often necessary

b. *[$_{N''}$ some thinking about PRO] could be useful

c. *[$_{N''}$ a lot of working on PRO] would be good

As the examples in (20) demonstrate, an arbitrary interpretation for the empty category in these nominals is impossible. It should be noted that the empty category in many of these examples stands in a government configuration with a preposition, so a redefinition of government to ensure that a local configuration of government is sensitive to elements outside of this local configuration; this result is not obviously desirable given that it is motivated solely on the basis of these examples. Apparently, we must reject the hypothesis that the empty category in these constructions is instantiated by PRO.

We have seen that the assumption that the empty category in this construction is marked [+anaphor, +pronominal] has led to a number of difficulties; namely, this empty category stands in a government configuration with some head and, unlike PRO, this empty category may not receive an arbitrary interpretation. Let us assume, despite what the functional definition of empty categories has led us to believe, that the empty category in question is a pure anaphor, [+anaphor, -pronominal] (i.e., NP-trace). NP-trace has the property of being locally A-bound by an element that does not receive an independent θ-role; since its

binder does not receive a θ-role in situ, it must form a chain with the trace and "share" its thematic role. This fact is quite apparent is we consider the structure of raising constructions:

(21)a. It seems that [John has left]
$_S$

b. John$_i$ seems [$_S$ $\bar{e}_i$ to have left]

The pleonastic element, <u>it</u>, occupies the matrix subject position is (21a), indicating that this position does not receive a θ-role from the VP, <u>seems that John has left</u>. Hence, the matrix subject position in raising constructions must be a $\bar{\theta}$-position. In (21b), on the other hand, a referential NP, <u>John</u>, occupies the matrix subject position. Since, as we have seen, the subject position is a $\bar{\theta}$-position, a referential subject must receive its θ-role by virtue of being linked to a θ-position. In order for the representation in (21b) to be well-formed, the matrix subject must form a chain with the empty category in the subject position of the embedded clause. Let us assume, then, that NP-trace always forms a chain with the argument that binds it. Given this thematic relation between the NP-trace and its binder, let us reconsider the examples in (19), repeated here as (22):

(22)a. *It could use [a good talking to John]
$_{N''}$

b. *It merits [some working on these ideas]
$_{N''}$

c. *It bears [some thinking about this problem]
$_{N''}$

d. *It needs [a thorough picking up of my room]
$_{N''}$

These examples led us to the conclusion that the subject position

these sentences is a θ-position. But if the subject position receives a θ-role independently of the empty category, then the two positions cannot form a chain, since such a chain would bear two θ-roles, which is impossible. Since NP-trace inevitably enters into a chain relation with some other position, it must be the case that the empty category is not an NP-trace.

We have examined two types of empty categories, PRO and NP-trace, and we have discovered that the empty category in this construction cannot be identified with either of them. Since PRO and NP-trace form a natural class with respect to the feature system in (11)--the class of nominal elements selected by the feature [+anaphor]--let us assume that this empty category is inherently marked [-anaphor]. The class of non-anaphoric empty categories is exhaustively made up by pro ([-anaphor, +pronominal]) and variable ([-anaphor, -pronominal]).

Suppose that the empty category is a pure pronominal. Thus, the structure in question would have the following representation:

(23)a. John_i could use [a good talking to $\underline{\text{pro}}_i$]$_{N''}$

b. These ideas_i merit [some working on $\underline{\text{pro}}_i$]$_{N''}$

c. This problem_i bears [a good deal of thinking about $\underline{\text{pro}}_i$]$_{N''}$

d. My room_i needs [a thorough picking up $\underline{\text{pro}}_i$]$_{N''}$

As a pure pronominal, it must obey condition B of the binding theory and, so, it cannot be bound in its governing category.

Recall however that, as the examples in (17) show, placing a subject inside the post-verbal NP in this construction invokes a violation of the Specified Subject Condition (SSC). These SSC violations indicate that the empty category must be bound in its governing category. If the empty category must find its antecedent within its local domain, then we cannot identify it with a pure pronominal. Let us adopt the reasonable assumption that pro must be locally identified.[12] The empty category in these examples has no local identifier. We must assume that this empty category is not pro.

2.3 A plausible alternative

We have seen that the empty category in examples like (16) is not pronominal and, for a variety of reasons, cannot be identified with an anaphor. Our remaining option is that the empty category is marked [-anaphor, -pronominal] (i.e., that it is a variable). Our rationale for rejecting this hypothesis was that the empty category was not $\bar{A}$-bound; this line of reasoning is difficult to maintain in light of examples like (24) and (25), where an empty category is $\bar{A}$-bound by a non-overt operator:

---

12 Chomsky (1982) assumes that pro may be identified by a sufficiently rich inflection or by a clitic. Huang (1984) hypothesizes that, in some languages, pro may also be identified by means of coindexation with a topic. English, however, lacks a sufficiently rich inflection, has no clitics, and does not allow pro to be coindexed with a topic; hence, it is commonly assumed that English lacks pro.

(24)a. John_i is easy $[_{\bar{S}}$ Op_i $[_S$ PRO to dislike $\underline{t}_i$]]

b. This lecture_i is hard $[_{\bar{S}}$ Op_i $[_S$ PRO to follow $\underline{t}_i$]]

(25)a. John bought the book $[_{\bar{S}}$ Op_i $[_S$ to read $\underline{t}_i$ to the children]]

b. A shovel_i is $[_{\bar{S}}$ Op_i $[_S$ PRO to dig with $\underline{t}_i$]]

In each of the above examples, an empty category is bound by a phonetically non-overt operator. Notice that, as argued in Chomsky (1982), these constructions allow parasitic gaps:

(26) This course_i is hard to follow $\underline{t}_i$ without having done the readings for $\underline{e}_i$

If it is true that parasitic gaps may be licensed only be elements in an $\bar{A}$-position, then examples like (26) are well-formed only if there is a non-overt operator to bind the the trace in the object position of the embedded clause.

An a priori plausible objection to the hypothesis that this empty category is a variable would be that the empty category is A-bound by the subject of the sentence. If the empty category is, in fact, A-bound by the subject, then the configuration violates condition C of the Binding theory. Given the existence of examples like (24-26), the condition that variables must be A-free must be weakened to the condition that variables must be A-free in the domain of the operator that binds them.[13] Hence,

[13] Strong cross-over effects, under this view, are limited to

this objection is ultimately untenable. Let us assume that the construction has the following representation:

(27) This problem$_i$ bears [Op$_i$ [some thinking about $\underline{t}_i$]]
N" N"

where Op is a phonetically null operator that binds a variable.[14]

If this approach to the problem is correct, then the construction bears a superficial resemblance to "tough" movement

---

the domain of an operator. If this approach is correct, then the ungrammaticality of (i) does not follow from the fact that the variable is bound by an A-position:

(i) *John$_i$ wonders [who$_i$ [$\underline{t}_i$ saw Mary]]

since the matrix subject position is outside the domain of the operator, who. This is the approach was first hypothesized in Chomsky (1981) and is developed in Aoun & Clark (1985), where it is argued that the ungrammaticality of (i) follows from the fact that the operator, which has an inherent range and may not be referentially dependent, is bound. Unlike overt operators, non-overt operators have no inherent range and, therefore, must be referentially dependent; this can be viewed as an identification requirement on non-overt operators. Compare the examples in (i) with those of (ii):

(ii)a. John$_i$ is easy [Op$_i$ [PRO to please $\underline{t}_i$]]

b. A shovel$_i$ is [Op$_i$ [PRO to dig with $\underline{t}_i$]]

where, in some sense, the subject delimits the range of the non-overt operator. The requirement that non-overt operators be identified is, I believe, related to the requirement that variables be "strongly bound" (in the terms of Chomsky, ms.). See Aoun & Clark (1985) for discussion.

14 A simple way to test the hypothesis that a non-overt operator is involved is to look for violations of weak cross-over. Here, the facts become rather complex; see below.

constructions as illustrated in (28):

(28)a. John_i is easy $[_{\bar{S}}\ \text{Op}_i\ [_S$ PRO to please $\underline{t}_i$]]

b. John_i could use $[_{N''}\ \text{Op}_i\ [_{N''}$ a good talking to $\underline{t}_i$]]

Like "tough" movement, the retroactive nominal construction is lexically governed; that is, the adjectives and verbs that may govern a retrograde nominal fall into a semantic class, which we will term "predicates of requirement". This is illustrated by the examples in (29):

(29)a. *John likes some talking to
b. *My room looks a little cleaning up
c. *This problem appears some thinking about

Some examples of the verbals that fall into this class are <u>need</u>, <u>bear</u>, <u>merit</u>, <u>could</u> <u>use</u>, and <u>could</u> <u>stand</u>. Given the apparent similarity between the two constructions, it is plausible that retroactive nominals are a sort of NP version of "tough" movement. Let us assume that NPs may have an internal Comp position.[15] Since this construction is lexically governed, it is

---

15 Unlike its clausal counterpart, this NP-internal Comp is not the head of a category. The assumption that NPs may contain an internal Comp presents something of a problem if we assume, first, that subcategorized positions must be governed and, second, that the head of a maximal projection may be governed from outside of that maximal projection. Under this view, the NP-internal Comp is not governed by the element that subcategorizes for it. Let us assume, for the sake of the discussion, that this is more a problem of execution than of insight. The hypothesis that NPs contain an internal Comp position might provide an account for NPs of the form:

(i)a. $[_{N''}$ how long a book] did Mary write?

b. Mary had never seen $[_{N''}$ so long a book].

reasonable to assume that verbs may subcategorize for this Comp position in much the same way that verbs may subcategorize for a sentential Comp position, as illustrated in (30):

(30)a. I wonder $[_{\bar{S}}$ $[_{\text{Comp}[+WH]}$ who$_i$] $[_S$ John saw $\underline{t}_i$]]

b. *I wonder $[_{\bar{S}}$ $[_{\text{Comp}[-WH]}$ that] $[_S$ John saw Mary]]

The verb, wonder, subcategorizes for a [+WH] Comp; since the Comp in (30a) contains a WH-word, it satisfies wonder's lexical requirements. The Comp in (30b) does not contain a WH-element and is, therefore, marked [-WH], which cannot satisfy the lexical requirements of the governing verb. Hence, (30b) is ungrammatical.

On this approach, a retroactive nominal would have a D Structure like (31):

(31) John needs $[_{N''}$ $[_{\text{Comp}}$ $\underline{e}$] a good talking to $\underline{Op}$]

WH-movement would apply to the non-overt operator to yield:

(32) John needs $[_{N''}$ $[_{\text{Comp}}$ $\underline{Op}_i$] a good talking to $\underline{t}_i$]

The non-overt operator would then be coindexed with the matrix subject by predication.

By hypothesis, an NP-internal $\bar{A}$-position is a marked structure that must be subcategorized for by a lexical item that governs the NP. In the unmarked case, the structure in (33) is impossible:

(33) *$[_{N''}$ ...$[_{\text{Comp}}$ $\underline{e}$]...]

Consider the sentences in (34):

(34)a. *some thinking about is often necessary
b. *a little working on would be useful
c. *looking into appears necessary

The examples in (34) indicate that retroactive nominals exhibit a subject/object asymmetry. The subject NPs in the examples of (34) are governed by Infl rather than a lexical element. The NP-internal Comp is, as a result, unavailable to these NPs. The D Structure of (34a), for example, is (35):

(35) [$_{N''}$ some thinking about Op] is often necessary

In order for the non-overt operator to move to a non-argument position, it must move to the matrix Comp. But this movement violates subjacency (or the Left Branch Condition) and is, therefore, impossible. The non-overt operator must remain in situ. Suppose, following Aoun & Clark (1985), that non-overt operators must have their range identified at S Structure. If so, then examples like (35) will be filtered because the non-overt operator lacks a range.

2.4 Some inadequacies

On the basis of the Binding theory and the θ-criterion, we have be forced to reject the hypothesis that the empty category in the retroactive nominal construction is a pure anaphor. On the one hand, the empty category could not be PRO because the position is governed. On the other hand, the empty category could not be a pure anaphor (NP-trace) because of the θ-criterion. Finally, the assumption that the empty category is a

pure pronominal could not be maintained because such a pronominal would violate condition B of the Binding theory and the condition that empty pronominals be identified. The only alternative that we are left with is that the empty category is a variable which is locally $\bar{A}$-bound by a non-overt operator. In this section, we will see that this hypothesis is not without its problems.

We have assumed that the non-overt operator is generated in an A-position within the nominal and is moved to an NP-internal Comp in the syntax. The operator is then associated with the matrix subject by predication. This analysis makes two rather clear predictions. First, if the non-overt operator undergoes WH-movement in the syntax, then it should be possible to move the operator successive cyclically to yield apparently unbounded movement. Second, if the non-overt operator does, in fact, occupy an $\bar{A}$-position, then it should license parasitic gaps. The ability to move successive cyclically, licensing of parasitic gaps and weak cross-over have, over the years, come to serve as diagnostics for operator-variable relations.[16]

---

16 The assumption that retroactive nominals involve an operator predicts that it should be subject to weak cross-over. Let us assume that weak cross-over involves the binding of two A-positions by a single operator. Example (i) illustrates this configuration:

(i) *who$_i$ does his$_i$ mother love $\underline{t}_i$

In example (i), the operator, who, locally binds the genitive pronoun, his, and a trace. If we assume that local A-bar-binding is a bijective relation, then we will correctly rule out (i).

Recall that, for possibly independent reasons, the

Turning first to the question of successive cyclicity, it should be adequate to test nominalizations of verbs that take sentential complements. If the empty category is associated with a non-overt operator, it should be possible to move the operator successive cyclically, leaving a gap in the embedded clause. Thus, prima facie, we would expect examples like those in (36):

---

retroactive nominal construction shows a strong Specified Subject Condition effect; thus, the environment where the nominal contains a subject which, in turn, contains a pronominal is ruled out for reasons which are independent of cross-over. Furthermore, a pronominal immediately contained in the NP will be in the same governing category as the subject of S, which is, itself, coindexed with the empty category in the nominal; hence, this configuration is already ruled out be condition B of the Binding theory. Finally, a pronominal contained in a restrictive relative will be locally bound by the trace left by the non-overt operator and, as a result, is irrelevant for weak cross-over since the local binding relation will still be bijective. The remaining possibilities are a pronominal contained in a non-restrictive relative or an adjunct prepositional phrases. Since weak cross-over effects tend to become more obscure the more deeply embedded the pronoun is, let us focus our attention on prepositional phrases.

The facts here are very difficult. Suppose we assume that an element c-commands throughout the maximal projection that contains it. Since NPs do not contain an internal maximal projection, the trace left by the non-overt operator will c-command a pronoun inside of an adjunct PP; hence, the pronoun will be locally bound by the trace of the non-overt operator and not be the operator itself. The following examples are, therefore, grammatical even under the assumption that a non-overt operator is involved in these constructions:

(ii)a. John$_i$ could always use some comforting $\underline{t}_i$ after his$_i$ sessions with the psychiatrist.

b. John$_i$ could use a good talking to $\underline{t}_i$ about his$_i$ many annoying habits.

(36)a. *John's paper could use some convincing (of) people that they should read [e]

b. *My office needs some conning (of) the janitor to clean [e] up

c. *This book deserves some persuading (of) the students to discuss [e] over beer.

The object nominal in example (36a) has the S Structure:

(37) $[_{N''}\ [_{Comp}\ Op_i\ ]$ some convincing (of) people $[_{\bar{S}}\ t_i$ that $[_S$ they should read $t'_i\ ]]]$

The non-overt operator first moves into the Comp of the embedded $\bar{S}$. Next, the non-overt operator moves into the NP-internal Comp. Since both steps obey the subjacency condition, derivations such as the one sketched for (37) should be ruled in since no first principles are violated; in order to prevent sentences such as those in (36), we will be forced to place an otherwise unwarranted constraint on successive cyclic movement. This result is undesirable in a theory which seeks to give an explanatory account of language acquisition since the child would be forced to acquire such a restriction on the basis of virtually non-existent evidence, given the limited distribution of retroactive nominals.

The second prediction made by this analysis is that the non-overt operator should license parasitic gaps. Parasitic gaps occur in the environment:

(38) $\ldots Op_i \ldots t_i \ldots e_i \ldots$

where t_i is the trace left by movement of the operator, Op_i, and e_i is a base-generated gap. In order for the base-generated gap

to be licit, it must be bound by the operator; furthermore, if we adopt the analysis of Chomsky (1982), the trace, $\underline{t}_i$, may not c-command the parasitic gap.

The environment in (38) can be trivially instantiated in structures where the parasitic gap, $\underline{e}_i$, is contained in an adjunct. This sort of structure is available in NPs if the NP is headed by a derived nominal:

(39)a. The burning of the tenement after the insurance money had been collected was in very poor taste.
b. The invasion of the country without informing the media was considered a major scandal.
c. Any torturing of the prisoners without getting the president's permission is frowned upon.

Retroactive nominals are equally capable of taking adjuncts:

(40)a. This problem deserves some thinking about without losing any sleep over it.
b. John could use a good talking to without treating him like a child.
c. This novel merits rewriting without changing it too much.

If the structure of retroactive nominals actually involves a non-overt operator, then we would expect the examples in (40) to have counterparts containing parasitic gaps:

(41) *This problem deserves some thinking about <u>t</u> without losing any sleep over <u>e</u>

(42)a. *John could use a good talking to <u>t</u> without scolding <u>e</u>
b. John could use a good talking to <u>t</u> without PRO being scolded

(43)a. *This novel merits some rewriting <u>t</u> without changing <u>e</u> too much
b. This novel merits some rewriting <u>t</u> without PRO being changed too much

Insofar as parasitic gaps are impossible in the retroactive

nominal construction, the hypothesis that this construction involves $\bar{A}$-binding is not supported.

There is a further inadequacy with the approach that assumes that the empty category in retroactive nominal constructions is a variable; namely, the fact that these NPs may not contain an overt subject. The relevant facts are repeated in (44):

(44)a. *John$_i$ could use [$_{N''}$ Op$_i$ [$_{N''}$ a competent psychiatrist's] talking to [e$_i$]]

b. *These ideas$_i$ merit [$_{N''}$ Op$_i$ [$_{N''}$ Bill's] working on [e$_i$]]

c. *This problem$_i$ bears [$_{N''}$ Op$_i$ [$_{N''}$ everybody's] thinking about [e$_i$]]

d. *My room$_i$ needs [$_{N''}$ Op$_i$ [$_{N''}$ the janitor's] picking [e$_i$] up]

If the retroactive nominal construction actually involves an NP-internal Comp position, then it would appear that this Comp is in complementary distribution with the subject of the NP. This apparent generalization does not follow from any principles, however. As in the case of successive cyclic movement (above), a special stipulation must be made in order for the analysis to be descriptively adequate.[17]

---

17 The facts at hand bear a striking resemblance to the construction:

The empirical inadequacies of the approach outlined in the previous section imply that a different approach to this construction needs to be taken. The evidence discussed in the first section that the empty category is an anaphor was abandoned largely on the basis of the θ-criterion. If the empty category is an anaphor, then it must form a chain with the matrix subject position since an NP-trace may not head a chain; this chain, however, would be doubly θ-marked. If the empty category forms a chain with some element other than matrix subject position, this problem may not arise. Since PRO, unlike NP trace, may head a chain, let us assume that the empty category is linked with a PRO inside the NP. If this is indeed the case, the anaphoric properties of the empty category is not inconsistent with thematic theory. Let us turn our attention to some properties of

---

(i) This painting is lovely to look at

as discussed in, for example, Chomsky (1977). Like retroactive nominalization, this construction is bounded:

(ii) *this painting is lovely to persuade people to look at

and the embedded clause may not have an overt subject:

(iii) *this painting is lovely for John to look at

Notice that this construction, unlike the nominals in question, allows for parasitic gaps:

(iv) This kind of painting is nice to look at t without thinking about e

I will therefore assume that the above construction has properties that are markedly different from those of retroactive nominals.

θ-marking and government within NPs.

2.5 "Move NP" inside NP

Since at least Lees (1960), certain parallels in structure between clauses and NPs have been observed. This parallelism between S and NP was one of the factors that motivated the introduction of $\overline{X}$-theory (Chomsky, 1970). Subsequent research (for example, Jackendoff (1977) and Stowell (1983)) has supported the assumption that close structural parallelisms exists across all grammatical categories.

One important set of data that has supported the assumption that S and NP have a similar syntactic structure is that verbs which allow passive often have related derived nominals which also have a passive-like form:

(45)a. John narrated the play.
b. The play was criticized by John.

(46)a. John's narration of the play...
b. The play's narration by John...

(47)a. John criticized the book.
b. The book was criticized by John.

(48)a. John's criticism of the book...
b. The book's criticism by John...

We can account for sentential passives (examples (45b) and (47b)) if we assume that the passive morpheme absorbs the Case assigned by the verb to the object and the θ-role assigned to the subject.[18] Since the subject position is a $\overline{\theta}$-position, a general

[18] We may assume, following Keyser & Roeper (1984), that the absorption of accusative Case and the blocking of the subject's θ-role are not, in fact, two independent operations, but, rather,

rule, "move alpha", may move the object NP to that position; this movement is, in fact, forced by the Case filter, given that the object position does not receive Case from the verb.

Analogous reasoning may be applied to the "passive nominals" in examples (46) and (48). Since a noun may not assign Case, we need not consider the Case filter; we must assume, however, that an NP has a subject position and that this subject position is a $\overline{\theta}$-position. Given these assumptions, the object of a derived nominal may optionally move into the subject position of the NP.

In light of these facts, we will assume that there are structural similarities across categories and that sub-theories of grammar are, in the main, indifferent to grammatical category; in other words, "move alpha", for example, will behave uniformly across categories.[19] Given this type of approach, particular gaps in paradigms must be derived from independent sub-theories of grammar. In order to make the discussion more concrete, let us

---

they follow from Burzio's generalization to the effect that the subject position does not receive a θ-role if accusative Case is not assigned to the object.

19 This should not be taken to mean that grammatical categories are irrelevant to all the various sub-theories of the grammar. Case theory, for instance, makes crucial reference to the features [+N, +V]. The hypothesis is that reference to grammatical category should be eliminated where possible. This hypothesis, then, calls into question whether categories like NP or S need to be mentioned in the definition of subjacency and whether the definition of proper government should stipulate the category of proper governors. Our hope is that these (apparently) unrelated stipulations will be derived from some more general principle of core grammar.

return to "passive nominals". As pointed out above, verbs that allow the passive construction often have derived nominals that have a "passive" form. This parallelism breaks down systematically, as shown in the following examples:[20]

(49)a. Many people rely on Mary.
b. Mary is relied on by many people.

(50)a. The reliance on Mary by many people (is unwise)
b. *Mary's reliance on by many people (is unwise)

(51)a. Ignorant people laughed at Galileo's ideas.
b. Galileo's ideas were laughed at by ignorant people.

(52)a. the laughter at Galileo's ideas by ignorant people (annoyed the more literate elements of the population)
b. *Galileo's ideas laughter at by ignorant people (annoyed the more literate elements of the population)

(53)a. Washington slept in this bed.
b. This bed was slept in by Washington.

(54)a. Any sleeping in this bed by Washington (will be considered an act of treachery)
b. *this bed's sleeping in by Washington (will be considered an act of treachery)

(55)a. Congress looked into this problem.
b. This problem was looked into by Congress.

(56)a. Any looking into this problem by Congress (will be considered a threat to the national security)
b. *this problem's looking into by Congress (will be considered a threat to the national security)

The verbs in examples (49), (51), (53) and (55) all allow pseudo-

---

20 The systematic gaps in nominal paradigms discussed here and below have received attention in a number of places; see, for example, Chomsky (1970), Postal (1974), Williams (1982) and Kayne (1984).

passive forms where the object of a preposition may be preposed to subject position. The related derived nominals in (50), (52), (54) and (56), however, do not allow the prepositional object to be preposed. Thus, there is a systematic gap in the pseudo-passive paradigm. If the hypothesis that sub-systems of core grammar are largely category neutral is correct, then we must search for some more abstract principle(s) from which we can derive the fact that derived nominals do not permit pseudo-passive forms.

The pseudo-passive construction is hardly the only construction that, while present with verb forms, is systematically missing from NPs. Kayne (1984) provides a sizable collection of such constructions. To take a few examples, dative shift, small clauses, tough movement, exceptional Case marking and raising are all absent from derived nominals, although the related verbal or adjectival root may permit such constructions. These facts are exemplified in the following examples:

(57)a. John gave the book to Mary.
b. John gave Mary the book.

(58)a. John's gift of the book to Mary...
b. *John's gift of Mary of the book...

(59) John considers Bill a threat.

(60) *John's consideration of Bill (of) a threat...[21]

[21] But compare (60) with:

(i) the consideration of Bill as a threat

(61) John is easy to irritate.

(62) *John's easiness to irritate..

(63) John believes Bill to be incompetent.

(64) *John's belief of Bill to be incompetent...

(65) John is likely to arrive late.

(66) *John's likelihood to arrive late...

Kayne (1984) argues that the systematic lack of these constructions by NPs may be accounted for if we assume that nouns may not govern across maximal projections.[22] Suppose, for example, that tough movement involves the presence, in the embedded Comp, of a non-overt operator which must be properly governed. Example (62) would have an S Structure like (67):

(67) [[John's] [easiness [Op [PRO to irritate
N" N" $\bar{N}$ $\bar{S}$ i S
$\underline{t}_i$]]]]

Since the noun, easiness, may not govern across a maximal projection and the operator is dominated by an $\bar{S}$, the non-overt operator is not properly governed and, as a result, (67) violates

---

The addition of as to the predicate seems to improve the NP. In itself, however, this fact should not be taken as an indication that nouns admit a particular type of small clause (i.e., those including an as).

22 In order for this analysis to work out completely, we must adopt Kayne's analysis of datives. Kayne assumes that give Mary the book has the following bracketing:

(i) [give [[Mary] [the book]]]
X NP NP

i.e., that Mary the book is a small clause and that the node labelled X in (i) is a maximal projection.

the Empty Category Principle.[23] But why should it be the case that nouns are incapable of governing across a maximal projection? This stipulation about nouns is at odds with the category neutral hypothesis; optimally, we would prefer to derive the apparent fact that nouns cannot govern across a maximal projection from some component of core grammar rather than, say, by stipulation in the definition of government.

Let us assume the following feature system for grammatical categories:

(68)

N	V	
+	+	= Adjective
+	-	= Noun
-	+	= Verb
-	-	= Preposition

Notice that the inability to govern across a maximal does not correlate with the category [+N] since adjectives may so govern:

(69)a. John_i is [$_{adj}$ likely] [$_{\bar{S}}$ $\underline{t}_i$ to win]

b. John_i is [$_{adj}$ easy] [$_{\bar{S}}$ Op_i [$_S$ PRO to please $\underline{t}_i$]]

If adjectives could not govern across a maximal projection, then the trace in (69a) and the non-overt operator in (69b) would not be properly governed (or governed at all); the examples in (69) would then be ruled out by the Empty Category Principle. Since the examples are grammatical, we must assume that adjectives may

---

23 We can derive the fact that nouns lack some form of dative shift, exceptional Case marking and small clauses if "of"-insertion requires government.

govern across a maximal projection.

Following the category neutral hypothesis, we will suppose that the fact that nouns may not govern across a maximal projection follows from some other sub-component of the theory of grammar and is not simply stipulated by the definition of government. We have already ruled out Case theory as a potential candidate for determining the class of categories that may not govern across a maximal projection; neither nouns nor adjectives may assign Case, but adjectives, unlike nouns, may govern across maximal projections.[24]

If nouns were alone in their inability to govern across a maximal projection, then this inability might well be considered an aberration that is best stipulated, perhaps in the definition of government. There is, however, some evidence that prepositions are also unable to govern across a maximal projection. Let us assume that, following Stowell (1981), that in order for a non-root $\overline{S}$ to have an empty Comp, the Comp must be properly governed. It will then follow that an $\overline{S}$ that is

---

24 But see Safir (1984) where it is argued that at least one adjective, namely worth, may assign Case. Safir also points out that adjectives may assign Case in German. For more on worth, see Safir (1984), Hantson (1984) and below. Notice also that if Case assignment were crucial for government across a maximal projection, then passive participles should be incapable of governing across such a projection. This would predict that Exceptional Case Marking verbs should not allow passive when the complement clause is infinitival. This is counterexemplified by (i):

(i) John is believed to have been abducted

governed by a verb may have an empty Comp:

(70)a. John believes (that) Bill is a spy.
b. John knows (that) Mary is in Europe.

An $\overline{S}$ in subject position will not be properly governed since Infl is not a proper governor in English; this treatment predicts that $\overline{S}$'s in subject position may never have an empty Comp, which is, in fact, correct:

(71)a. *(that) Bill is a spy is obvious to John.
b. *(that) Mary is in Europe is quite likely.

In many dialects of English, the preposition <u>in</u> may be used to introduce adjunct clauses:

(72)a. In that John treated her so rudely, Mary left early.
b. You will have to take the exam later, in that you are three hours late.

If prepositions could govern across maximal projections, then we would expect that the Comp of such a clause could be empty:

(73)a. *in [<u>e</u>] John treated her so rudely, Mary left early
b. *you will have to take the exam later, in [<u>e</u>] you are three hours late.

Since the sentential object of <u>in</u> obligatorily has an overt complementizer, the conclusion is that the preposition does not properly govern the Comp position. This fact is consistent with the hypothesis that prepositions may not govern across maximal projections.[25]

---

[25] Cinque (1984) argues that parasitic gaps are impossible in Italian when the adjunct clause is introduced by the preposition <u>per</u>:

(i) *questo libro$_i$ è troppo di parte per adottare $\underline{e}_i$

Even in semantically plausible environments, prepositions may not take small clauses, as illustrated in (74):

(74)a. *Mary will always think of John an innocent fool
b. Mary will always think of John as an innocent fool.

(75)a. *we argued about John guilty
b. We argued about John's being guilty.

The absence of small clause complements of prepositions is consistent with the hypothesis that prepositions may not govern across maximal projections.[26]

It would appear, then, that nouns and prepositions form a

"which book is too biased to adopt?"

(ii) Questo libro $_i$ é troppo di parte per adottarle$_i$
"Which book is too biased to adopt it?"

If the generation of parasitic gaps involves non-overt operators, then this fact indicates that the non-overt operator is not properly governed. Notice, however, that prepositions are not proper governors in Italian independently of their inability to govern across maximal projections.

26 I have not included examples of dative shift, tough movement and raising simply because I could not find any examples that were semantically plausible. The inability to govern into Comp and the absence of small clause complements are sufficient to demonstrate the point.

Notice that these observations should be taken as applying to the core case. Tim Stowell (p.c.) points out that there is at least one preposition in English that does allow small clauses:

(i)a. With John out of the way, we can safely continue our conversation.
b. With the alarm system broken, anyone could just walk in.

My judgement is that with only takes a small clause when it occurs in an adjunct position. The relevance of this fact should be apparent shortly.

class with respect to their inability to govern across maximal projections. This class is selected by the feature [-V]. Now we can inquire as to what property, aside from sharing the feature [-V], the class of nouns and the class of prepositions have in common. Following, for example, Jaeggli (1984), let us assume that prepositions may assign a θ-role in composition with some other θ-assignor. Consider, for instance, an indirect object in a dative construction:

(76) John gave the book to Mary

We will assume that Mary receives a θ-role from the verb, give; the preposition, to, merely acts as an intermediary in the process of θ-role transmission. Similarly, as Jaeggli argues, the head of the by phrase in passive constructions acts simply to transmit the external θ-role of the passivized verb; he shows that the θ-role assigned to the object of by varies according to the external θ-role on the verb:

(77)a. Bill was killed by Mary
b. The letter was received by Bill
c. The package was sent by John
d. That professor is feared by all students

The underlined NPs in (77a-d) receive the agent, goal, source and experiencer θ-roles, respectively; these θ-roles correspond to the external θ-roles assigned by the verbs kill, receive, send and fear.

If we examine prepositional phrases more generally, it becomes apparent that if the head assigns a θ-role, it must be a fairly "schematic" one, open to a variety of interpretations.

Consider a preposition like in:

(78)a. John was found in the woods.
b. Bill was wounded in the rioting.
c. In some respects, John is anachronistic.
d. Bill made faces at the judge in an act of contempt.

While in the woods in (78a) is clearly a locative, the prepositional phrases in (78b-d) are locative in only the most abstract sense. Given the variety of interpretations that a single preposition, in this case in, may receive, it is unclear what inherent thematic property, if any, unifies the various roles it may play. It would appear, rather, that their semantic function is determined compositionally on the basis of the predicates which they modify. Given the correctness of this line of reasoning, let us assume that prepositions do not assign a θ-role to their object.[27]

If we are correct in this hypothesis, then prepositional objects pose an apparent problem for the θ-criterion since we have not yet explicated a method for assigning any θ-role to

---

27 This hypothesis should be read with the proviso in note 25 kept in mind; that is, in the core case, prepositions do not independently assign a θ-role to their complements. Some prepositions like on, over, above, etc. can assign a θ-role to their objects directly. Notice that the ability of these prepositions to assign a θ-role may be over-ridden when the preposition occurs as the complement of a verb, as can be seen in the ambiguity of:

(i) They decided on the boat.

where on the boat may occur as a locative PP or as a complement of the verb decide. The θ-role assigned to the boat differs across these two cases.

prepositional objects. There are two possible cases. In the first case, the preposition is the head of a subcategorized complement; in the second case, the preposition is the head of a phrase in an adjunct position. We will assume in the first case, when the prepositional phrase is subcategorized by some element, X, that the preposition acts as an intermediary in the transmission of a θ-role from the head, X, to the object of the preposition. Thus, the prepositional object will receive a θ-role indirectly from the head that subcategorizes for a PP. In the second case, when the PP is in an adjunct position modifying some projection, XP, we will assume that a θ-role is randomly generated on the prepositional object; this θ-role is then checked at LF, when interpretive rules apply.

In both of the above cases, the prepositional object obeys the θ-criterion, by virtue of bearing a θ-role. In neither case, however, is the preposition determining the θ-role on its object. Higginbotham (1983) argues that similar considerations apply to θ-roles carried by nominal complements. He observes that arguments to nouns, unlike those of verbs, are always optional. If we assume, as seems reasonable, that the θ-criterion has the effect of making θ-role assignment obligatory, the optionality of nominal complements would appear to be problematic; either the apparently absent complements to N are realized as PRO (leading to the problems discussed above) or nouns freely violate the θ-criterion. Let us assume, however, that the thematic properties of nouns are similar to the thematic properties of prepositions:

Nominal complements are not assigned a θ-role directly by the head N; a θ-role is generated on the nominal complement directly, subject to a checking process at LF. This checking process compares the θ-role on the nominal complement with the argument structure of the head noun and determines if the complement is compatible with the argument structure of the head.[28]

We now have a property that unites prepositions and nouns beyond the categorial feature [-V]; neither category is capable of independent θ-role assignment. Their common property with respect to government may now follow not merely from their common categorial feature, but from properties of θ-role assignment. Our hypothesis, then, is that θ-theory, like Case theory, is sensitive to the grammatical category of elements; only items marked [+V] are capable of independent θ-role assignment:

(79)

N	V	Case Assigner	θ Assigner
+	+	-	+
+	-	-	-
-	+	+	+
-	-	+	-

In other words, Case theory specifies that only elements marked [-N] are capable of assigning Case directly and Thematic theory

28 An interpretive approach to the thematic roles associated with nominal complements would provide a research strategy for investigating the range of interpretations found for the subject of NP, as observed in Williams (1982).

specifies that only elements marked [+V] are capable of assigning a θ-role directly.

It remains to define "government across a maximal projection" in such a way as to exclude elements that do not directly θ-mark their complements:

(80) A maximal projection, XP, is transparent to government by an element Z if and only if Z assigns a θ-role to XP.

It follows from (80) that nouns and prepositions will not be able to govern into a complement phrase because they cannot θ-mark they cannot directly θ-mark their complements. Formally, we can indicate thematic role assignment in representations by means of superscripts. Following Kayne (class lectures; 1983), we will stipulate the following convention on θ-role assignment:

(81) If Z assigns a θ-role to XP, then cosuperscript Z and XP.

Principle (80) may then be restated in terms of superscripts:

(82) XP^{i} is transparent to government by Z^{j} if and only if $\underline{i} = \underline{j}$.

Although nouns and prepositions may perfectly well govern a sister node, it follows from (81) and (82) that nouns will be incapable of governing into a maximal projection.

In this section, we have seen that nouns, unlike verbs and adjectives, do not directly θ-mark their complements. This incapacity has the consequence that maximal projections are always absolute barriers to government by nouns. We will now turn to the investigation of how the ability to govern across

maximal projections relates to the structure of retroactive nominals.

2.6 PRO in NP

We have identified a cluster of properties associated with retroactive nominals. First, these nominals apparently contain an empty category with anaphoric properties, in that the empty category is sensitive to Specified Subject Condition effects. Second, these nominals must be governed by a verbal element that falls into a restricted semantic class ("requirement verbals"). Finally, as the examples below illustrate, the head noun of retroactive nominals has a restriction on its morphological form:

(83)a. *the baby could use some attention to
b. *these diplomats deserve some conversation with
c. *these ideas don´t need any arguments about
d. *the status quo doesn´t merit any revolution against

None of the derived nominals in (83) involve ing affixation whereas the examples of retroactive nominals given above all involve ing nominalization. Non-ing derived nominals are not absolutely barred from the object position position of "requirement" verbals as the examples in (84) show:

(84)a. The baby could use some attention.
b. This scoundrel deserves a prompt conviction by the jury.
c. John fully merits ostracization by the government.
d. This body needs a rapid cremation by the mortuary.

The contrast between (83) and (84) indicates that only ing derived nominals allow preposition stranding within the NP. Compare the examples in (83) with those in (85):

(85)a. The baby could use some attending to.
b. These diplomats deserve some conversing with.
c. These ideas don´t need any arguing about.
d. The status quo doesn´t merit any revolting against.

The examples in (84) do share at least one property with other retroactive nominals; they show a Specified Subject Condition effect:[29]

(86)a. *this scoundrel deserves the jury´s prompt conviction
b. *John fully merits the government´s ostracization
c. *this body needs the mortuary´s rapid cremation

We can account for the contrast between (84) and (86) if we assume that there is an empty category in the object position of these nominals and that this empty category must be locally bound. Notice that the examples in (86) have a syntactically well-formed representation which leads to a semantically anomolous interpretation; namely, the genitive NP can be taken as the antecedent of the empty category in the object position of the nominal. Example (86b) would then have the representation in (87):

(87) [John]$_{N''}$ merits [[the government´s]$_{N''}$]$_{N''}$
$_i$ $_i$
ostracization [e]$_i$]

The interpretation of (87) would be that John merits the government´s being ostracized. This interpretation indicates

---

29 The nominal attention seems less sensitive to SSC effects, as shown in (i):

(i) ?The baby could use the nurse´s attention

I have no account for why this should be.

that retroactive nominals have a passive interpretation. Suppose, for the sake of discussion, that the structure of a retroactive nominal is:

(88) $[_{N''}\ [_{N''}\ \underline{e}]_i$ a good talking to $\underline{t}_i\ ]]$

The trace in the object position of the preposition is locally A-bound by a null element in the Specifier of the NP. Since the empty category is locally A-bound, it will be marked as a pure anaphor. Hence, the SSC effects and the locality condition, noted above, follow directly from the Binding theory and subjacency. Consider an example like:

(89) *[John's paper]$_i$ could use $[_{N''}\ [_{N''}\ \underline{e}]_i$ some convincing (of) people $[_{\bar{S}}$ that $[_{S}$ they should read $\underline{t}_i\ ]]]$

The trace, $\underline{t}_i$, must be bound in the minimal maximal category containing a governor and a SUBJECT. A trace in the governing domain of the head noun will be bound in its governing category. A trace in a relative clause or a clausal complement to the head noun will have as its governing category the embedded clause, not the superordinate NP; hence, such a trace will not be bound in its governing category and will either violate condition A of the binding theory or be interpreted as a PRO. Furthermore, the movement of the null element to the Spec of the superordinate NP will violate subjacency.

The assumption that retroactive nominals involve NP-internal A-binding has a number of other consequences. Recall that an

overt subject seems impossible in these examples:

(90) *John$_i$ could use [[a competent psychiatrist's]$_{N''}$ $_{N''}$ talking to [e]$_i$]

For reasons that will be discussed below, the non-overt NP may not remain in situ as the object of a preposition; the non-overt NP must move to an A-position in the Spec of the NP. The subject position of the object NP in (90) is already filled, however. Hence the non-overt nominal has no landing site and must remain in place. Since this situation is, by hypothesis, impossible, we can account for the ungrammaticality of (90). Furthermore, given that the non-overt element occupies an A-position in the Spec of the retroactive nominal, we can account for the unacceptability of parasitic gaps (cf. the examples in (41), (42) and (43)) since elements in A-positions are incapable of licensing parasitic gaps.

We are now confronted with the question of the status of the non-overt element in the Spec of NP. Since this element binds a nominal trace, we will assume that it is, itself, an NP.[30] This

---

30 Here, we assume, following Kayne (1984), that a potential binder of a trace must be of the same category as the trace. He observes that this accounts for the contrast in (i):

(i)a. [Poland's]$_{N''}$ $_i$ invasion [e]$_{N''}$ $_i$ by Germany...

b. *the [Polish]$_{Adj}$ invasion [e]$_{N''}$ by Germany...

In (ia), the trace in the object position is A-bound by the genitive NP. In (ib), the trace is A-free because the adjective, Polish, cannot bind the trace of an NP. Example (ib),

non-overt NP is, in turn, A-bound by the subject of the superordinate S. As argued above, the subject position of sentences containing retroactive nominals receives an independent θ-role (cf. the examples in (19)). Since the non-overt NP in the subject position of the retroactive nominal is A-bound by an NP bearing an independent θ-role, it must be interpreted as PRO.

If the non-overt nominal in the Spec of NP is, in fact, PRO, then we are faced with an apparent dilemma. The definition of government given in Aoun and Sportiche (1983) allows a head to govern throughout its maximal projection. This definition was intended, in part, to rule out a PRO in the Spec position of NP; it was argued that such a PRO should admit arbitrary interpretation which predicts that a pronoun bound by that PRO should receive an arbitrary interpretation:

(91) [un livre sur sa vie] vient d'être publié
a book on {his/*one's} life has just been published.

Stowell (1983) observes, however, that some nominals in subject position may contain a pronoun with arbitrary interpretation:

(92)a. [un livre sur sa vie] est toujours utile.
a book on one's life is always useful
b. [hatred of oneself] is dangerous.

Following Horvath (1981), Stowell proposes that government contains a directional parameter. We may, therefore, propose the following definition of government:

---

therefore, violates the Binding theory.

(93) A governs B if and only if
(i) both A c-commands B and B c-commands A;
(ii) either A precedes B and the language is head-initial or B precedes A and the language is head-final.

Since English is a head initial language, the above definition has the consequence that elements occurring in the specifier of a head will not be governed by the head since the head does not precede them.

Let us, now, reconsider the proposed structure of retroactive nominals:

(94) $[_{N''}$ PRO$_i$ a good talking to t_i]

Although the PRO and the head noun, talking, stand in a symmetrical c-command relation, the PRO precedes the head and is, therefore, not governed by it. If we assume the PRO theorem (Chomsky, 1982), then the movement from object (of a preposition or of the noun itself) is forced. Consider the D Structure associated with (94):

(95) $[_{N''}$ a good talking to PRO]

If the PRO remains in situ, it will occur in a governed position. Movement to the ungoverned Spec of NP position is, therefore, forced by definition of government and by the Binding theory. Since movement to the Spec of NP is forced under this analysis, the SSC facts follow immediately, as argued above.[31]

---

31 J. Aoun (p.c.) and T. Hoekstra (p.c.) have observed that under the assumption that governed PRO is a pure anaphor, the SSC facts may still be made to follow. See the discussion, below.

2.7 Thematic properties and proper government

We observed in the previous section that, although any NP could occur in the retroactive construction when the trace is a right-sister of the head, only ing nominals allowed preposition stranding. Nothing in our analysis so far predicts this curious asymmetry between ing nominals and other types of derived nominals; nevertheless, there is a strong contrast in grammaticality between (96a) and (96b):

(96)a. *The baby needs some attention to.
b. The baby needs some attending to.

A contrast as strong as that found in (96a) and (96b) demands an explanation: We would like to capture the fact that preposition stranding is "unusual" for PPs inside many derived nominals, but that ing nominals exceptionally allow preposition stranding.

It has often been noted by a number theorists (van Riemsdijk (1978), Hornstein & Weinberg (1981), Stowell (1981), Jaeggli (1982), Kayne (1984), to name but a few) that preposition is a marked property of English. It is commonplace for preposition stranding to be impossible in the majority of the world's languages. There have been a number of attempts to capture the ability of English to strand prepositions while properly reflecting the fact that this is a marked property of grammar. Suppose, following Chomsky (1981), that empty

---

In either event, there seems to be good evidence that PRO exists in this construction.

categories must be properly governed by some head and that English, unlike most languages, allows prepositions to act as proper governors. Example (96a) indicates, however, that such an approach to preposition stranding will be of little use to us in dealing with the problem at hand. The trace in the object position of the PP will not only be locally A-bound by the PRO in the Spec of the NP, but it will be properly governed by the preposition, since, under this view, English admits prepositions into the class of proper governors.

An alternative approach is provided by Hornstein and Weinberg (1981). Unlike the previous analysis, they make no appeal to any special governing properties on the part of prepositions. Their analysis hinges on the interaction between Case theory and a marked rule of reanalysis. They argue for a filter of the form:

(97) $*[\underset{NP}{\underline{e}}\ \ \underset{oblique}{\ }]$

The filter is intended to rule out traces marked with oblique Case. If we assume that prepositions generally assign oblique Case to their objects, then filter (97) rules out preposition stranding across the board; it should be noted that, even if prepositions are taken to be proper governors, this filter will independently rule out preposition stranding. English is exceptional, however, in that it allows a marked rule of reanalysis to apply to certain strings, thus converting these strings into single words:

(98) V --> V* (where V c-commands all elements in V*)

This rule will reanalyze (99a) as (99b):

(99)a. [John [[talked] [to Harry] [about Fred]]]
S VP V PP PP

b. [John [[talked to Harry about] Fred]]
S VP V

The NP, Fred, in (99b) is governed by a complex verb, talked to Harry about. Since the NP is governed by a verb, it will be assigned accusative Case, rather than the oblique Case assigned by a preposition. If the NP is extracted, the trace that it leaves behind will also be marked accusative and filter (97) will not apply to the representation. English, therefore, should allow preposition stranding insofar as the environment of the reanalysis rule is satisfied.

In general, NPs do not satisfy the environment of the reanalysis rule, since reanalysis is defined only for verbs. This result seems correct since, as observed above, NPs lack pseudo-passive constructions:

(100)a. *the diplomats´ conversation with t by John...
b. *the legislation´s arguments about t by the congress...
c. *the problem´s thought about t by John...

Since the prepositions in the above examples will not reanalyze with the head noun, the trace left by "move alpha" will be marked with oblique Case and the representation will be ruled out by filter (97). Unlike the ECP analysis of preposition stranding (above), the Case analysis will rule out all cases of preposition stranding in NPs, barring the interaction of independent

principles.

Recall that the retroactive nominal construction is lexically governed, that is, a verbal of the appropriate semantic class must govern the retroactive nominal in order to trigger the NP-movement of PRO within the nominal. Furthermore, the head noun must be an <u>ing</u> nominal in order to allow preposition stranding. One could complicate the reanalysis rule (98) by directly stipulating that reanalysis applies in the following environment:

(101) V [+F] N-ing P --> V*

where [+F] is a lexical feature denoting the semantic class of verbs that allow retroactive nominals as complements. This rule would convert the string in (102a) to the string in (102b):

(102)a. [this problem [[deserves] [thinking
S VP V NP
[about PRO]]]]
PP
b. [this problem [[deserves thinking about]
S VP V*
PRO]]

Because of its stipulative character, this analysis fails to provide any illumination on a number of explanatory issues. This analysis, for example, fails to provide an explanation as to why it should be the case that only <u>ing</u> nominals satisfy the environment for reanalysis. Aside from the stipulative, and, hence, unexplanatory nature of this analysis, it also has a number of empirical and conceptual problems. For example, how can it be that the PRO that is the D Structure object of the

preposition can move inside of the reanalyzed word?

For the moment, let us assume a form of the ECP which requires that an empty category be governed by a coindexed element. Coindexation will result in one of the following two ways:

(103) An element, A, is coindexed with an element, B only if either:
(i) A assigns a θ-role to B
or
(ii) A binds B.

The requirement that empty categories must be governed by a coindexed element has the result that , all else being equal, nouns and prepositions will not count as proper governors for their complements because, as argued above, they are normally incapable of directly assigning a θ-role to their complements.

Consider the case of a prepositional phrase that is a complement to a verb, as in:

(104) $[_{S}$ John $[_{VP}$ $[_{V}$ talked$]$ $[_{PP}$ to Bill$]]]$

The verb, talk, will assign a θ-role to its complement, to Bill. By convention, the verb and the preposition phrase will be cosuperscripted under θ-role assignment:

(105) $[_{S}$ John $[_{VP}$ $[_{V}$ talked$]^{k}$ $[_{PP}$ to Bill$]^{k}$ $]]$

The θ-role assigned to the prepositional phrase will, under standard assumptions, percolate to the head of that category, the preposition, to. Let us suppose that English exceptional allows direct θ-role transmission by the preposition; that is, the θ-

role assigned to the PP may be transmitted by the head to its complement. Given this mechanism, English exceptionally allows prepositions to be cosuperscripted with their complements:

(106) $[_{S}$ John $[_{VP}$ $[_{V}$ talked$]^{k}$ $[_{PP}$ to^{j} Bill$^{j}]^{k}]]$

Notice that the object of the preposition is cosuperscripted with the head in (106). Thus, the process of θ-role assignment ultimately conditions the distribution of empty categories.

As Hornstein and Weinberg observe, following Chomsky (1965), adjunct PPs may not be stranded:

(107)a. *what time did John arrive at
b. *which inning did the Yankees lose the ball game in

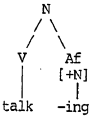
The facts in (107) follow from our analysis; since these PPs are not assigned a θ-role by the verb, the head preposition cannot transmit a θ-role to its complement and the head and its complement will not be cosuperscripted. I will assume that this process subsumes reanalysis, at least insofar as preposition stranding is concerned.

Despite the inconclusive nature of the analysis up to this point, we have isolated some of the ingredients of a complete analysis of retroactive nominalization. First, the retroactive nominal must be governed by an element that is a member of a particular semantic class of verbals. To capture this fact, let us assume that these verbals have some common element in their argument structure (i.e., the set of thematic roles assigned by a particular element). We will label the thematic element common

to this class of verbals "θr".

Second, we have seen that when ing nominals receive the θ-role, θr, they take on "verb-like" properties in that they apparently permit a form of reanalysis. Following Williams (1981), let adopt the Righthand Head Rule (RHR) which defines the head of a morphologically complex lexical item to be the rightmost element in the item. It will follow, then, that the nominalizing affix, ing, is the head of an ing nominal:

(108)



Let us suppose that assignment of θr to ing nominal renders the N boundary dominating the V+Affix invisible just in case the Affix is ing. This process is analogous to Exceptional Case Marking which renders an $\overline{S}$ boundary transparent to government (and Case assignment) only if the Comp node is [-finite].[32]

Given that the N boundary has been rendered invisible by

---

32 In particular, the Comp must subcategorize for a to infinitive. This analysis of θ-role assignment and ing nominals raises an interesting possibility. Following recent work by Kayne, let us suppose that the possibility of Exceptional Case Marking is interconnected with preposition stranding and the possibility of dative shift constructions. We would assume that the ability to make the N boundary transparent is also tied in with the permissibility of these other constructions. Thus, a language that admits one of these constructions should allow all of them. A full investigation of this is beyond the scope of this work.

virtue of the assignment of θr by the governing verbal, the [+V] stem is now free to assign a θ-role to its complement. Since θ-role assignment by the nominal is now licensed, the complement which receives the θ-role will be cosuperscripted with the head, as in (109):

(109) [John [[needs]^{k} [a good talking^{j} [to
S VP V NP PP
PRO$^{j\,k}$]]]]

Since the head of the PP complement to the noun now bears a fixed θ-role, it is now able to assign that θ-role to its complement; hence, the preposition, <u>to</u>, and its NP complement, the PRO, will be cosuperscripted as in (110):

(110) [John [[needs]^{k} [a good talking^{j} [to^{i}
S VP V NP PP
PRO$^{i\,j\,k}$]]]]

"Move alpha" may now freely apply to the PRO, moving it into the Spec of NP; since the head of the prepositional phrase and the trace left by this movement are cosuperscripted, the version of the ECP formulated above will be satisfied:

(111) [John [[needs]^{k} [PRO$_{1}$ a good talking^{j}
S VP V NP

[to $\underline{t}_{1}^{\,i\,i\,j\,k}$]]]]
PP

In (111), the empty category, $\underline{t}_{1}$ is governed by the preposition, <u>to</u>. In addition, because of a conspiracy between exceptional θ-

role assignment by the ing nominal and the marked ability of prepositions to transmit θ-roles directly in English, the trace is governed by the empty category in the way that is relevant to the ECP. Finally, the trace is A-bound in its governing category by the PRO in the Spec of NP. Since the trace obeys both condition A of the Binding theory and the ECP, the grammar will allow the representation in (111).

The D Structure in (109) should be contrasted with the following D Structure:

(112) $[_{S}\ [_{NP}$ some thinking $[_{PP}$ about PRO] is necessary]

An ing nominal occupies the subject position in (112). An NP in this position will not receive the θ-role, θr; since θ-marking by the head noun to its complements presupposes this θ-role, the ing nominal will not assign a θ-role to the complement PP and the head will not be cosuperscripted with its complement. Since the PP is not directly assigned a θ-role by the head noun, the head of the PP cannot directly transmit that θ-role to its complement. As a result, the preposition and its complement will not be cosuperscripted. Extraction of the prepositional complement, as in (113), will therefore violate the ECP:

(113) $[_{S}\ [_{NP}$ PRO$_1$ some thinking $[_{PP}$ about $\underline{t}_1$]] is necessary]

Movement of the PRO to the Spec of NP will trigger a violation of the ECP; the PRO, however, may not remain in situ since it will be governed by the preposition. There is no well-formed

derivation associated with example (112) since all the possible representations violate either the Binding theory or the ECP. A conspiracy between properties of θ-role assignment and the requirement that empty categories be governed by a coindexed element guarantees that the retroactive nominal construction will be isolated to the object position of verbals that assign the θ-role, θr.

Consider the case of a derived nominal which is not headed by -ing; compare the D Structure in (109) with:

```
(114) [ [   these people] [   [  merit] [   some
        S NP                VP  V          NP
        conversation [   with PRO]]]]
                      PP
```

The verb, merit, in (114) will assign the θ-role, θr, to its complement NP. The head of the noun, conversation, is not, however, ing. Since transperancy of the N boundary presupposes that the head is ing, conversation will not assign a θ-role to its complement PP. From this step, the proof of the ungrammaticality of (114) is identical to that of example (112). The head of the PP, with, will not assign a θ-role to its complement; any resulting representation associated with this example will violate the ECP:

```
(115) [ [   these people] [   [  merit] [   PRO  some
        S NP                VP V          NP    1
        conversation [   with t ]]]]
                      PP       1
```

Finally, we can consider the case of extraction from a phrase which is not a complement of the head N. Consider examples of the following form:

(116)a. John falls asleep [$_{PP}$ during his matrix algebra class]

b. Most farmers work [$_{PP}$ after a good rain]

c. Most of the audience left [$_{PP}$ before John's film]

The PPs in the above examples are adjuncts which do not receive a θ-role from any head element in the sentence. As a result, it is impossible to extract the complement NPs, allowing the head preposition to be stranded:

(117)a. *Which class$_i$ did John fall asleep [$_{PP}$ during $\underline{t}_i$]

b. *What kind of rain$_i$ do most farmers work [$_{PP}$ after $\underline{t}_i$]

c. *Which film$_i$ did most of the audience leave [$_{PP}$ before $\underline{t}_i$]

The ungrammaticality of the examples in (117) follows immediately from our assumptions about the ECP and the nature of θ-role assignment: Prepositions may directly assign a θ-role only insofar as they are transmitting a θ-role from a head. Since the PPs in (117) are not complements of a head, they are not in a position to transmit a θ-role. Hence, in each of the examples in (117) the head of the PP will not be cosuperscripted with its complement. Since every empty category must be governed by an element which is coindexed with it, none of the empty categories in (117) will pass the ECP.

Our analysis predicts that stranding of a preposition in

retroactive nominals is possible only if the PP is assigned a θ-role by the head noun of the construction. A prepositional adjunct to the head noun should not allow stranding:

(118)a. *no class in matrix algebra deserves any falling asleep during
b. *most of the farmers agree that a good rain merits a lot of working after
c. *even though it was poorly made, John's film did not deserve any leaving before by the audience

The examples in (118) are out for essentially the same reasons as examples (112) and (114). The PRO that is the D Structure object of the prepositional adjunct will violate the ECP if it moves out of that adjunct and it will violate the Binding theory (by being in a governed position) if it remains in situ.

It is quite generally impossible to strand a preposition in retroactive nominal constructions unless the PP is directly θ-marked by the head noun; it is, therefore, impossible to strand a preposition that is inside a complement of the head noun:

(119)a. *Nixon deserves some rewriting of [a book [about e]] (cf. "Many books about Nixon deserve some rewriting.")
b. *John merits some talking to [friends [of e]] (cf. "Friends of John merit some talking to.")
c. *This foundation needs some rethinking of [gifts [to e]] (cf. "Gifts to this foundation need some rethinking.")

In each of the examples in (119), the PP containing the stranded preposition has not been θ-marked. Once again, any representation will ultimately involve a violation of the ECP (if movement occurs) or a violation of the Binding theory (if the PRO object of a preposition remains in situ). Either way, there is

no mechanism that will save the representations.

2.8 Governed PRO

We have assumed, up to this point, that PRO may not be governed since it is a pronominal anaphor. A pronominal anaphor, by definition, is subject to conditions A and B of the Binding theory; in others words, a pronominal anaphor must be both bound and free in its governing category. In order to circumvent this contradiction, it must be the case that PRO lacks a governing category. The one sure way to avoid having a governing category is to avoid being governed. The distribution of PRO is, thus, limited to ungoverned positions like the subject position of non-finite clauses and, as argued in the previous sections, the subject position of NP.

In recent work, Bouchard (1984) and Sportiche (1983), among others, have argued that there is no reason to assume that an element with the features [+anaphor, +pronominal] exists. All the properties commonly attributed to PRO may be derived independently via the interaction of the Binding theory, Case theory and thematic theory. The approach developed by Bouchard and Sportiche obviate the requirement that PRO be ungoverned; as we will see, they assume that it is quite possible for what is commonly referred to as PRO to be governed.[33] If we grant the

[33] If PRO may be governed, then our argument that government is directional may be weakened. It does not follow, however, that government is not directional. In this connection, see Stowell (1983).

assumption that PRO may be governed, then it is of some interest to discover whether or not the analysis of retroactive nominals given above may be maintained. If our analysis does not carry over to the "governed PRO framework", then it will pose an interesting research problem for the proponents of that theory.

Bouchard (1984) discusses the following sentences of obligatory control:

(120)a. John promised [PRO to leave early]
b. John wanted [PRO to leave early]
c. Mary persuaded John [PRO to leave early]

In all the examples in (120), John is the antecedent for PRO. The examples in (120) should be contrasted with those in (121):

(121)a. John knows [how [PRO to do the proof]]
b. John wondered [how [PRO to do the proof]]
c. John showed us [how [PRO to do the proof]]

In the examples in (121), the subject of the subordinate clause is not necessarily taken to be John; example (121a) means that John knows how anyone would go about doing the proof. The sentences in (121) illustrate examples of arbitrary PRO, meaning that PRO is interpreted as ranging over the set of humans. I will refer to this last type of control as "optional control".

Superficially, however, the sentences in (120) differ from those in (121) minimally. The one difference between the two classes of examples is that the sentences in (121) all have an overt complementizer for the embedded clause. Let us assume Comp is present only if it is independently required by, for example,

subjacency.[34] There is no reason to assume that the embedded clauses in (120) have a Comp node; given this line of reasoning, in fact, the embedded clauses in (120) must be bare S´s, not $\overline{S}$´s. If this is so, then the matrix verbs in (120) govern the subject positions of the embedded clauses. The embedded clauses in (121) all have an overt complementizer and, hence, must be $\overline{S}$´s. We have pinpointed a clear structural difference between the obligatory control sentences in (120) and the arbitrary control sentences in (121).

Bouchard argues that controlled PRO is always governed while optionally controlled PRO is ungoverned. To put this in line with the Binding theory, he proposes that:

(122)a. Governed PRO is a pure anaphor.
b. Ungoverned PRO is a pure pronominal.

(122a) will derive the fact that the PROs in (120) all find their antecedents in the superordinate S. Since each of the PROs are governed by the matrix verb and have an accessible SUBJECT in the matrix, the matrix S is the governing category for the anaphoric PROs and they must be bound in that category. We thus account

---

34 The assumption that Comp is absent unless required by independent principles is far from harmless. In section 5 we argued, following Stowell (1981), that subject sentences required an overt complementizer because an empty Comp in that position violated the ECP. This account won´t hold if we assume that that infinitivals do not have a Comp obligatorily (it is obviously perverse to say that infinitivals in subject position require a Comp in order to violate the ECP). Prima facie, we are in the embarrassing position of having to stipulate that clauses in subject position have an overt complementizer rather than deriving this fact from a basic principle like the ECP. For further discussion, see Stowell (1985).

for the fact that obligatory control of PRO does not take place over arbitrarily long distances.[35]

(122b) treats ungoverned PRO as a pure pronominal. If this PRO has an antecedent, its antecedent may be arbitrarily far away; we have a simple account of examples like (123):

(123) [they were offended that [I thought that [Mary
S_i S S
claimed that [[PRO feeding each other] would be
S_i
difficult]]]]

In example (123) a PRO in the sentential subject of a deeply embedded clause is taken as coreferent with the matrix subject. As example (124) shows, this coreference relationship is not obligatory:

(124) [they were offended that [I thought that [Mary
S_i S S
claimed that [[PRO feeding oneself] would be
S_{arb}
difficult]]]]

These facts follow from (122) since a PRO in a non-finite sentential subject is ungoverned and, by (122b) must be a pronominal. Given its pronominal status, an ungoverned PRO may freely corefer with another NP over an unbounded distance.

Recall that a central part of the analysis of retroactive nominals given above was that a PRO object of the noun or

---

35 It should be noted that a theory which indicates the controller of an obligatorily controlled PRO by designating an element of the control verb's argument structure will also derive the above locality constraint. See, for example, the theory of control developed in Manzini (1983).

prepositional complement was forced to move to the Spec of NP in order to avoid being governed. This NP movement of PRO was used to derive two facts: First, the SSC effect that retroactive nominals display; second, the fact that only ing nominals allow preposition stranding. The treatment of PRO outlined above, however, allows a controlled PRO only if it is governed. If we adopt the hypothesis that PRO may be governed, then we apparently cannot appeal to a movement analysis of retroactive nominals to derive the SSC and preposition stranding facts.

We can, however, appeal to other subcomponents of grammar to derive these facts. First, let us adopt Bouchard's argument that elements that receive Case must be phonologically realized (the Principle of Lexicalization). It follows from the Principle of Lexicalization that PRO may never receive Case.[36] Second, modifying a proposal by Hornstein & Weinberg (1981) and Kayne (1984), let us assume that English allows the head of a PP that is cosuperscripted with some lexical element (by θ-role assignment) to take on the Case assigning properties of that head. From this it will follow that the head of a PP cosuperscripted with a noun will not assign Case to its object the ing nominal to the prepositional phrase. Given that the head noun and the PP are cosuperscripted, the head of the PP, to, takes on the Case assigning properties of the noun; in other

---

36 We will assume that a variable transmits its Case to the operator which binds it.

words, the preposition will not assign Case to the PRO. It follows from the Principle of Lexicalization that the PRO will not be phonologically realized. The PRO object of the preposition is governed and, by (122a), it is a pure anaphor and must be bound in its governing category, the matrix S. Since the PRO is bound by John, the representation (125a) does not violate any of the relevant principles and is ruled grammatical.

In (125b), however, assignment of θr to the object NP will not license θ-role assignment by the head noun since the head is not an ing nominal. Since θ-role assignment by the head noun is not possible, the noun and its complement PP will not be cosuperscripted. The head of the PP, with, will therefore retain its Case assigning properties. The PRO object of with will violate the Principle of Lexicalization and the representation associated with (125b) will be ruled out. It would appear that the governed PRO framework can account for the preposition stranding facts by appealing to Case theory and the Principle of Lexicalization.

The second puzzle that the governed PRO framework must account for is the apparent SSC effects shown by retroactive nominals. For convenience, I repeat some of the relevant examples:

(126)a. *John$_i$ could use [[a competent psychiatrist´s]$_{N''}$ talking to PRO$_i$]$_{N''}$

b. *these ideas$_i$ merit [[Bill´s] working on PRO$_i$]$_{N''}$

c. *this problem bears [[everybody´s] thinking about PRO]
i N" N" i

d. *my room needs [[the janitor´s] picking PRO up]
i N" N" i

In all the examples of (126), the matrix verbal assigns θr to its object, thus licensing θ-role assignment by the head noun of the object NP; hence, the head noun and its complement will be cosuperscripted. By our assumption that cosuperscripting between a head and its complement gives the complement the same Case feature as the head, it follows that the PRO object in the above examples will not be assigned Case. Hence, the Principle of Lexicalization will not be violated by any of the PROs in (126).

Although none of the PROs in (126) receive Case, each of them is governed. It follows from (122a) that each of the PROs is a pure anaphor and, by condition A of the Binding theory, must be A-bound in its governing category. Since each of the object NPs has a prenominal genitive NP which will act as a SUBJECT for the PRO, it follows that the governing category for the anaphors is the object NP. Since the matrix subject is outside of the governing category for the PRO, the indexing shown in (126) is impossible since the PRO will be free in its governing category, in violation of condition A.

We have shown that the governed PRO framework rules out one possible indexing in the examples like (126). Consider the predictions made with respect to the following assignment of

indices:

(127)a. John could use [[a competent
N" N"
psychiatrist´s] talking to PRO]
i i

b. these ideas merit [[Bill´s] working on
N" N" i
PRO]
i

c. this problem bears [[everybody´s] thinking
N" N" i
about PRO]
i

d. my room needs [[the janitor´s] picking PRO
N" N" i i
up]

In each of the examples in (127), the anaphoric PRO is bound by the SUBJECT of its governing category, the prenominal genitive NP. Since each of the anaphors is bound in its governing category, no condition A violation occurs and the examples in (127) should be ruled grammatical. In the PRO theorem framework, these examples were ruled ungrammatical because the PRO was in a governed position and could not move to an ungoverned landing site. Thus, the governed PRO framework predicts that example (127a) should be grammatical with the following interpretation:

(128) John could use a competent psychiatrist´s talking to
j i
himself .
i

We should be precise about the sense in which <u>grammatical</u> should be taken in the preceding paragraph. Given a modular theory of grammar, the status of <u>grammatical</u> must be relativized to a component or interacting set of components. Thus, one could argue that <u>grammatical</u> in the preceding paragraph should be read

as "grammatical with respect to the constraints placed on S Structure and LF". This should not be taken as implying that some other component does not mark the examples in (127) as deviant, if not ungrammatical in the sense of syntactic theory. One could argue, for example, that the examples in (127) are pragmatically odd, although syntactically well-formed.[37] In support of this, we could point out the following example:

(129) The district attorney needs a quick conviction.

Sentence (129) is, of course, ambiguous. It could mean that the district attorney is on trial and is so heinously guilty that the conviction cannot be quick enough; this would be the retroactive nominal reading. Suppose, however, that the state has placed a quota system on the district attorney; out of a certain number of trials, the district attorney's office must win some proportion of convictions or the district attorney will be fired. Sentence (129) could mean that the district attorney needs for someone in his office to win a quick conviction of some alleged criminal. This reading is similar to the interpretations for the examples in (127) insofar as the subject of the sentence is not associated with an argument of the object nominal.

The above discussion merely serves to illustrate the obvious

---

37 "Pragmatically odd" is so vague that I would prefer an analysis that does not make reference to it. I believe that a rational research strategy is to try to make use of core principles wherever possible. It is not, however, incoherent to appeal to pragmatic principles however poorly understood these principles may be.

point that a particular construction does not have any a priori status with respect to the theory. One approach may derive a particular range of facts as following from one component while another approach may derive the same facts from a separate component. It may even be that an approach may choose to set aside some range of facts treated by other frameworks. The data in itself cannot decide between competing theories; theories must be evaluated in terms of the problems that they solve and their explanatory power.

I will not attempt to decide what status the facts in (127) should have in the theory since their status ultimately depends on the choice between a theory incorporating the PRO theorem and a theory incorporating governed PRO; although it is certainly true that evidence based on the nominals discussed in this chapter could play a part in resolving this theoretical issue on empirical grounds. The range of data and the explanatory problems that must be taken into account in deciding between these two theories is, no doubt, very extensive and would take us far afield. My point in this section is that, whichever of the two is ultimately selected, certain aspects of the analysis of retroactive nominals remain constant across the two theories. First, the proper analysis of retroactive nominals involves control of PRO--however PRO is ultimately analyzed. This conclusion implies that PRO occurs inside Noun Phrases. Second, thematic theory constrains the functioning of other components of the grammar in a way that extends beyond the θ-criterion and the

Projection Principle. In a framework that invokes the PRO theorem, thematic theory constrains the ability to extract an element from a phrase. We have so far assumed that this ability to extract from a phrase follows from a rather simple formulation of the Empty Category Principle; thus, thematic theory interacts with the ECP. In a framework that invokes a governed PRO, thematic theory constrains the assignment of abstract Case. For the remainder of this chapter, I will assume that the PRO theorem is correct and, hence, that thematic theory constrains the functioning of the ECP. As will become apparent in the last chapter of this work, I do not take the PRO theorem as a given.

2.9 A note on "worth"

Throughout the discussion of retroactive nominals, I have systematically avoided examples involving worth. At first, it would appear that worth falls solidly into the class of verbals that license retroactive nominals:

(232)a. This problem is worth working on.
b. Picasso´s later paintings are worth looking at.
c. John´s arguments aren´t worth any worrying about.

As pointed out by Safir (1984), worth shows some systematic differences from the other class of verbals. First, the complement to worth does not show Specified Subject Condition effects:

(233)a. This problem is worth Bill´s working on.
b. Picasso´s later paintings are worth your students´ looking at.
c. John´s arguments aren´t worth his wife´s worrying about.

Second, the movement inside the complement to worth may be unbounded:

(234)a. This problem$_i$ is worth [persuading some students [PRO to work on t_i]]

b. Picasso´s later paintings$_i$ are worth [forcing [some students to look at t_i]]

c. John´s arguments$_i$ aren´t worth [convincing [anyone to worry about t_i

Third, the subject position of S may be a $\overline{\theta}$-position since pleonastic it may occur there:

(235)a. It is worth working on this problem.
b. It is worth looking at Picasso´s later paintings.
c. It isn´t worth worrying about John´s arguments.

These "worth" sentences are strikingly similar to tough movement (for further arguments see Safir (1984)) and dissimilar to retroactive nominals which have none of the above properties. Strikingly, however, if we disambiguate the complement of worth as to whether it is a true ing nominal or a gerund, we begin to get a number of contrasts. Thus, if the complement to worth is unambiguously an ing nominal, the properties of retroactive nominals return:

(236)a. This problem is worth some working on.
b. Picasso´s later paintings are worth a lot of looking at.
c. John´s arguments aren´t worth any worrying about.

(237)a. *This problem is worth the working on.
b. *Picasso´s later paintings are worth the looking at.
c. *John´s arguments aren´t worth the worrying about.

(238)a. *this problem$_i$ is worth [some persuading of some students [PRO to work on $\underline{t}_i$]]

b. *Picasso´s later paintings$_i$ are worth [some forcing of [some students to look at $\underline{t}_i$]]

c. *John´s arguments$_i$ aren´t worth [any convincing [anyone to worry about $\underline{t}_i$

(239)a. *it is worth some working on this problem.

b. *it is worth a little looking at Picasso´s later paintings.

c. *it isn´t worth any worrying about John´s arguments.

Briefly, we can encapsulate the properties of worth as follows: Unlike the other verbals that license retroactive nominals, worth allows a true gerund as its complement; the other verbals require an NP. If a gerund occurs as the complement of worth, a non-overt operator may undergo wh-movement inside the complement. This movement is analogous to tough movement. If a true ing nominal occurs as the complement to worth, then the construction is another instance of retroactive nominals. This analysis makes an interesting prediction. Recall that retroactive nominals do not felicitously license parasitic gaps, while parasitic gaps may occur in tough movement constructions. We should, then, get a contrast depending on whether the complement to worth is a gerund or an ing nominal:

(240)a. This problem$_i$ is worth working on $\underline{t}_i$ without losing any sleep over $\underline{e}_i$

b. *?this problem$_i$ is worth some working on $\underline{t}_i$ without losing any sleep over $\underline{e}_i$

The contrast between (240a) and (240b) indicates that worth does, in fact, subcategorize either for an NP or for a gerund and, further, that movement of a non-overt operator to an $\bar{A}$-position, in a way which is analogous to tough movement, is possible in gerunds.

2.10 Summary of Chapter 2

We began this chapter by considering what properties the empty category found in retroactive nominals must have. We found that, although the empty category has anaphoric properties, treating it as a pure anaphor raised a number of problems for the theory of empty categories since it appeared to be A-bound in its governing category by an element with an independent θ-role. This indicates that the empty category in question must be a PRO. This PRO, however, appeared to be in a governed position, a situation which is impossible in a theory which incorporates the PRO theorem. A systematic investigation of the other possible empty categories showed that none of them could completely account for the facts. A null pronominal in that position would violate condition B of the Binding theory and, furthermore, could not meet the requirement that pro be locally identified. An NP-trace would violate the θ-criterion since it would form a chain with its binder. The assumption that the empty category was a variable could not account for the SSC effects, the boundedness of the movement or the relative unacceptability of parasitic gaps in this construction.

We were therefore led to the assumption that the construction involved NP movement of a PRO, leaving behind an NP-trace (notice that this analysis forces us to the conclusion that retroactive nominals involve two empty categories--an NP-trace and a PRO). This assumption immediately accounted for the bounded nature of the movement and the SSC effects since (1) NP movement leaves behind a trace which is subject to condition A of the Binding theory; and (2) the PRO must be the subject of NP at S Structure, thus accounting for the fact that an overt structural subject is impossible. The set of assumptions we had made forced us to adopt the proposal that government is directional which derived the fact that the subject position of NP is ungoverned.

A final empirical problem was to account for the complex of preposition stranding facts found in retroactive nominals. Only ing nominals allowed this preposition stranding and, furthermore, only a preposition that headed a PP complement to the head noun could be stranded; an prepositional adjunct to the NP remained an island to extraction. We accounted for these facts by assuming that the ECP requires a form of antecedent government. That is, an empty category is properly governed with respect to the ECP only if it is either governed by its antecedent or if it is governed by a lexical head which assigns the position occupied by the empty category a θ-role. Since ing nominals may exceptionally assign a θ-role to their complements, the ECP would be exceptionally satisfied in the retroactive nominal

construction. This formulation of the ECP requires an intimate connection between the theory of movement and thematic theory.

Finally, we showed that many of the fundamental aspects of the analysis of retroactive nominals were compatible with a theory that assumes that governed PRO is possible. The main empirical problems for such a theory are to account for the SSC effects and the preposition stranding facts. Under this theory, the SSC effects do not follow from syntactic constraints on indexing since coindexing the governed PRO with the subject of NP is permissible. This means that the SSC effects must be made to follow from some other component of grammar, such as the pragmatic component. The preposition stranding facts followed under the assumption that θ-marking by the head noun deprived the preposition of the ability to assign abstract Case. Given Bouchard's Principle of Lexicalization, this means that the PRO object of the preposition will be phonologically null. If the PRO were assigned Case, it would, by necessity, be lexicalized. Thus, the thematic conditions that a theory assuming the PRO theorem would place on movement will be placed on Case assignment in a framework assuming a governed PRO. In either event, thematic relations play a central role in governing the distribution of empty categories. The next chapter will be devoted to exploring the nature of the relationship between thematic theory and the distribution of empty categories.

Chapter 3
The Boundary Condition

In Chapter 2, we investigated the properties of what I have been calling retroactive nominals. We attributed the preposition stranding facts, exemplified in (1), to the effects of the Empty Category Principle:

(1)a. This problem$_i$ needs some thinking about $\underline{t}_i$

b. *this problem$_i$ needs some thought about $\underline{t}_i$

In particular, the empty category inside the NP in (1a) is properly governed because the preposition can transmit a thematic role from the head noun to the object of the preposition. This exceptional thematic role transmission was attributed to special properties of the verbal governing the entire NP and to special properties of <u>ing</u> nominals. Thus, the noun, <u>thought</u>, in (1b) may never assign a thematic role to its PP complement and the

preposition, as a result, cannot transmit this thematic role to its object. The empty category in (1b) is, therefore, not governed by a thematic role assigner and the structure violates the Empty Category Principle.

In this chapter, we will probe the hypothesis that the alternation in (1) is attributable to the workings of the Empty Category Principle. We will begin by considering the hypothesis that proper government involves either government by an antecedent or government by a thematic role assigner. We will ultimately reject this view of the Empty Category Principle on the basis of a variety of constructions where an empty category is properly governed by an element that does not directly assign it a thematic role. Nevertheless, we will encounter cases where thematic role assignment conditions the possibility of extraction from a phrase (i.e., extraction of a sub-part of that phrase). We will go on to define a boundary condition on the basis of these examples and show how this boundary condition can provide an account of examples that fall under the Left-Branch Condition, the Constraint on Extraction Domains, the reformulation of Subjacency found in Chomsky (1985) in addition to providing an account of the retroactive nominal cases. Finally, we will compare this boundary condition with certain configurational accounts, in particular, the connectedness framework of Kayne (1984) and the Constraint on Extraction Domains (Huang, 1982).[1]

---

[1] By <u>configurational</u> I mean analyses which make crucial

3.1 Reconsidering the ECP

On an intuitive level, the Empty Category Principle is intended to place a recoverability of deletion requirement on syntactic representations. Recoverability of deletion has long played a role in generative grammar (see, for example, Chomsky (1965)); but, despite its intuitive appeal, the requirement that deletions must be of recoverable material has been notoriously difficult to formalize in any precise way (but, cf. Chomsky & Lasnik, 1977). The introduction of trace theory, however, made it possible to state a constraint that would have the effect of the recoverability of deletion requirement. In essence, trace theory forces syntactic positions to be constant across levels of derivation. Since syntactic positions may not "disappear" in the course of a derivation, but may, rather, contain a phonetically unrealized element, recoverability requirements could be stated as conditions on the identification of empty categories. Furthermore, since positions remain constant, trace theory would allow recoverability requirements to be stated in a purely local way.

The ECP hypothesizes that the presence of an empty category

---

reference to tree structure. Thus, the connectedness framework makes crucial reference to "canonical government configurations" which are defined in terms of tree geometry. I place the Constraint on Extraction Domains in this category because of its crucial reference to proper government, a notion that crucially refers to tree structure.

will always be "signaled" by the local environment around the empty category:

(2) The Empty Category Principle
[e] must be properly governed.

(3) Proper Government
A properly governs B if and only if A governs B and A is lexical.

The ECP as defined in (2) and (3) means that an empty category is only legitimate if it is governed by a lexical element.[2] The question is, of course, what exactly is meant by lexical in definition (3).

Consider contrasts like the following:

(4)a. Who_i did John say t_i saw Mary?
b. *who_i did John saw that t_i saw Mary
(5)a. Who_i did John say Mary saw t_i ?
b. Who_i did John say that Mary saw t_i ?

The examples in (4) involve extraction from the subject position of an embedded clause while the examples in (5) involve the extraction of the object of an embedded clause. In (4a) the extraction is good while the extraction of the subject in (4b) is

---

2 We must, of course, exclude PRO from the ECP if we assume that PRO is obligatorily ungoverned (cf., Stowell, 1985). Chomsky (1981) does this by means of a "generalized ECP" which stipulates that PRO must be ungoverned. Notice that PRO has certain semantic restrictions; for example, it must be interpreted as [+human] in the unmarked case. We might assume, therefore, that PRO has certain lexical features and, as a result, counts as lexical. There are other possible solutions, none of which will be examined here since I assume that the problem is largely mechanical, but see the account of control developed in Chapter 5.

impossible. The examples in (5) are unproblematic given that the trace in both cases is governed by a lexical element, the verb. How can we explain the contrast in (4)? We account for this asymmetry if we suppose that the empty category in (4b) is not properly governed; in other words, the presence of the complementizer, <u>that</u>, in the embedded Comp in (4b) prevents the empty category in subject position from being properly governed, as in the account in Chomsky (1981) and Pesetsky (1981).

The subject of tensed clauses is governed by Infl; it is arguably the case that Infl is not sufficiently lexical for purposes of the ECP. If this is correct, then proper government of a trace in subject position of a tensed clause must be by an element that is external to S. Given the subjacency constraint, we know that the (4a) must have an S Structure similar to:

(6) $[_{\bar{S}}$ who$_i$ $[_S$ John say $[_{\bar{S}}$ $\underline{t}_i$ $[_S$ $\underline{t}_i$ saw Mary]]]]

The embedded Comp in (6) contains a trace left by successive cyclic movement of the wh-element from the embedded subject position. The S Structure representation of (4b) looks like:

(7) $[_{\bar{S}}$ who$_i$ $[_S$ John say $[_{\bar{S}}$ $\underline{t}_i$ that $[_S$ $\underline{t}_i$ saw Mary]]]]

Notice that the trace in Comp in (6) may c-command (and, hence, govern) out of Comp since the first branching node that dominates the trace is $\bar{S}$. In (7), however, the presence of both the trace and the complementizer forces the Comp to branch; hence, neither the complementizer nor the trace may govern out of Comp. The

trace in the embedded subject position of (7) is, therefore, not properly governed since its only governor, Infl, is not sufficiently lexical to count as a proper governor.

In (6), however, we can assume that the trace in the embedded subject position is governed both by Infl and by the trace in Comp. We have already ruled out Infl as a proper governor for the empty category since Infl is not sufficiently lexical to count as a proper governor. But surely the trace in Comp is no more lexical than Infl; how can it be, then, that the trace in Comp can properly govern the trace in subject position? Notice that the trace in Comp not only governs the trace in embedded subject position but also binds it. It is not unnatural to assume that binding satisfies recoverability of deletion. Hence, we can extend the definition of proper government to include binding:

(8) A properly governs B if and only if A governs B and either:
  (i) A and B are coindexed;
  or:
  (ii) A is lexical.

Despite the introduction of binding into the definition of proper government, we still must define how lexical is to be interpreted in (8). Consider, in this light, the following contrasts:

(9)a. [For whom]$_i$ did he vote t_i ?
  b. Who$_i$ did he vote for t_i ?
(10)a. [pour qui]$_i$ a-t-il voté t_i ?
  b. *qui$_i$ a-t-il voté pour t_i

The examples in (9) show that prepositions may count as proper governors in English. As noted in Kayne (1981), prepositions in French, however, do not count as proper governors as the examples in (10) illustrate. It is not entirely transparent how to interpret the meaning of lexical in (8); it seems strange to claim that prepositions are more lexical in English than in French. We may suppose, as is standard, that lexical in (8) is replaced by a variable, X^0, which ranges over the lexical categories. Whether or not a particular lexical category counts as a proper governor would then be a matter of parametric variation and particular instantiations for X would be subject to markedness constraints.

This approach to the ECP poses a number of problems. It is a matter of some interest, for example, to determine the possible range of variation for the instantiations of X in (8). On the face of it, we might expect to find languages which, for example, allowed only nouns and prepositions to be proper governors. A related question is whether or not it is possible for nonlexical categories to instantiate X. Huang (1982) argues that Infl is a proper governor in Chinese since it lacks the subject/object asymmetries illustrated for English in the discussion of (4) and (5) above. Given that the full range of possible instantiations for X has been determined, our next task would be, of course, to try to explain the range of categories involved. Why should it be that verbs and adjectives are always proper governors while

prepositions and Infl are only proper governors in some languages? A possible approach would be to develop a markedness theory that would, for example, derive the fact that some categories are always proper governors.

A second matter of interest is the disjunction in (8). That is, an element may be properly governed if it is governed by its antecedent or if it is governed by an element of some particular category. That the antecedent of a gap should play some role in licensing that gap has a certain intuitive appeal. Much recent work has exploited the relationship between antecedents and gaps to derive the ECP, although this relationship was implicit in the classical ECP of Chomsky (1981). For example, Path Theory, developed in Pesetsky (1982) and the Connectedness framework of Kayne (1984) both rely on aspects of tree geometry to license the antecedent/gap relationship. The version of the ECP developed in Lasnik & Saito (1984) explicitly marks empty categories according to whether they have an antecedent. Finally, Aoun (forthcoming) reduces the ECP to the Binding theory, a move which would eliminate clause (ii) of (8) entirely. Definition (8) poses the following riddle: How is government by a lexical element like government by an antecedent?

In the first part of this chapter, we tried to reduce the definition of proper government to clause (i) of (8). An empty category counted as properly governed just in case it was governed by a coindexed element. This coindexation could arise

in either of two ways. First, the empty category could be bound by its governor. This is the case for subject extraction as illustrated in examples (4) and (5) above. Second, the empty category could be governed by an element which directly assigns it a θ-role. Since direct θ-role assignment implies cosuperscripting between the role assigner and the element which receives the role, this case reduces to government by a coindexed element, as in Stowell (1981) and the references cited there. The parametric variation found in the class of proper governors would, thus, be an artifact of parametric variation in θ-theory. For example, we have been assuming that in the unmarked case languages do not allow prepositions to directly θ-mark their objects. English is odd in that it allows direct θ-role transmission by prepositions. The intent of this is, of course, to eliminate clause (ii) of (8) while still deriving some of its effects.

Before exploring some of the effects of this version of the ECP, let us consider in more detail the nature of the cosuperscripting induced by θ-role assignment. In particular, we want to show that this notational device does, in fact, reflect a linguistically significant relation and is not merely a post facto device designed to simulate some extraction asymmetries.

Let us identify the following three types of configurations:[3]

---

[3] Recent work on various structural configurations of some

(11)a. $\ldots[_{\bar{X}}\ X\ YP]\ldots$

b. $\ldots[_{XP}\ \ YP\ [_{\bar{X}}\ldots$

c. $\ldots[_{XP}\ \ XP\ YP]\ldots$

The configuration in (11a) is a configuration where X subcategorizes for and θ-marks YP. The configuration in (11b) is a structure of predication where $\bar{X}$ (or a maximal projection contained therein) predicates a role for YP (see Williams (1980) for some discussion of the predication relation). Finally, the configuration in (11c) is an adjunction structure where YP is an adjunct on the head, XP.

The configuration in (11a) is one which permits direct θ-role assignment by the head, X. Following Stowell (1981), let us assume that the thematic roles of the head are organized into a thematic grid. Each thematic role is associated with a position in the thematic grid. We will further assume that direct θ-role assignment is actually a rule which copies the index from the position "assigned" the thematic role to the corresponding slot on the thematic grid. This copying rule will allow us to dispense with the cosuperscripting mechanism entirely since cosuperscripting repeats the information that can be found by inspecting the index on a position and the thematic grid of the governing head. Recall that cosuperscripting was crucial for

---

linguistic significance includes Marantz (1984), Stowell (1981), Williams (1980) and Williams (1984).

government across a maximal projection; we may restate the condition for transparency of government as follows (cf., Kayne (1984) and Stowell (1985)):

(12) X^n is transparent to government by Y^n if and only if Y^n bears the same index as X^n.

By the above conventions, thematic role assignment by Y^n to X^n will result in Y^n bearing the same index as X^n; it follows that X^n will be transparent to government by Y^n. Similarly if Y^n binds X^n, then X^n will be transparent to government by Y^n.

As an example of transparency of government (and, hence, proper government) under coindexation, let us consider the problem of <u>en</u> cliticization in French. Following Kayne (1975), we will assume that clitics are dominated by the verb (i.e., V^0). Given the system developed in Borer (1984), we will take this as implying that a clitic may govern complements to the verb. Consider a contrast like the following:

(13)a. J´en ai trois
´I have three of them´
b. *j´ai trois

Let us assume, with Kayne (1975, 1984), that the structure of <u>trois</u> involves an empty category:

(14) [[trois] <u>e</u>]
NP Q

Finally, following work found in Belleti & Rizzi (1981) and Jaeggli (1984b), let us assume that this empty category must be identified and that this identification may taken place only under government by a coindexed element. We thus have a fairly

simple account for the contrast found in (13). In (13a), the NP, [[e] de livres], is governed and assigned a thematic role by
NP Q
the verbal complex, [en ai], which induces coindexation between
V
the two elements. The maximal projection, NP, is, as a result, transparent to government by the verbal complex. This implies that the clitic, en, may identify the empty quantifier, since en is part of the verbal complex which governs into the NP. In (13b), the verb may govern into the object NP by virtue of coindexation under thematic role assignment. The verb, however, does not include the clitic en; hence, the empty quantifier will not be identified and the example is, as a result, ruled out by the identification requirement.

Now, consider the contrast in (15) and (16):

(15)a. La première partie de ce livre rappelle les romans policiers.
'The first part of this book recalls detective novels.'

b. *la première partie en rappelle les romans policiers
('The first part of it recalls detective novels.')

(16)a. Ce roman m'en rapppelle la première partie de son histoire.
'This novel reminds me of the first part of his story.'

b. Ce roman m'en rappelle la première partie.
'This novel reminds me of the first part of it.'

The facts in (16) are fairly easy to account for under the assumption that the clitic, en, identifies an empty category inside the object NP. Once again, since the verb assigns a thematic role to the object NP, the clitic will be able to govern this empty category despite the intervening maximal projection,

NP. In (15), the subject contains an empty category which must be identified by en. But the subject is not governed by the verb; hence, the empty category may not be identified by the clitic and the example is ruled out.

Our set of assumptions would seem to rule out any association between en and the subject position. This is not the case:

(17)a. La forme de ce poème n´est pas admirable.
´The form of this poem is not admirable.´
b. La forme n´en est pas admirable.
´The form of it is not admirable.´

In general, en may be associated with the subject of predicative verbs like être (Kayne (1975)). If we adopt the rule of predication of Williams (1980) which coindexes the subject and the predicate in certain structures, then (17b) will be associated with the following representation:

(18) [la forme [e]$_i$] [n´en est pas admirable]$_i$

In (18), the subject and the predicate are coindexed; hence, the subject is transparent to government by the predicate. Given the assumption that Infl is the head of the predicate and the n´en est is dominated by Infl, it is reasonable to assume that en may identify the empty category inside the subject NP.

Extending the discussion a bit further, consider a contrast like:

(19)a. Il en a connaît l´auteur.
´He knows the author of it.´
b. *il en a téléphoné à l´auteur
(´He telephoned the author of it.´)

Recall that we have assumed that prepositions do not directly assign a thematic role to their complements. Hence, the object of the preposition, à, in (19b) will not be transparent to government. The prepositional object in (19b) will thus not be transparent and there is, therefore, no way to identify the empty category inside the NP.

We have tried to show that the assumption that the ECP is tied in with thematic theory can provide a foundation for an interesting theory of the distribution of empty categories that goes some way in deriving certain asymmetries involving subjects versus objects, nominal and preposition complements versus verbal and adjectival complements and arguments versus adjuncts. We will now go on to show that our framework is, on the one hand, too strong and, on the other hand, perhaps a bit too weak. From this, we will factor the constraints that thematic theory places on the distribution of empty categories out of the ECP and into a separate component.

Let us first consider cases where the theory developed up to this point is too strong; that is, cases where legitimate extractions are incorrectly ruled out. Along the way, we will note certain suspicious parallelisms that we will later use to justify the assumption that our theory is also too weak.

Recall that, for our purposes, proper government consists solely in government by a coindexed category, as in Stowell (1981). Thematic theory enters into the picture because θ-role

assignment allows a head to copy the index of its complement into a thematic grid and is, therefore, coindexed with the complement. Consider, first, the structure of small clauses like:

(20)a. John considers [Bill a fool]
b. John saw [Mary leave]
c. John heard [the men in the room]

Following Chomsky (1981) and Stowell (1981, 1983), we will assume that the bracketed phrases in (20) form a constituent with the left-most NP in each acting as the subject of that constituent. This assumption accounts for the fact that subject of a small clause obeys the Subject Condition (or Left Branch Condition) which prohibits the extraction of a subpart of a subject (cf., Kayne, 1984):

(21)a. *of whom does John consider [[friends t] fools]
b. *of whom did John see [[friends t] leave]
c. *of whom did John hear [[relatives t] in the room]

Notice that the examples in (21) follow from our account of proper government and thematic role assignment. The head noun may not directly assign a θ-role to its object; hence, the head noun will not contain the index of its complement in its thematic grid. If this is so, then the head noun may not properly govern its complement.

The right-most constituent in each of the examples of (20) is the head of the bracketed phrases and, furthermore, assigns a θ-role to the subject. Given these assumptions, it follows that the verb may not assign a θ-role to the subject of a small clause since the subject would then bear two thematic roles. (Note that the verb assigns a thematic role to the entire small clause,

however.) The system developed up to this point makes a very clear prediction about extraction of the subject of a small clause: Since the verb does not assign the subject a θ-role and since it has no antecedent internal to the small clause, extraction of the subject should be impossible. Given our assumptions about transparency of maximal projections, this prediction is patently false:

(22)a. Who_i does John consider [$\underline{t}_i$ a fool]

b. Who_i did John see [$\underline{t}_i$ leave]

c. What_i did John hear [$\underline{t}_i$ in the room]

(23)a. Bill_i is considered [$\underline{t}_i$ a fool] [4]

b. *Mary_i was seen [$\underline{t}_i$ leave]

c. The men_i were heard [$\underline{t}_i$ in the room]

The examples in (22) and (23) show that the subject of a small clause may be extracted even though it is not governed by a coindexed element. Notice that if we return to the definition of the ECP which requires either government by an antecedent or government by a lexical element then the facts in (22) and (23) follow immediately. The verb assigns a θ-role to the small

---

4 I have no account for the impossibility of (23b). Notice that <u>see</u> does admit passive:

(i) Mary_i was seen $\underline{t}_i$

so it cannot be because <u>see</u> does not have a passive participle. Nor is the impossibility of passive a general property of perception verbs, as example (23c) shows.

clause which is, as a result, transparent to government by the verb. Thus, the verb will govern the subject of the small clause. Since verbs are the paradigmatic proper governors, it follows that the subject position of the small clause will be properly governed by the verb. This has the result, however, that the Subject Condition must stand as an independent principle to rule out the examples in (21) since the head noun in each of the examples may properly govern the empty category inside the subject.

In a similar vein, consider the class of Exceptional Case Marking structures:

(24)a. John believes [$_S$ Bill to be the biggest fool in the class]

b. John considers [$_S$ Mary to be the most talented student]

c. The men proved [$_S$ John to be incapable of building an acceptable set of shelves]

Each of the underlined NPs in the examples in (24) is assigned a θ-role internally to the embedded S. Notice that the matrix verb places no selectional restrictions upon that position as can be seen in (25) which show that pleonastic it and the reserved element there may occur in this position:

(25)a. John believes [$_S$ it to be raining]

b. The men proved [$_S$ there to be severe nutritional problems among the lower economic classes]

We will, therefore, assume that the position in question is the subject position of an embedded clause. Notice that this position has no way of receiving Case within the embedded clause since the subject may receive nominative Case from a tensed Infl and the clauses in question are infinitivals. The subject must receive Case from an element which is external to the embedded clause. Following standard practice, we will assume that the embedded subjects in this class of structures receives accusative Case from the verb of the superordinate clause, see the analysis of Exceptional Case Marking in Chomsky (1981). Thus, Bill in (24a) will receive accusative Case from the verb believes.

If the superordinate verb assigns Case to the embedded subject then it must be true that the verb governs this position since Case assignment transpires under government. This is perfectly consistent with our assumptions about transparency of government since the superordinate verb assigns a θ-role to the embedded clause; the clause boundary is, therefore, transparent to government by the verb. Exceptional Case Marking structures, like small clause structures, represent instances wherein a verb may govern, but not assign a θ-role to, a particular position.

The version of proper government developed in this chapter clearly predicts that extraction of the embedded subject should be impossible since in Exceptional Case Marking structures. The empty category left in the embedded subject will not be governed by its antecedent, since the embedded clause, under standard assumption lacks a Comp which will contain a trace that can

antecedent govern the empty category. Nor is the empty category governed by an element which has its index as an element in its thematic grid, since its governor, the superordinate verb, does not assign it a θ-role. As is the case with small clauses, this prediction is not viable:

(26)a. $[_{\bar{S}}$ who$_i$ $[_S$ does John believe $[_S$ $\underline{t}_i$ to be the biggest fool in the class]]]

b. $[_{\bar{S}}$ who$_i$ $[_S$ does John consider $[_S$ $\underline{t}_i$ to be the most talented student]]]

c. $[_{\bar{S}}$ who$_i$ $[_S$ did the men prove $[_S$ $\underline{t}_i$ to be incapable of building an acceptable set of shelves]]]

(27)a. Bill$_i$ is believed $[_S$ $\underline{t}_i$ to be the biggest fool in the class]

b. Mary is considered $[_S$ $\underline{t}_i$ to be the most talented student]

c. John was proven $[_S$ $\underline{t}_i$ to be incapable of building an acceptable set of shelves]

The examples in (26) and (27) show that both wh-movement and NP-movement are possible from the embedded subject position of an Exceptional Case Marking structure. These facts will follow immediately in a framework which allows proper government by a lexical element, whether θ-marker or not. Notice, furthermore, that, like the subject of a small clause, the embedded subject of an Exceptional Case Marking structure falls under the Subject Condition:

(28)a. $[_{\bar{S}}$ who$_i$ $[_S$ does John believe $[_S$ [a friend of $\underline{t}_i$] to be the biggest fool in the class]]]

b. $[_{\bar{S}}$ who$_i$ $[_S$ does John consider $[_S$ [relatives of $\underline{t}_i$] to be the most talented students]]]

c. $[_{\bar{S}}$ who$_i$ $[_S$ did the men prove $[_S$ [students of $\underline{t}_i$] to be incapable of building an acceptable set of shelves]]]

Like the subjects of small clauses, the embedded subjects of Exceptional Case Marking structures are islands to extraction.

Consider, now, Raising structures as in:

(29)a. Bill$_i$ seems $[_S$ $\underline{t}_i$ to be the biggest fool in the class]

b. Mary$_i$ appears $[_S$ $\underline{t}_i$ to be the most talented student]

c. John$_i$ is certain $[_S$ $\underline{t}_i$ to be incapable of building an acceptable set of shelves]

The superordinate verbal in each of the examples in (29) may not assign a θ-role to the embedded subject position since that position receives a θ-role internal to the embedded clause. Furthermore, the verbal may not assign a θ-role to the raised subject since such θ-role assignment would violate the θ Criterion which forbids a single argument from receiving two thematic roles; the raised subject receives its θ role by virtue of forming a chain with the trace in the embedded subject position. This analysis receives support from the fact that pleonastic <u>it</u> appears when Raising fails to apply:

(30)a. It seems [$_S$ Bill is the biggest fool in the class]

b. It appears that [$_S$ Mary is the most talented student]

c. It is certain that [$_S$ John is incapable of building an acceptable set of shelves]

As the examples in (29) show, extraction is possible in Raising structures even though the resulting empty category is not antecedent governed either by its own antecedent (due, under our assumptions, to the presence of the intervening VP) or by a θ-marker. Once again, a definition of proper government which requires government by an antecedent alone cannot account for the facts in (29). If the definition of proper government is broadened to include government by a lexical category, then the facts follow trivially; the superordinate verbal assigns a θ-role to the embedded clause which is, therefore, transparent to government by the superordinate verbal. If this is so, then the empty category in the embedded subject position is governed by the superordinate verbal, as required by the definition of proper government which allows government by a lexical category as well as government by an antecedent.[5]

---

5 Illustrating the effects of the Subject Condition is more difficult in Raising structures than in Exceptional Case Marking structures and small clause structures. This is because an NP must receive abstract Case, which is unavailable in Raising structures because the embedded clause is infinitival (hence, nominative Case will not be assigned) and the superordinate verbal will not assign accusative Case.

Finally, let us consider the problem of traces in Comp which result from successive cyclic movement of a element:

(31)a. $[_{\bar{S}}$ who$_i$ $[_S$ did John say $[_{\bar{S}}$ t_i $[_S$ t_i´ saw Mary]]]]

b. $[_{\bar{S}}$ what$_i$ $[_S$ does John believe $[_{\bar{S}}$ t_i $[_S$ t_i´ hit Bill]]]]

The embedded Comp in both of the examples in (31) contains a trace left by successive cyclic movement of a wh-element. Notice that in both instances the trace in the intermediate Comp is independently necessary to antecedent govern the trace in the embedded subject position. Since the trace in the embedded subject position is antecedent governed, we may assume that it is properly governed and passes the ECP. What of the trace in the embedded Comp position? Notice that this trace is not assigned a θ-role by the superordinate verb. If it were assigned a θ-role by the superordinate verb, then the position which it occupies would count as an A-position. But then the Comp and the embedded subject position would form an A-chain which bears two distinct θ-roles. If the trace in the embedded Comp is not assigned a θ-role by the superordinate verb, then this verb may not, by our assumptions, count as its antecedent governor. Notice that we cannot say that the element in the superordinate Comp may count as an antecedent governor for purposes of the ECP given the well-formedness of:

(32) $[_{\bar{S}}$ who$_i$ $[_S$ did John say $[_{\bar{S}}$ t_i that $[_S$ Mary said $[_{\bar{S}}$ t_i´ $[_S$ t_i" saw Bill]]]]]]

If a trace in a superordinate Comp may act as an antecedent governor for a trace in an embedded Comp (as in Lasnik & Saito, 1984), then we would expect (32) to be an ECP violation. The trace in the lowest Comp, $\underline{t}'$, could not be antecedent governed by the trace in the superordinate Comp since the latter trace cannot c-command out of Comp due to the presence of the complementizer <u>that</u>. We may assume, then, that a trace in Comp is neither antecedent governed (although it may be antecedent bound) by an element in a superordinate Comp nor is it governed by an element that assigns it a θ-role (and is, therefore, coindexed with it).

Notice that in each of the cases in (31) and (32) the embedded clause is θ-marked by a verb. Hence, the $\bar{S}$ containing the trace in Comp will be transparent to government by a verb. Given the old definition of proper government, the trace in Comp would be lexically governed by the superordinate verb and would, as a result, pass the ECP. Notice that, as in the case of small clauses and Exceptional Case Marking structures, elements in Comp show Subject Condition effects; a subpart of a phrase that has been moved to Comp may not be extracted:

(33)a. $*[_{\bar{S}}$ who$_i$ $[_S$ does John think $[_{\bar{S}}$ [to $\underline{t}_i$] $_j$] $[_S$ Mary gave the money $\underline{t}_j$]]]]

b. $*[_{\bar{S}}$ who$_i$ $[_S$ did John say $[_{\bar{S}}$ [with $\underline{t}_i$] $_j$ $[_S$ Mary went to the film $\underline{t}_j$]]]]

Once again, a system that includes both lexical government (without reference to θ-marking) and some form of the Subject

Condition appears to make the right distinctions; traces in Comp would be lexically governed by a verb or adjective and extraction of a subpart of a phrase in Comp would be ruled out by the Subject Condition.

Let us suppose then that we return to a system that includes lexical government as part of the definition of proper government and some form of the Subject Condition. In order to be properly governed, it is sufficient to be governed by a lexical category. Which categories count as lexical categories will be determined by fixing certain parameters in the definition of proper government, presumably with some markedness conditions placed on these parameters.[6] Thus, no language will have only prepositions as proper governors without also having verbs as

---

6 Recent work by Ken Wexler (talk given at UCLA; 1985) suggests that markedness theory could be derived from the Subset Condition. This condition requires that, given two possible settings, A and B, of some parameter, P, such that the language generated by setting P with A is a subset of the language generated by setting P with B, the initial setting of the parameter will be with A. If the parameter were set with B, then there can be no evidence to force setting the parameter with A since every structure compatible with A is compatible with B (but not vice versa). It isn't obvious, at this point, how to derive a markedness hierarchy for proper government based on the Subset Condition. It is clear, however, that a language generated by allowing only verbs to be proper governors is a subset of a language in which both verbs and prepositions are proper governors. Thus, we may independently set up a hierarchy (based on empirical evidence) of categories: V < P <..., where '<' means "less marked than". We predict that we cannot find the following situation: In language X, prepositions but not Infl is a proper governor while in language Y, Infl but not prepositions is a proper governor. In this case, X is not a subset of Y and Y is not a subset of X; thus, neither language is more marked than the other language.

proper governors. Presumably, the "weaker" (in some as yet to be established definition of "weak") the thematic assigning properties of a category are the more marked that category is as a proper governor. Since prepositions only "weakly"--in the sense that they either transmit a θ-role assigned to them from a head or have their argument structure checked at LF-- assign a θ-role to their objects, prepositions are a highly marked category with respect to the definition of proper government. Clearly, this approach is, at this point, only worked out at a very impressionistic level.

Under this approach, prepositions are proper governors across the board in English rather than being proper governors relative to their status with respect to a head (as in the approach to the ECP we adopted earlier). The extraction facts we have considered will, as a result, have to be handled by some fairly diverse mechanisms in the grammar. As pointed out, above, the Subject Condition effects in (34) will be ruled out by some independent principle of core grammar (possibly subjacency):

(34)a. *who$_i$ did [friends of t_i] surprise John

b. *who$_i$ does John believe [[relatives of t_i] to be incompetent]]

c. *who$_i$ does John consider [stories about t_i] offensive

Extraction from an adjunct, as in (35), may be handled either by subjacency or by a combination of subjacency and a

filter that rules out traces marked with oblique Case (as in Hornstein & Weinberg (1981), see the discussion, above):

(35)a. *which inning$_i$ did the Yankees lose the ball game in $\underline{t}_i$

b. *which critical paper$_i$ did you write a review of Walter´s new novel [without consulting $\underline{t}_i$]

c. *where$_i$ did John get lost [walking to $\underline{t}_i$]

We could rule out the above examples if, following Chomsky (forthcoming), we assume that any category may be a bounding node with respect to subjacency if that category is not a complement of some head. Since adjuncts are by definition not complements of a head, they will always count as bounding nodes, regardless of the category of the adjunct. In the examples in (35), the wh-element had to cross the matrix S and the maximal projection of the adjunct, both of which will count as bounding nodes, in order to land in Comp; hence, each of the examples will count as subjacency violations.

Finally, there is the problem of the retroactive nominals discussed at length in the first half of this chapter. Recall that the structure of a well-formed example of these nominals will be as in (36a) while an ill-formed example is shown in (36b):

(36)a. ...needs [PRO$_i$ some [working [on $\underline{t}_i$]]]...
N" N´ PP

b. *...needs [PRO$_i$ some [discussion [about $\underline{t}_i$]]]...
N" N´ PP

Both of the traces in (36a) and (36b) could be properly governed since prepositions act as lexical governors in English; hence, the ungrammaticality of (36b) does not follow from a violation of the ECP. Consider the status of (36a) with respect to the version of subjacency outlined above. The verb that governs the nominal, need, licenses thematic role assignment by the ing nominal, working, to the prepositional phrase headed by on. This thematic role assignment licenses transmission of the thematic role assigned to the PP to the object of the preposition. Since the status of a maximal projection as a bounding category is determined by thematic role assignment, the PP will not be a bounding node with respect to the NP in (36a). Hence, no bounding node lies between the NP-trace and its antecedent in (36a) and the derivation is, thus, well-formed as it does not violate subjacency.

Let us now consider the status of (36b) with respect to subjacency, in particular, the version of subjacency proposed in Chomsky (1985). The governing verb, need, does not license thematic role assignment by the head noun, discussion, since this noun is not an ing nominal. Hence, the noun will not directly assign a thematic role to its complement PP. If this is so, then the head of the PP, about, may not transmit a thematic role to its object. Now, the PP will count as a bounding node with respect to extraction of its object. Notice, however, that the only maximal projection that intervenes between the trace and its

antecedent is the maximal projection of the preposition. Hence, only one bounding node is crossed in the derivation of (36b). Since a subjacency violation requires that at least two bounding nodes be crossed, example (36b) is not a violation of subjacency.[7] The ill-formedness of (36b) is not a consequence either of the ECP or of subjacency; we must make an appeal to some principle other than the ECP if we are to successfully filter out the derivation of (36b).

Let us take a moment to consider the results of this section. We began by attempting to reduce proper government to some notion of government by an antecedent (i.e., government by a category coindexed with the position in question). This was intended both to derive certain facts regarding retroactive nominals and, more generally, to eliminate a disjunction in the definition of proper government. This latter point brought the definition of proper government in line with a (possibly incorrect) intuition that recoverability of deletion involves a close thematic connection between an empty category and the element that signals its presence. This reformulation of proper government, despite whatever intuitive appeal it may have, is too strong in that it rules out a number of well-formed derivations involving extraction of a subject in small clause, Exceptional

---

7 One could generalize the definition of subjacency in such a way as to make any projection, including non-maximal projections, count as bounding nodes under the appropriate circumstances. As I will argue below, this generalization of subjacency is too strong insofar as it rules out well-formed cases of extraposition

Case Marking Structures, Raising structures and, in addition, it seemed to rule out successive cyclic movement of wh-elements. While a return to the classical definition of proper government (which involves both antecedent government and lexical government) correctly derived the previously mentioned class of structures, we were forced to make appeal to some version of the Subject Condition (perhaps in the guise of a modified formulation of Subjacency which makes reference to thematic theory). This position, however, fails to derive certain facts about retroactive nominals. In order to derive these facts, we must make appeal to some as yet unknown principle of grammar. Thus, the facts derived by the reduction of proper government to antecedent government means that the explanation of Subject Condition effects and retroactive nominals will no longer be a unified explanation. It should be noted that I will continue to assume that the ECP and subjacency are independently necessary, but I will assume that many of the facts discussed in this chapter do not necessarily fall under these constraints. In the next section, I will attempt to unify these facts by making appeal to thematic theory.

3.2 Some possible thematic effects

In the previous section we saw that thematic theory constrains subjacency, but not, apparently, proper government. Recall, however, that thematic theory does affect the definition of government itself, since a maximal projection is transparent

to government under conditions of thematic marking. Furthermore, thematic theory appears to constrain the representation of what I have been calling retroactive nominals. Recall, also, that this latter effect of thematic marking does not follow from subjacency. Finally, we retreated to a classical ECP account of certain extraction effects; this move leaves us with the riddle of why some categories are sufficiently "lexical" to act as proper governors while other categories are not. What, exactly, is the content of "lexical" in the definition of proper government?

Throughout the preceding discussion, I have been tacitly assuming the thesis that thematic theory acts as an important constraint on movement. This thesis is based on facts like the following:

(37)a. *which manual$_i$ did you fill out the forms [without reading $\underline{t}_i$]

b. *this issue$_i$ needs [PRO$_i$ some discussion about $\underline{t}_i$]

Example (37a) involves extraction out of an adjunct, which, presumably, is not θ-marked and (37b) involves extraction out of a PP which, again, is not θ-marked. A survey of the various structures relevant to thematic theory yields as number of correlations.

Let us, first, consider the case where an element is directly θ-marked by a head, for example, when an element is a subcategorized complement of a verb. As is well known,

extraction from a θ-marked position is quite free, keeping the effects of subjacency constant and ignoring, for the moment, certain non-bridge verbs:

(38)a. [$_{\bar{S}}$ who$_i$ [$_S$ did John see $\underline{t}_i$]]

b. [$_{\bar{S}}$ who$_i$ [$_S$ did Mary say [$_{\bar{S}}$ [$_S$ John saw $\underline{t}_i$]]]]

c. [$_{\bar{S}}$ who$_i$ [$_S$ does Bill think [$_{\bar{S}}$ [$_S$ Mary said [$_{\bar{S}}$ [$_S$ John saw $\underline{t}_i$]]]]]]

As examples (38b) and (38c) show, extraction out of a subcategorized complement is also quite free. This fact holds constant across the categories of direct θ-role assigners; thus, extraction out of a complement to an adjective is also quite free:

(39)a. [$_{\bar{S}}$ who$_i$ [$_S$ is John$_j$ [$_{Adj''}$ most eager [$_{\bar{S}}$ [$_S$ PRO$_j$ to impress $\underline{t}_i$]]]]]

b. [$_{\bar{S}}$ who$_i$ [$_S$ could [eating these oats] be [$_{Adj''}$ good [$_{PP}$ for $\underline{t}_i$]]]]

Notice that the category of the subcategorized complement does not appear to cause changes in the extraction facts. Not only are clauses and prepositional phrases transparent to extraction when they occur in a θ-marked position, but indefinite NPs are also quite transparent:

(40)a. [$_{\bar{S}}$ who$_i$ [$_S$ did you give [$_{N''}$ pictures of $\underline{t}_i$] to Mary]]

b. [$_{\bar{S}}$ which president$_i$ [$_S$ did John write [$_{N''}$ a shocking exposé on $\underline{t}_i$]]]

c. [$_{\bar{S}}$ which public figures$_i$ [$_S$ does John enjoy [$_S$ PRO writing [$_{N''}$ off-color jokes about $\underline{t}_i$]]]]

As (40c) shows, the extraction site may be quite deeply embedded.

Directly θ-marked positions are, then, quite transparent to extraction. Positions that receive their thematic roles by means of transmission of that thematic role (i.e., object of subcategorized prepositions) begin to show some opacity to extraction:

(41)a. ??[$_{\bar{S}}$ who$_i$ [$_S$ did you give some pencils [$_{PP}$ to [$_{N''}$ friends of $\underline{t}_i$]]]]

b. *[$_{\bar{S}}$ of whom$_i$ [$_S$ did you give some pencils [$_{PP}$ to [$_{N''}$ friends $\underline{t}_i$]]]]

Notice, however, that the complement of the preposition, itself, may be freely extracted:

(42) [$_{\bar{S}}$ who$_i$ [$_S$ did you give pencils [$_{PP}$ to $\underline{t}_i$]]]

---

8 There is a dramatic contrast in grammaticality between (41a) and (41b). This contrast appears to be the result of <u>of</u> insertion. Thus, I also find a small contrast in the following examples:

(i)a. ?of whom$_i$ did you take [$_{N''}$ pictures $\underline{t}_i$]

b. who$_i$ did you take [$_{N''}$ pictures of $\underline{t}_i$]

I have, however, no clear hypothesis about the mechanism that conditions these contrasts nor why the contrast between (41a) and (41b) is so great while the contrast between (ia) and (ib) is so small.

It appears, then, that thematic role transmission sets up a weaker relationship between head and complement than direct thematic role assignment, although this difference must be made precise.

We turn, now, to environments of predication, like the subject of S or the subject of a small clause. We have already discussed some Subject Condition effects; casting this condition in slightly different terms, we can say that the subject of a predication is an island to extraction:

(43)a. *[$_{\bar{S}}$ who$_i$ [$_S$ did [$_{N''}$ relatives of $\underline{t}_i$] walk into the room]]

b. *[$_{\bar{S}}$ of whom$_i$ [$_S$ did [$_{N''}$ relatives $\underline{t}_i$] walk into the room]]

c. *[$_{\bar{S}}$ who$_i$ [$_S$ do you consider [$_{Adj''}$ [$_{N''}$ relatives of $\underline{t}_i$] obnoxious]]]

d. *[$_{\bar{S}}$ of whom$_i$ [$_S$ do you consider [$_{Adj''}$ [$_{N''}$ relatives $\underline{t}_i$] obnoxious]]]

The "connectedness" framework of Kayne (1984) attributes the ungrammaticality of structures like those in (43) to the fact that the extraction is of a subpart of a left-branch. The connectedness ECP requires that a "canonical government configuration" be projected from the site of extraction to the landing site of the extraction. Kayne argues that the canonical government configuration for English has the governor on the left and the governed element on the right. Elements inside a left

branch may not project up the tree since the governor of the left branch is on the right, which is the opposite of the canonical government configuration for English. This approach to Subject Condition effects has the result of making left-branches absolute barriers to extraction.

While a left-branch does limit the extraction possibilities, it does not appear to be true that left-branches are absolute barriers to extraction. In particular, consider pairs of sentences like:

(44)a. $[_{NP}$ many books $[_{\bar{S}}$ which I've enjoyed reading]] have been published recently

b. $[_{NP}$ many books $\underline{t}_i$] have been published recently $[_{\bar{S}}$ which I've enjoyed reading$]_i$

Assuming with Gueron & May (1984) that the relative clause on the subject NP in (44a) has been extraposed by "move alpha", then it is not the case that left-branches are absolute barriers to movement. The rightward movement of a subpart of a left-branch is, nevertheless, strictly bounded; the movement is bounded by the first clause-like boundary (cf., the discussion of the Right Roof Constraint in Ross (1967))[9]:

---

9 Because of its boundedness, extraposition from subject position does not, however, create a problem for Kayne's connectedness ECP. The set of nodes projected up from the empty category must be able to "connect" with the antecedent. In extraposition from subject structures, the set of nodes projected from the empty category will stop at the subject NP. The node that immediately dominates the extraposed element is, presumably, S. Since S also immediately dominates the subject NP, the set of

(45) *the man $[_{\bar{S}}$ who $[_{S}$ claimed that $[_{NP}$ many books $\underline{t}_i]$ have been published recently]] died $[_{\bar{S}}$ which he has enjoyed reading$]_i$

This extraposition from subject is not limited to the subject of full clauses; extraposition may apply from the subject of a small clause:

(46)a. John considers [[many books $[_{\bar{S}}$ which I've enjoyed reading]] foolish]

b. John considers [[[many books $\underline{t}_i$] foolish] $[_{\bar{S}}$ which I've enjoyed reading$]_i$]

A subject of predication is, then, not an absolute barrier to extraction, although the extraction is of an extremely local nature (as noted in Rizzi, 1982). Let us refer to S and small clauses as predication structures; we may state the following descriptive generalization:

(47) A subpart of a subject may not be extracted from a predication structure containing that subject.

Let us also note, if only for the record, that subjects of predication and elements that receive their thematic role by means of "transmission" (the objects of subcategorized prepositions) seem to class together in the sense that it is difficult to extract a subpart of these elements.

---

nodes dominating the empty category and the node dominating the antecedent connect, as required. See Longobardi (1985) for some discussion.

The revised version of subjacency noted above would seem to capture these facts in a fairly direct fashion. Assuming that the subject NP is always a bounding node for subjacency and that, as a result, S is a bounding node for elements extracted from within the subject NP, then it follows that movement of a subpart of a subject NP will violate subjacency since such a movement would have to traverse two bounding nodes. Simple adjunction to S, as in extraposition from subject, will cross only one bounding node--the subject NP. Hence, extraposition from subject will always be strongly bounded by S. Since the algorithm for determining whether a particular maximal projection is a bounding node is stated in category neutral terms, we may assume that the account for extraposition from the subject of a small clause is identical, in crucial respects, to extraposition from the subject of S.

We turn now to elements which, we assume, are neither directly assigned a θ-role in the syntax nor are they the subjects of predication. We have assumed, for example, that nominal complements, in the unmarked case, are not assigned a thematic role, their interpretation being a matter of rules of interpretation applying at LF or possibly some later level. As we have seen, nominal complements seem to be islands for extraction. Thus, NP movement may not apply in the following examples:

(48)a. $*[_{N''}$ the boat's $[_{i}\ _{N'}$ decision [on $\underline{t}_i$]]]

b. *[the men´s [conversation [with $\underline{t}_i$]]]
N" $_i$ N´

c. *[the issue´s [arguments [over $\underline{t}_i$]]]
N" $_i$ N´

d. *[the problem´s [thought [about $\underline{t}_i$]]]
N" $_i$ N´

We will also assume that the subject of NP does not receive a thematic role by predication since there is no maximal projection which may serve as its predicate; I am assuming, here, that the subject of a small clause receives its thematic role under predication since the projection of the predicate may be analyzed as X" which counts as a maximal projection. Notice that the subject of NP is an absolute barrier to extraposition from subject:[10]

(49)a. [[a friend of John]´s book]
N" N"

b. *[[a friend $\underline{t}_i$]s book [of John]$_i$]
N" N"

We, next, turn to a consideration of sentential adjuncts. Although adjuncts may predicate something of some subject or, possibly, act in the creation of complex predicates, they do not receive a thematic role from any element. Consider the examples in (50):

---

10 Lightfoot (1979) notes that extraposition from the subject of NP was not always ruled out in English; Middle English allowed the following sort of construction:

(i)a. kyng Priamus sone of Troy
´king Priamus of Troy´s son´

b. the Prolog of the Clerkes tale of Oxenford
´the prolog of the clerk of Oxford´s tale´

I have, as yet, no principled analysis of why the examples in (i) were grammatical.

(50)a. John only eats meat [rawer than rare]
b. Bill carried the table [without any help]
c. Ron twisted Daniel's arm [to get him to cry "uncle"]

The bracketed phrases in each of the examples is an adjunct. Notice that extraction from the bracketed phrases is ungrammatical:

(51)a. *what$_i$ does John only eat meat [rawer than $\underline{t}_i$]
b. *what$_i$ did Bill carry the table [without $\underline{t}_i$]
c. *what$_i$ did Ron twist Daniel's arm [to get him to cry $\underline{t}_i$]

Under the framework of assumptions developed in this chapter, it seems clear that elements that are not assigned a thematic role either by direct θ-role assignment, transmission of a θ-role by a preposition or assignment of a θ-role by predication are absolute barriers to extraction.

In this section, I have attempted to devise a taxonomy of extraction facts based on mechanisms of thematic role assignment. The facts seem to be arrayed in the following way: Elements that are directly assigned a thematic role, for example by a verb, are transparent to extraction. Elements that are indirectly assigned a thematic role, or assigned a thematic role by virtue of being the subject of a predication, may be extracted themselves, but extraction of a subpart of such a phrase yields marginal results. Finally, elements that are not assigned a thematic role are opaque to extraction.

3.3 Thematic domains

Let us start with the idea that movement inside a particular subtree is sensitive to thematic role assignment. In particular, we have grounds for assuming that movement out of a phrase is legitimate only under conditions of thematic role assignment. To this end, we define the following:

(52) Thematically Transparent Domain
A maximal projection, XP, is a thematically transparent domain if and only if
(i) the head of XP assigns a thematic role, or;
(ii) XP is directly assigned a thematic role.

It follows, from the definition in (52) that VP is always a thematically transparent domain, since the head of VP directly assigns a thematic role, and that the object of a verb is a thematically transparent domain, since it is directly assigned a thematic role by the verb. We will say that movement is restricted in the following way:

(53) Condition on Thematic Domains
An element, Y, may be extracted within a thematically transparent domain with respect to Y.

Consider, first, extraction from a direct object position:

(54) $[_{\bar{S}}$ who $_i$ $[_S$ John $[_{Infl'}$ Infl $[_{VP}$ see t_i]]]]

where Comp is taken as the head of $\bar{S}$ and Infl is taken as the head of S. The VP will be a thematically transparent domain with respect to extraction of the direct object because the head of VP, see, directly assigns a thematic role to its object. Hence, we may extract the object out of VP. In order to move the direct object to Comp, however, it must be the case that S is a

thematically transparent domain. Recall, however, that S is the domain for predication. We can say that VP compositionally assigns a thematic role to the subject of S; indeed, it appears that we must say this, given the selectional restrictions a verb phrase may place on the subject of S, as illustrated by the following:

(55)a. #the rock saw John
b. #the table demanded a glass of beer
c. #the value of the dollar eradicated the meaning of the sentence

(The '#' sign is used to indicate anomaly.) I said that the selectional restrictions follow from the form of the verb phrase given the relative lack of anomaly of the following:[11]

(56)a. This table demands to be refinished.
b. The value of the dollar ultimately eradicated family farming in the US (because of the trade imbalances it triggered).

Thematic role assignment to the subject of S crucially involves the form of the VP. We may assume, following a number of researchers, that Infl transmits the thematic role compositionally determined by the VP. If this is the case, then S is also a thematically transparent domain for extraction of the direct object. Hence, the direct object may be extracted at least as far as Comp.

Consider, now, successive cyclic movement, as in:

---

11 Marantz (1984) presents a number of arguments that thematic role assignment to subject is contingent on the form of the VP rather than simply the lexical properties of the verb, although the latter play a part in determining the form of the VP.

(57) [$_{\bar{S}}$ who$_i$ [$_S$ Mary [$_{Infl'}$ Infl [$_{VP}$ think [$_{\bar{S}}$ $\underline{t}_i$ [$_S$ John
[$_{Infl'}$ Infl [$_{VP}$ saw $\underline{t}_i$]]]]]]]]

We know, from the discussion of (54) that extraction of the direct object of the embedded clause may go at least as far as the Comp of the embedded clause. The embedded $\bar{S}$ is directly assigned a thematic role by the verb, <u>think</u>. By clause (ii) of definition (52), the embedded $\bar{S}$ is a thematically transparent domain and an element in Comp may be extracted out of that domain. The matrix VP, headed by <u>think</u>, is a thematically transparent domain because <u>think</u> assigns a thematic role. Therefore, extraction may take a phrase from the embedded Comp to a position external to the matrix VP. By reasoning parallel to that involved with (54), the matrix Infl transmits a thematic role to the subject; hence, the matrix S is a thematically transparent domain. But, given that the matrix S is a transparent domain, a phrase may be extracted from the embedded Comp to the matrix Comp. The assumptions about extraction defined in (52) and (53), therefore, allow successive cyclic movement just so long as the embedded $\bar{S}$ is assigned a thematic role. In general, this will be just in case the embedded $\bar{S}$ is a subcategorized complement.

Let us turn to extraction of a subpart of direct object as in (58):

(58) [$_{\bar{S}}$ of whom$_i$ [$_S$ Mary [$_{Infl'}$ Infl [$_{VP}$ take
[$_{NP}$ pictures $\underline{t}_i$]]]]]

The direct object, headed by pictures, is directly assigned a thematic role by the verb, take. By clause (ii) of definition (52), the direct object NP is a thematically transparent domain and a phrase may be extracted out of this domain. Here the reasoning is exactly the same as seen above: The VP is a thematically transparent domain because its head, take, assigns a thematic role; S is a thematically transparent domain because its head, Infl, transmits a thematic role to the subject. Hence, a subpart of a direct object may be moved to Comp. I will spare the reader a proof of the fact that a subpart of a direct object may be moved successive cyclically since it, again, parallels the reasoning behind the discussion of (57).

We turn now to consider the extraction of an object of a subcategorized preposition as in:

(59)a. Who$_i$ did John give the books to t_i

b. [$_{\bar{S}}$ who$_i$ [$_S$ John [$_{Infl'}$ Infl [$_{VP}$ give the books [$_{PP}$ to t_i]]]]]

The prepositional phrase, headed by to, is directly assigned a thematic role by the verb give. By clause (ii) of definition (52), the prepositional phrase is a thematically transparent domain. As a result, the object of the preposition may be extracted out of the prepositional phrase. From here on, the proof is essentially identical to those given above--each maximal projection dominating the extraction site is a thematically transparent domain, so movement to Comp is possible from the

object position of a subcategorized preposition. Since movement to Comp is possible from this position, we would expect that successive cyclic movement is also possible:

(60) $[_{\bar{S}}$ who$_i$ $[_S$ did Mary say $[_{\bar{S}}$ that $[_S$ John thought $[_{\bar{S}}$ that $[_S$ Bill gave the books $[_{PP}$ to $\underline{t}_i$]]]]]]]

As (60) attests, the facts fall as predicted.

In addition, we predict that extraction of a subpart of such a phrase is not good:

(61)a. ??who$_i$ did John give pencils to [friends of $\underline{t}_i$]

b. *of whom$_i$ did John give pencils to [friends $\underline{t}_i$]

c. ??who$_i$ did John talk about [relatives of $\underline{t}_i$]

d. *of whom$_i$ did John talk about [relatives $\underline{t}_i$]

We assume that the head noun in these constructions does not assign a thematic role to its complement, thus the NP does not satisfy clause (i) of definition (52). Furthermore, the preposition does not directly assign a thematic role to this NP; the thematic role is transmitted by the preposition from the verb which subcategorizes for this PP. Hence, the NP also fails clause (ii) of definition (52). It must be the case, then, that the NP is not a thematically transparent domain. If this is true, however, then no phrase may be extracted from this opaque domain, by virtue of the Condition on Thematic Domains. We can, therefore, account for the examples in (61). Note that we do not account for the differences between the examples; nothing in our

analysis explains why (61b) is worse than (61a), although a genuine contrast seems to exist. It is possible to derive the contrast from the recoverability of deletion. The empty categories in (61a) and (61c) have no governing element that will signal their presence; the head noun, itself, does not assign a thematic role to this position. In (61b) and (61d), on the other hand, the Case-marker <u>of</u> is stranded. This fact is, I suppose, sufficient to signal the presence of an empty category.

We turn, now, to extraction of a subject. Consider, first, the extraction of a subject of S, as in:

(62) $[_{\bar{S}}$ who$_i$ $[_S$ $\underline{t}_i$ saw John]]

The subject of S receives a thematic role within S, albeit indirectly. This thematic role is assigned via Infl, the head of S. Hence, S is a thematically transparent domain by clause (i) of (52). We may extract a subject out of S at least as far as the Comp of that S. Example (62) thus obeys our constraint that extraction be within a thematically transparent domain.

By reasoning similar to that found above, successive cyclic extraction of a subject is also possible:

(63) $[_{\bar{S}}$ who$_i$ $[_S$ does Mary think $[_{\bar{S}}$ $\underline{t}_i$ $[_S$ $\underline{t}_i$ ´ saw John]]]]

We have already seen, with respect to (62), that extraction is possible from the subject position of the embedded clause to the Comp position of the embedded $\bar{S}$. The proof that movement may precede from the embedded Comp to the superordinate Comp should be fairly familiar: The embedded $\bar{S}$ is directly assigned a

thematic role by the verb think; by clause (ii) of (52), this direct thematic role assignment renders the embedded $\overline{S}$ a thematically transparent domain. The matrix VP is, then, a thematically transparent domain with respect to extraction from the embedded $\overline{S}$. Since Infl transmits a thematic role, S is a thematically transparent domain. We conclude that extraction from the embedded Comp may go at least as far as the superordinate Comp. But, then, successive cyclic movement of an element originating in a subject position is also possible, as desired.

Extraction from the subject position of a small clause, as in (64) is parallel to extraction from the subject position of S:

(64) $[_{\overline{S}}$ who$_i$ $[_S$ John $[_{Infl'}$ Infl $[_{VP}$ consider $[_{NP}$ $\underline{t}_i$ a fool$]]]]]$

The small clause, [$\underline{t}$ a fool], is directly assigned a thematic role by the verb, consider; by virtue of this thematic role assignment, the small clause is a thematically transparent domain (clause (ii) of (52)) and VP is a thematically transparent domain (clause (i) of (52)). Since the matrix Infl transmits a thematic role, the matrix S is a thematically transparent domain. Hence, extraction may take proceed from the subject of a small clause to a superordinate Comp. The proof that successive cyclic movement of an element originating as the subject of a small clause is possible should be sufficiently familiar to allow me to forgo the details.

Of more interest is the status of extraction of a subpart of a subject with respect to the analysis developed so far. Consider, first, the extraction of a subpart of a subject of a clause:

(65) $*[_{\bar{S}}$ who$_i$ $[_S$ do $[_{NP}$ pictures of $\underline{t}_i$] annoy John]]

Recall that, by assumption, the head noun of the subject does not assign a thematic role to its complement. Hence, the subject NP fails to satisfy clause (i) of (52). Infl, furthermore, does not directly assign a thematic role to the subject NP, so the subject does not live up to clause (ii) of (52). Thus, the subject NP is not a thematically transparent domain; by the Condition on Thematic Domains then, it is not possible to extract an element out of the subject NP. The Condition on Thematic Domains and the definition given in (52) ensure that although a complete subject may be extracted, no subpart of the subject may be extracted.

The behavior of the subpart of a subject of a small clause with respect to extraction is, as we have seen, virtually identical to the behavior of a subpart of the subject of a clause:

(66) $*[_{\bar{S}}$ which man$_i$ $[_S$ does John consider $[_{AdjP}$ $[_{NP}$ relatives of $\underline{t}_i$] idiotic]]]

As with the subject of S, the head of the subject of the small clause will not assign a thematic role to its complement. Thus the subject of the small clause fails clause (i) of definition (52). Furthermore, the subject of the small clause receives a

thematic role by predication, a process that is, by assumption, different from direct thematic role assignment by a head to its complement. This means that the subject of a small clause also fails clause (ii) of definition (52) and must not be a thematically transparent domain. The Condition on Thematic Domains will rule out extraction out of the subject position of a small clause. Definition (52) and the Condition on Thematic Domains thus predict that the behavior of a subject of a small clause is parallel to that of the subject of S, as desired.

Notice that the account of extraction I have been developing in this chapter does not rely on elements of tree geometry. It focuses, rather, on thematic relations between elements; we may treat the precise details of tree structure as purely epiphenomenal, although thematic relations may be represented by means of a tree. The exact position of a phrase in a tree has no direct effect on extraction of a subpart of that phrase. In particular, the framework developed here predicts that a postposed subject should be as much an island for extraction as when the subject occurs in preverbal position.

Following Belletti & Rizzi (1981), let us assume that there is a rule of ne cliticization which adjoins the clitic, ne, to the verb in Italian. This rule accounts for sentences of the following form:

(67)a. Gianni è rimasto tre settimane a Milano.
'G. remained three weeks in Milan.'
b. *Gianni è rimasto tre a Milano.
'G. remained three in Milan.'

c. Gianni ne è rimasto tre a Milano.
'G. ne remained three in Milan.'

As Belletti & Rizzi observe, a preverbal (subject) NP may not be associated with ne:

(68)a. Tre settimane passano rapidamente.
'Three weeks elapse rapidly.'
b. Tre passano rapidamente
'Three elapse rapidly.'
c. *tre ne passano rapidamente
'three ne elapse rapidly'

Recall that Italian freely allows subjects to occur postverbally as in:

(69)a. Tre ragazze hanno parlato.
'Three girls spoke.'
b. Hanno parlato tre ragazze.
'Three girls spoke.'

Notice that it is not possible to attach ne to the verb from a post-verbal subject, as (70) illustrates:

(70) *ne hanno parlato tre
'three spoke.'

This follows because, no matter what the position of the subject NP, its head does not assign a thematic role nor is the NP directly assigned a thematic role. By (52), it will not be a thematically transparent domain and is, as a result, opaque to extraction.[12] Hence, ne may not be extracted from a subject, even if this subject occurs postverbally. An exception to this generalization are "unaccusative" verbs. These verbs occur with the auxiliary essere.

---

[12] There are some trivial counterexamples to this result which I will eliminate below.

(71) Ne sono passate tre.
ne are elapsed three.

Following Burzio (1980), let us assume that the surface subjects of these verbs originate inside VP, as elements that are directly assigned a thematic role by the verb, but not Case. In order to receive Case, the NP must either move to preverbal position (and receive Nominative Case from Infl) or it must be associated with an expletive in subject position (see Belletti & Rizzi (1981) for some discussion). In this latter case, we may assume that the post-verbal NP remains in situ and Case is transferred to this NP by the null expletive. If this is the case, then the post-verbal "subject" of these verbs should be a thematically transparent domain since they are directly assigned a thematic role by the verb. Hence, we should expect that ne cliticization is possible in this instance.

Let us turn, now, to adjuncts. It is safe to say that adjuncts, whatever their thematic relation is to the clause containing them, are not directly assigned a thematic role by any head. Suppose that some adjunct were directly assigned a thematic role; its presence would then be obligatory because it would be satisfying a lexical requirement of some head and the Projection Principle would demand that that lexical requirement be satisfied. Given this fact, we know automatically that an adjunct will never satisfy clause (ii) of definition (52). The question of the thematic transparency therefore reduces to whether the head of the adjunct assigns a thematic role and,

therefore, comes to satisfy clause (i) of (52). The only way an adjunct may come to be a thematically transparent domain is if its head assigns a thematic role.

Consider, first, an adjunct with sentential structure, as in:

(72) *[$_{\bar{S}}$ which how-to book$_i$ [$_S$ did you build these shelves [without [$_{\bar{S}}$ t_i [$_S$ PRO reading t_i carefully]]]]]

I will not bore the reader with a discussion of how the object may be moved to the Comp of the adjunct $\bar{S}$ since it would be identical with the discussion of example (54). Let us consider the movement from the adjunct Comp to the Comp of the matrix. Assuming that an element that moves into the head of Comp is the head of Comp (as argued in Chomsky, 1985), it is fairly safe to assume that the head of Comp does not assign a thematic role to the S. Hence, the $\bar{S}$ fails clause (i) of definition (52). Let's suppose that <u>without</u> does not assign a thematic role to the $\bar{S}$; if this is the case then the $\bar{S}$ also fails clause (ii) of definition (52) and fails to be a thematically transparent domain. The Constraint on Thematic Domains will rule out any extraction from the adjunct $\bar{S}$. I will (somewhat unfairly) hold off any discussion of the case where <u>without</u> assigns a thematic role until the next section.

Finally, consider the by now familiar case of a stranded preposition in an adjunct position, the standard judgement being that such preposition stranding is impossible:

(73) *[$_{\bar{S}}$ which class $_{i}$ [$_{S}$ does John fall often asleep
[$_{PP}$ during $\underline{t}_{i}$]]]

We are operating under the assumption that the head of the PP, during, does not directly assign a thematic role to its object. Hence, the prepositional phrase does not live up to clause (i) of definition (52). Since the PP is a true adjunct (by assumption), it will not receive a thematic role, much less be directly assigned a thematic role by some head. Thus, the PP also fails to satisfy clause (ii) of definition (52). Since the PP does not satisfy either of the clauses of definition (52), it may not count as a thematically transparent domain. So, by the Constraint on Thematic Domains, we may not extract anything out of the PP. The ungrammaticality of adjunct preposition stranding, then, follows from our framework.

So far, there seems to be no empirical difference between the framework developed here and the modification of subjacency proposed by Chomsky (1985). Chomsky, it will be recalled, has redefined the notion of bounding node to be sensitive to thematic role assignment. A node that does not receive a thematic role directly will count as a bounding node for subjacency. Thus, in example (73), the PP does not receive a thematic role and, therefore, counts as a bounding node for subjacency. Since S always counts as a bounding node in English, an element extracted from the object position of the preposition, during, would have to cross two bounding nodes: PP and S. "Move alpha", however,

is permitted to move an item across only one bounding node at a time. Hence, (73) violates subjacency. Similar demonstrations could be carried out for the other examples given in this section.

There is one data set where the framework adopted here and the modified version of subjacency diverge in their predictions. The facts of preposition stranding retroactive nominals, as pointed out earlier, does not follow from subjacency even if subjacency is made sensitive to thematic role assignment. This is simply because only one bounding node will be crossed. The crucial case, remember, involves extraction of a prepositional complement of a non-<u>ing</u> derived nominal, as in:

(74) *these ideas$_i$ don´t merit [$_{NP}$ PRO any arguments [$_{PP}$ about $\underline{t}_i$]]

In the subjacency version, the preposition complement to <u>argument</u> will, indeed, be a bounding node for subjacency because it is not assigned a thematic role. But the PRO need only cross this one maximal projection to land in the Spec of NP. A bona fide violation of subjacency would require the PRO to cross two bounding nodes before landing in its final destination.

Consider the derivation of (74) within the framework of assumptions developed in this chapter. The prepositional phrase is not assigned a thematic role by the head noun, <u>argument</u>, nor does the preposition assign a thematic role. Thus, the PP fails to meet either clause (i) or clause (ii) of definition (52) and

is not a thematically transparent domain. By the Constraint on Thematic Domains, then, no element may be extracted out of the PP. Hence, the movement from object of the preposition to Spec of NP is simply not allowed. As usual, an example like (74) must be contrasted with a well-formed example:

(75) these ideas$_i$ don't merit [$_{NP}$ PRO any arguing [$_{PP}$ about $\underline{t}_i$]].

As I argued above, the verb, merit, assigns the designated thematic role, θr, to the object NP. This renders the N boundary transparent, allowing the stem, argu(e), to assign a thematic role to its complement, the prepositional phrase headed by about. But, by clause (ii) of definition (52), this means that the prepositional phrase is a thematically transparent domain. The Constraint on Thematic Domains will allow us to extract the PRO from the object position of the preposition to the Spec of NP.

The assumptions discussed in this section seem to be making the right cut. Subjects, adjuncts and certain other positions, like complements to nouns and prepositions, are islands to extraction. Direct arguments, like subcategorized complements of a verb, may be freely extracted from. Notice that the effects of definition (52) and the Constraint on Thematic Domains closely parallel those of the modified version of subjacency. It is reasonable to ask whether we could simply replace the modified version of subjacency with the notion of thematically transparent domain and the Constraint on Thematic Domains; both frameworks,

after all, proscribe certain boundary conditions on syntactic movement and their effects seem almost identical.

The answer to this question is clearly negative. Once a maximal projection is declared thematically transparent, extraction may apply freely. A set of nested thematically transparent domains may, itself, be a thematically transparent domain and the assumptions developed above say nothing about extracting an element from the most deeply embedded domain out of the topmost domain in one step. In other words, the framework developed in this section does not place a maximal distance condition on extraction inside of transparent domains; but this is completely unlike subjacency. It is useful to consider a concrete example:

(76) $*[_{\bar{S}}$ which book$_i$ $[_S$ do you wonder $[_{\bar{S}}$ who$_j$ $[_S$ $\underline{t}_j$ read $\underline{t}_i]]]]$

Example (76) is a violation of subjacency since the wh-quantified NP, <u>which book</u>, must cross two S boundaries in one step to get to the matrix Comp. The example, however, does not violate the Constraint on Thematic Domains. This is simply because the wh-quantified NP is never extracted from a maximal projection that is not a thematically transparent domain, as the reader may verify.

One could imagine ways to modify (52) and the Constraint on Thematic Domains to capture the effects of subjacency. Some degree of caution, however, is in order. It is well-known, for

example, that subjacency must be parameterized to account for the fact that in some languages both S and $\overline{S}$ are bounding nodes while only $\overline{S}$ counts as a bounding node in other languages (see, for example, Rizzi (1982) or Sportiche (1981)). Given the assumptions of this chapter, such a variation must arise from parameters in thematic theory. The exact nature of such a parameter or set of parameters is unclear and would take much care to explore.

For present purposes, I will simply advocate a return to an older version of subjacency such as that discussed in Chomsky (1981):

(77) No rule may involve A, B in the structure (order irrelevant):

$$\ldots A \ldots [\underset{X}{\ldots} [\underset{Y}{\ldots} B \ldots] \ldots] \ldots$$

where X and Y count as bounding nodes for subjacency.

I will assume that languages select the bounding nodes for subjacency from the set {$\overline{S}$, S, NP}. Subjacency and the framework developed in this chapter are both elements of the module of grammar concerned with boundary conditions.

3.4 Some amendments

Extraposition structures have long been studied in generative grammar. Consider the examples in (78) and (79):

(78)a. That John will borrow the keys is likely.
b. That Bill stole the keys is certain.
c. That Mary took the book is obvious.

(79)a. It is likely that John will borrow the keys.
b. It is certain that Bill stole the keys.
c. It is obvious that Mary took the book.

It is quite well-known that the sentential subjects in (78) are islands to extraction:

(80)a. *what$_i$ [$_{\bar{S}}$ that John will borrow $\underline{t}_i$] is likely

b. *what$_i$ [$_{\bar{S}}$ that Bill stole $\underline{t}_i$] is certain

c. *what$_i$ [$_{\bar{S}}$ that Mary took $\underline{t}_i$] is obvious

If we assume that these sentential subjects originate in subject position at D Structure and receive their thematic roles via transmission of a composite thematic role through Infl, then the facts in (80) are exactly what we would expect. The sentential $\bar{S}$ will not qualify as a thematically transparent domain and are, therefore, islands to extraction. Recall that the surface position of a subject makes no difference as to its status as a thematically transparent domain (see above); prima facie, we would expect that the extraposed sentences in (79) are islands to extraction. This is not the case:

(81)a. what$_i$ is it likely [$_{\bar{S}}$ that John will borrow $\underline{t}_i$]

b. what$_i$ is it certain [$_{\bar{S}}$ that Bill stole $\underline{t}_i$]

c. what$_i$ is it obvious [$_{\bar{S}}$ that Mary took $\underline{t}_i$]

In the discussion, above, we concluded that a non-root $\bar{S}$ could have an empty Comp only if the Comp were properly governed (see also Stowell (1981)). The Comp, itself, is properly governed only if the $\bar{S}$ is transparent to government. Notice that, under this approach, we must claim that the extraposed $\bar{S}$ is transparent to government by the adjective in each of the above cases:

(82)a. *John will borrow the keys is likely.
b. *Bill stole the keys is certain.
c. *Mary took the book is obvious.

(83)a. It is likely John will borrow the keys.
b. It is certain Bill stole the keys.
c. It is obvious Mary took the book.

But transparency of a maximal projection to government by a head presupposes thematic role assignment to that maximal projection by the head. The facts in (82) and (83) indicates that the adjectives directly assigns a thematic role to the extraposed $\overline{S}$. But this conclusion conflicts with the assumption that the $\overline{S}$ in this structure originate in subject position at D Structure. Extraction out of an extraposed $\overline{S}$ in <u>it</u> extraposition structures contrasts sharply with extraction out of phrases in other kinds of extraposition structures:

(84)a. A man with a glass eye arrived.
b. A man arrived with a glass eye.
c. There arrived a man with a glass eye.

(85)a. *what kind of glass eye$_i$ did a man with $\underline{t}_i$ arrive
b. *what kind of glass eye$_i$ did a man arrive with $\underline{t}_i$
c. *what kind of glass eye$_i$ did there arrive a man with $\underline{t}_i$

The D Structure of (78a) is:

(86) [$_{\overline{S}}$ [$_S$ [$\underline{e}$] Infl is [$_{Adjp}$ likely [$_{\overline{S}}$ that [$_S$ John will borrow the keys]]]]

Since an $\overline{S}$ does not require Case, the $\overline{S}$ complement of the adjective may remain in situ (see Safir, 1985 on the assumption that $\overline{S}$ need not be assigned Case). In this case, since an empty

expletive is impossible in English, presumably because it would violate the ECP, the expletive element, it, must be inserted in subject position:

(87) $[_{\bar{S}}$ $[_{S}$ it Infl is $[_{AdjP}$ likely $[_{\bar{S}}$ that $[_{S}$ John will borrow the keys]]]]

Since the matrix subject does not receive a thematic role, nothing prevents the complement $\bar{S}$ from moving into the subject position:

(88) $[_{\bar{S}}$ $[_{S}$ $[_{\bar{S}}$ that $[_{S}$ John will borrow the keys] $_{i}$ Infl is $[_{AdjP}$ likely t_{i}]]]

Suppose that, as in (87), the sentential complement remains in situ. It receives a thematic role directly from the head of the adjective phrase and is, therefore, a thematically transparent domain; hence, we would expect extraction to be possible out of the $\bar{S}$. Suppose, on the other hand, that the sentential complement moves to subject position; then it will receive a thematic role by virtue of being linked with a trace. We can interpret this as meaning that it no longer receives a thematic role directly from the head of the phrase which contains it. If we take this relationship between an antecedent and a trace as an instance of indirect thematic role assignment, it follows that a sentential subject in the cases above will not count as a thematically transparent domain.

It should be obvious that phrases do not count as thematically transparent domains simply by virtue of being linked

to a trace which directly receives a thematic role. Consider, for instance, cases of passivization. In the standard analysis of passives (see Jaeggli (1984) and Keyser & Roeper (1984) for some recent expositions), the ability to assign an external thematic role to the subject is blocked and the Case feature of the verb is absorbed. These facts, along with the Extended Projection Principle, guarantee a D Structure of the following form for a passive sentence:

(89) $[_{\bar{S}}\ [_{S}\ [_{NP}\ \underline{e}]\ \text{be}\ [_{VP}\ \text{broken}\ [_{NP}\ \text{the vase}]]]]$

The passive participle, broken, will assign a thematic role directly to its object. Given the passive morphology, it is unable to assign Case to its object. In order to avoid the effects of the Case Filter, the object must move to a position which receives Case. In order to satisfy the uniqueness condition on thematic roles that the θ-criterion places on arguments, the position that the object moves to must be a $\bar{\theta}$-position. Since assignment of an external thematic role to subject is impossible in passives, the object may move to subject position to receive nominative Case from Infl:

(90) $[_{\bar{S}}\ [_{S}\ [_{NP}\ \text{the vase}]_{i}\ \text{be}\ [_{VP}\ \text{broken}\ \underline{t}_{i}\ ]]]$

If we interpret the relationship between the surface subject and the trace in object position as involving thematic role transmission and not direct thematic role assignment, then the surface subject will fail to be a thematically transparent

domain, as desired since extraction is impossible:

(91)a. *who$_i$ were [$_{NP}$ stories about $\underline{t}_i$] told $\underline{t}_j$

b. *who$_i$ were [$_{NP}$ pictures of $\underline{t}_i$] taken $\underline{t}_j$

c. *what$_i$ were [$_{NP}$ cartons of $\underline{t}_i$] loaded $\underline{t}_j$ into the truck

Passive sentences raise an interesting point. Suppose we have a D Structure of the following form:

(92) [$_{\bar{S}}$ [$_S$ [$_{NP}$ $\underline{e}$] were [$_{VP}$ told [$_{NP}$ stories about who]]]]

The D Structure object is directly assigned a thematic role and, by definition, counts as a thematically transparent domain. It should be possible, then, to extract the wh-element inside the NP:

(93) [$_{\bar{S}}$ who$_i$ [$_S$ [$_{NP}$ $\underline{e}$] were [$_{VP}$ told [$_{NP}$ stories about $\underline{t}_i$]]]]

From here, the derivation proceeds like that of a normal passive sentence. The object is moved to subject position in order to receive Case:

(94) [$_{\bar{S}}$ who$_i$ [$_S$ [$_{NP}$ stories about $\underline{t}_i$] were [$_{VP}$ told $\underline{t}_j$]]]

It would appear that we have a well-formed derivation of the ungrammatical examples in (91).

There is, however, no problem with ruling out the representation in (94), if the Constraint on Thematic Domains is taken as a constraint on representations. The logical object is the surface subject; the surface subject does not directly receive a thematic role in S since its thematic role is

transmitted by virtue of the subject's binding a trace in the object position. The surface subject of (94) may be interpreted as a thematically opaque domain with respect to S since Infl, the head of S, does not directly assign a thematic role. This problem may be reproduced by using not only passives, but also ergatives and middles as in (95):

(95)a. Bureaucrats$_i$ bribe $\underline{t}_i$ easily.

b. The frigate$_i$ sank $\underline{t}_i$.

In the examples in (95), an NP is directly assigned a thematic role by the verb and then moves to the subject position for reasons of Case assignment. If the Constraint on Thematic Domains were taken as a constraint on rules, it should be possible to violate the Subject Condition (subsumed the Constraint on Thematic Domains):

(96)a. *which department$_i$ do [bureaucrats from $\underline{t}_i$] bribe easily

b. *what country$_i$ did [a frigate from $\underline{t}_i$] sink yesterday.

The above example demonstrates that the Constraint on Thematic Domains cannot be taken as a constraint on rules but, rather, must be interpreted as a constraint on representations since the functioning of the rules never violated a core principle. This fact gives us a further argument from distinguishing the Constraint on Thematic Domains from subjacency. Subjacency is commonly viewed as a constraint on

rules (see, for example, Lasnik & Saito (1984)). Given the conceptual difference between constraints on rules and constraints on representations, the derivation of the S Structure in (94) provides us with a good argument for keeping the Constraint on Thematic Domains distinct from subjacency.

The problem brought up by middles, ergatives and passives should bring to light an additional problem with the system developed so far, in particular with the definition of "thematically transparent domain". We observed, above, that $\bar{S}$ is a thematically transparent domain by virtue of the fact that its head, Infl, transmits a thematic role to the subject of $\bar{S}$. Passives, ergatives and middles share the property that the subject of $\bar{S}$ is a $\bar{\theta}$-position; in other words, Infl in these sorts of structures does not, in fact, transmit a thematic role to the subject. If this is so, then we would expect on the fact of it that $\bar{S}$ should be an absolute barrier to extraction in such structures. This prediction is obviously false:

(97)a. $[_{\bar{S}}$ who $_i$ $[_{\bar{S}}$ $\underline{t}_i$ bribes $\underline{t}_i$ easily]]

b. $[_{\bar{S}}$ what $_i$ $[_{\bar{S}}$ $\underline{t}_i$ sank $\underline{t}_i$]]

c. $[_{\bar{S}}$ which vase $_i$ $[_{\bar{S}}$ $\underline{t}_i$ was broken $\underline{t}_i$]]

The above also holds for so-called extraposition structures since, under the analysis presented in this chapter, the subject position is a $\bar{\theta}$-position; nevertheless, as we saw above, it is possible to extract out of the subordinate clause:

(98)a. what $_i$ is it likely $[_{\bar{S}}$ that John will borrow $\underline{t}_i$]

b. what_i is it certain [$_{\bar{S}}$ that Bill stole $\underline{t}_i$]

c. what_i is it obvious [$_{\bar{S}}$ that Mary took $\underline{t}_i$]

Thus, the transparency of S with respect to extraction does not hinge on whether or not the head of S transmits a thematic role to the subject. It would be more accurate to say that S is always transparent to extraction from within an S to the Comp that is a sister of the S.

Having established the fact that the thematic properties of Infl are irrelevant to the transparency of S with respect to extraction, we are faced with the question of why S is consistently transparent (barring the effects of subjacency). We can answer this by considering the relationship between Comp and S a little more closely. It has been observed that the form of Comp affects the realization of Infl (see, for example Stowell (1981)). For example, a Comp headed by <u>for</u> obligatorily requires an infinitival S:

(99)a. It would be a mistake [$_{\bar{S}}$ for [$_S$ John <u>to</u> refuse the offer]]

b. *it would be a mistake [$_{\bar{S}}$ for [$_S$ John refuses the offer]]

Similarly, a Comp headed by <u>that</u> requires a tensed S:

(100)a. It is strange [$_{\bar{S}}$ that [$_S$ John would refuse the offer]]

b. *it is strange [$_{\bar{S}}$ that [$_S$ John to refuse the offer]]

c. *it is dangerous [$_{\bar{S}}$ that [$_{S}$ PRO to refuse the offer]]

We may assume that the form of Comp ultimately determines the form of S. The relationship between Comp and Infl appears to be a subcase of standard subcategorization relationships and is quite comparable to the relationship between a verb and its object. In both cases, the presence of the subcategorized element is obligatory, given the lexical properties of the head. As we saw in examples (99) and (100), the lexical properties of the complementizer are inviolable: The lexical properties of <u>that</u> are not satisfied by an infinitival nor are the lexical properties of <u>for</u> satisfied by a tensed clause.

We will assume, then, that the relationship between Comp and Infl is one of strict subcategorization and that the lexical properties of Comp must be satisfied in the syntax. The analogy between Comp/Infl and verb/object suggests that S counts as a thematically transparent because of clause (ii) of (52) which stipulates that an XP is a thematically transparent domain if and only if it receives a thematic role. Intuitively, however, the relationship between Comp and S is not one of thematic role assignment since S is not in any clear sense an argument of S in the way that an object, for example, is an argument of a verb. A more general notion than thematic role assignment is needed, one in which thematic role assignment counts as a special case. Given that S satisfies the lexical requirements of Comp in the syntax, we can restate clause (ii) as:

(52)(ii)′ XP satisfies the lexical properties of a head in the syntax.

From the modified definition of "thematically transparent domain", we can deduce that S is always transparent because it must satisfy the lexical restrictions placed upon it by Comp.

We can now begin to rethink clause (i) of definition (52) which allows an XP to be a thematically transparent domain if the head of XP assigns a thematic role. As it stands, this clause only works to ensure that VP is always a thematically transparent domain. But the head of VP must satisfy lexical requirements placed upon it by Infl. Thus, a tensed Infl subcategorizes for a VP headed by a verb that is morphologically marked as tensed:

(101)a. John {sees/saw} the accident.
b. *John {seeing/seen/see} the accident

Similarly, if Infl contains a modal, the VP must be headed by a verb in the stem form. In general, we have the following dependencies between auxiliary verbs and their complements (the notation is adapted from Baker (1981)):

(102)a. be: [progressive] < __ V <ing> >

b. be: [passive] < __ V <en> >

c. have: [perfective] < __ V <en> >

d. modal: < __ V <stem> >

Thus, assuming an analysis under which Infl strictly subcategorizes for features of the XP which follows it, VP will

be a thematically transparent domain since it, too, satisfies the lexical features of a head that subcategorizes for it.

We are, therefore, in a position to restate the definition of "thematically transparent domain" as follows:

(103) Thematically Transparent Domain
A maximal projection, XP, is a thematically transparent domain if and only if XP satisfies the lexical properties of a head in the syntax.

This simplification of the definition of thematically transparent domain will allow us to simplify certain assumptions made earlier in this chapter. Specifically, I assumed that prepositions never directly assign a thematic role to their complements; at best, a preposition could transmit a thematic role from some other element, say a verb, to the object of the preposition. This assumption worked to guarantee that a prepositional complement to a noun would be a thematically transparent domain, all other things being equal. Thus, in an ill-formed sentence like:

(104) *John$_i$ could use [$_{NP}$ PRO$_i$ some conversation [$_{PP}$ with t_i]]

The prepositional phrase headed by with would not count as a thematically transparent domain because with does not assign a thematic role. It followed that NP-movement from the object of the preposition to the Spec of NP would be ill-formed. It is now irrelevant whether or not the preposition assigns a thematic role; the PP in (104) does not satisfy lexical features in the syntax and, therefore, does not count as a thematically transparent. We can, therefore, rule out movement to the Spec of

NP without consulting the thematic properties of the head of the opaque projection; it is sufficient to verify whether or not the projection satisfies lexical requirements in the syntax.

It should be noted that we rule in examples like:

(105) John$_i$ could use [$_{NP}$ PRO$_i$ some conversing [$_{PP}$ with $\underline{t}_i$]]

The matrix verbal, could use, assigns the thematic role, θr, to the object NP. This, in turn, licenses thematic role assignment by the ing nominal to the preposition phrase headed by with. Since this thematic role assignment transpires in the syntax, the PP counts as a thematically transparent domain. Hence, NP-movement from the object of the preposition to the Spec of NP is well-formed.

Finally, let us turn to adjuncts headed by without, as in:

(106) *[$_{\bar{S}}$ which how-to book$_i$ [$_S$ did you build these shelves [without [$_{\bar{S}}$ $\underline{t}_i$ [$_S$ reading $\underline{t}_i$ ´ carefully]]]]

Earlier, I showed that extraction from the adjunct headed by without was impossible given that without does not assign a thematic role to the $\bar{S}$ that follows it. Now, the thematic properties of without are irrelevant to determining its status as a thematically transparent domain; the adjunct does not satisfy lexical requirements in the syntax and, hence, is not a thematically transparent domain. Under any assumption about the status of without with respect to thematic role assignment, the

projection of without will be opaque to extraction.

3.5 Tree Geometry

In this chapter, we have attempted to account for a number of boundary phenomena associated with the extraction of sub-constituents. In particular, we have attempted to account for why a sub-part of certain types of constituents (in particular, a subject, an adjunct or certain nominal complements) may not be extracted from that constituent. In doing so, we have appealed to thematic relations as they are expressed at S Structure; a constituent that is directly assigned a thematic role (e.g., the object of a verb) is transparent to extraction whereas constituents that are indirectly assigned a thematic role (subjects, objects of certain prepositions, nominal complements or elements which receive a thematic role under mediation by an empty category) or not assigned a thematic role at all (e.g., an adjunct) are boundaries to extraction. This account does not make reference to purely configurational relations such as "is properly governed by" or "is in a government configuration with". Briefly put, our account seeks to abstract away from specifics of tree geometry. It will therefore be useful to contrast the account of boundaries developed here with two other accounts of extraction phenomena that are more directly based on structural notions such as government--namely, the "connectedness" framework of Kayne (1984) and the Constraint on Extraction Domains (CED) of Huang (1982).

We will turn first to the "connectedness" framework (Kayne,

1984), which represents a serious attempt to draw on the relationship between an empty category and its antecedent in order to delimit the possible distribution of empty categories. In doing this, the framework relies very heavily on certain purely structural considerations; roughly put, an empty category must stand in a particular configuration with a governor and this structural configuration must be preserved throughout the tree, until the antecedent for the empty category is located. It will be useful to introduce the technical mechanisms by which this is accomplished. He first defines the notion of g-projection:

(107) Y is a g-projection of X iff
  a. Y is a projection of X (in the usual sense of X´-theory) or of a g-projection of X
     or
  b. X is a structural governor and Y immediately dominates W and Z, where Z is a maximal projection of a g-projection of X, and W and Z are in a canonical government configuration.

The definition of canonical government configuration is as follows:

(108) W and Z (Z a maximal projection and W and Z immediately dominated by some Y) are in a canonical government configuration iff
  a. V governs NP to its right in the grammar of the language in question and W precedes Z
     or
  b. V governs NP to its left in the grammar of the language in question and Z precedes W.

Notice that English falls under clause (a) of the definition in (108), since the object occurs to the right of the verb, so the canonical government configuration for English is $[_{Y}\ W\ Z]$. The connectedness ECP is defined over g-projection sets (G_B) of a

category $\underline{B}$, where $\underline{g}$ governs $\underline{B}$:

(109)a. For all $\underline{p}$, if $\underline{p}$ is a g-projection of $\underline{g}$ the $\underline{p}$ is an element of G_B

b. $\underline{B}$ is an element of G_B

b.´ if $\underline{d}$ dominates $\underline{B}$ and $\underline{d}$ does not dominate $\underline{g}$ then $\underline{d}$ is an element of G_B

Lastly, we come to the statement of the connectedness ECP itself:

(110) Let $\underline{B}_1 \ldots \underline{B}_j, \underline{B}_{j+1} \ldots \underline{B}_n$ be a maximal set of empty categories in a tree T such that there is a c-commanding $\underline{a}$, such that for all j, $\underline{B}_j$ is locally bound by $\underline{a}$. Then

a. the union of all the g-projection sets of the $\underline{B}_j$ must constitute a subtree of T and

b. there must exist a $\underline{p}$ such that $\underline{p}$ is an element of the set of all the g-projection sets of the $\underline{B}_j$ and $\underline{p}$ dominates $\underline{a}$.

In other words, one must be able to construct a g-projection set for some empty category in such a way that its antecedent is contained in this set and the set of nodes contained in this set make up a continuous subtree of the structure in which the empty category occurs.

Let us consider how the "connectedness" framework works for some simple examples (irrelevant details omitted):

(111)a. [$_{\bar{S}}$ who$_i$ [$_S$ did John [$_{Infl'}$ Infl [$_{VP}$ take [$_{NP}$ pictures [$_{PP}$ of $\underline{t}_i$]]]]]]

b. *[$_{\bar{S}}$ who$_i$ [$_S$ were [$_{NP}$ pictures of $\underline{t}_i$] [$_{Infl'}$ Infl [$_{VP}$ taken]]]]

In determining the g-projection set, one begins from the empty category and works upward, so, in (111a), we will begin with the trace left by wh-movement of who to Comp. First, the empty category stands in a government configuration with of so by the above definitions, the entire PP is an element of the g-projection set of the trace. The head noun of the object NP and the PP stand in a canonical government configuration so, by the above definitions, we are entitled to add the entire NP to the g-projection set of the trace. Next, the verb stands in a canonical government configuration with the object NP, so we are now allowed to add the VP to the g-projection set of the trace. Since Infl stands in a canonical government configuration with the VP, we can add the projections of Infl (including S) to the g-projection set of the trace. Finally, Comp stands in a canonical government configuration with S so we can add the $\bar{S}$ to the g-projection set of the empty category. Thus, excluding non-maximal projections, the g-projection set of the empty category in the first example is:

(112) $\{\underset{j}{t}, PP, NP, VP, S, \bar{S}\}$

which constitutes a coherent subtree of the entire S. Since the antecedent of the empty category, who, is in Comp (we will identify it as the head of Comp) and since Comp is immediately dominated by $\bar{S}$, the entire representation passes the connectedness ECP.

The second example is a violation of the Subject condition.

The g-projection of the empty category in (111b) begins with the trace itself which is governed by the preposition, of; thus the PP is part of the g-projection set of the empty category. The PP stands in a canonical government configuration with the head of the subject NP; hence, the NP is part of the g-projection set of the empty category. Now, the subject NP does not stand in a canonical government configuration with Infl since the NP occurs to the left of Infl and, as observed above, canonical government precedes from left to right in English. Hence, S is not part of the g-projection set of the empty category. The entire g-projection set of the empty category (once again, omitting non-maximal projections) is:

(113) $\{t_i, PP, NP\}$

The antecedent of the empty category, who, is not dominated by any member of the g-projection set in (113). Example (111b), therefore, fails the connectedness ECP and is ruled out as ill-formed.

It should be fairly easy to see the fate of empty categories which are separated from their antecedents by a left-branch; in all such cases, the g-projection set of the empty category will stop at the left branch and the antecedent will not be dominated by any member of the g-projection set of the empty category. The subtree consisting of the g-projection set of the empty category and the node dominating the antecedent of the empty category will not constitute a coherent, continuous subtree of the larger

structure. The connectedness framework thus relies on the fact that subject NPs in English and other well-studied languages occur in a non-canonical government configuration. The framework, therefore, makes a number of clear predictions about extraction asymmetries. First, in a language where the subject of a predication structure may occur in a canonical government configuration with the predicate, there should be no Subject Condition effects. Second, if an adjunct occurs in a canonical government configuration then it should not form an island to extraction.

As we have seen throughout this chapter, the last prediction is clearly incorrect. Adjuncts form absolute barriers to extraction no matter where they occur in the tree. In order to account for this, Longobardi (1985) proposes that the following be added to the definition of g-projection:

(114) W governs Z.

Under the assumption that adjuncts are never governed, the g-projection set of an empty category inside an adjunct will never be able to include nodes that are not dominated (or equal to) the maximal projection of the adjunct; hence, a trace inside an adjunct will never connect with an antecedent outside the adjunct (except in the special case of parasitic gaps; see Kayne (1984) and Longobardi (1985) for discussion).

Kayne's framework, along with Longobardi's revisions, make a clear prediction: If we can find a case where a phrase is not

directly θ-marked but is both (properly) governed and in a canonical government configuration, then it should be possible to extract a sub-part of that phrase. We can find this sort of situation in Null-Subject Languages like Italian and Spanish (see Rizzi, 1982 and Jaeggli, 1982). Put briefly, these languages typically allow subjects to appear post-verbally:

(115)a. Juan compró una casa ayer.
J. bought a house yesterday
'Juan bought a house yesterday.'

b. Compró una casa ayer Juan.
bought a house yesterday J.
'Juan bought a house yesterday.'

As argued in Rizzi (1982), this ability to postpose the subject exempts the subject from familiar ECP effects like the Complementizer-Trace Filter:

(116) Quién$_i$ dijiste que llegó ayer $\underline{t}_i$
who you-said that arrived yesterday
'Who did you say arrived yesterday?'

Since English has no mechanism corresponding to post-verbal subjects that would allow subjects to circumvent the effects of the ECP, a word-for-word translation of (116) is, of course, ungrammatical:

(117) *who$_i$ did you say that $\underline{t}_i$ arrived yesterday

The explanation for (116) runs as follows: A post-verbal subject is properly governed, presumably by the verb; therefore, the variable left by extracting the post-verbal subject is properly governed and obeys the ECP. Interested readers should consult Rizzi (1982) and Jaeggli (1982) for an extensive discussion of

the ECP and post-verbal subjects.

For our purposes, it is sufficient to note first that post-verbal subjects are (properly) governed and second that a post-verbal subject stands in a canonical government configuration with VP.[13] That is, a post-verbal subject stands to the right of VP and the canonical government configuration in these languages is $[_{Y}\ W\ Z]$. The predictions made by the connectedness framework are quite clear. When a subject occurs preverbally, it should be an island to extraction because it is not in a canonical government configuration (i.e., it is a left-branch and canonical government in Spanish goes from left to right). When a subject occurs post-verbally, however, it should be easy to extract a sub-part of the subject because it is both governed and in a canonical government configuration (i.e., on a right-branch).

Consider, now, the following pair of sentences:

(118)a. *de qué pintor$_i$ dijiste que varios dibujos $\underline{t}_i$
of which painter you-said that several drawings
son caros
are expensive

---

[13] I am assuming that post-verbal subjects are adjoined to VP, although the argument still goes through under the assumption that the subject somehow winds up inside the VP or, provided we can guarantee proper government, is adjoined to Infl. What is important is simply that the subject is in a canonical government configuration with some other element.

b. *de qué pintor$_i$ dijiste que son caros

of which painter you-said that are expensive

varios dibujos $\underline{t}_i$

several drawings

´*by which painter did you say that several drawings are expensive´

In (118a), a sub-part of a preverbal subject has been extracted and the result, as predicted by the connectedness framework, is ungrammatical. In (118b), the subject is post-verbal and hence, as we decided above, both governed and in a canonical government configuration. Nevertheless, the resulting structure is not acceptable. This is not expected within the connectedness framework since the subject is on a right-branch.[14] The connectedness framework cannot, therefore, place (118a) in the

---

14 I am indebted to Osvaldo Jaeggli for the above example. He observes that extraction from within an NP is a rather complicated matter in Spanish and appears to be subject to a number of rather mysterious conditions. There does seem to be some variation in the acceptability of extracting from subjects and slight effect depending on whether the subject is preverbal or post-verbal. Depending on the nature of the embedded predicate in examples like (118) the acceptability may range from "*" for extraction to a preverbal subject and "???" for extraction from a post-verbal subject. In short, the contrasts do not seem to be nearly as dramatic as one would expect given the connectedness framework. In all fairness, the framework developed here does not predict any contrast at all. It should be pointed out that argument structure (namely, the properties of the predicate) does seem to determine the nature of the contrast where there is a contrast. For example, extraction from a post-verbal subject isn´t particularly bad when the construction is a middle (with se morphology). This is as we would expect given the underlying contention of this chapter—namely that argument structure determines the possibility of extracting from a particular phrase. As usual, further work is required before we can completely clear the matter up.

same class as (118b) but, rather, must attribute the ungrammaticality of (118b) to some principle other than the ECP.

The above result with respect to the examples in (118) is exactly what we would expect given our definition of Thematically Transparent Domain, the Constraint on Thematic Domains and the fact that the subject does not directly satisfy the lexical requirements of any head, no matter whether it occurs preverbally or post-verbally. Thus, our account of extraction is, to a considerable degree, independent of particular facts about tree configurations. No matter where the subject hangs in the tree, it does not count as a thematically transparent domain and, hence, is opaque to extraction.

The examples in (118) also pose a problem for a framework like that of Huang (1982). Huang accounts for the impossibility of extracting from a subject or an adjunct by means of the Constraint on Extraction Domains (CED):

(119) Constraint on Extraction Domains
A phrase A may be extracted out of domain B only if B is properly governed.

Given that subjects and adjuncts are, by hypothesis, not properly governed, it follows from the CED that a sub-constituent may not be extracted out of a subject or an adjunct. But we saw, above, that post-verbal subjects must be properly governed in order to account for the fact that they are exempt from Complementizer-Trace effects. Hence, by the CED, it should be possible to extract a sub-part of a post-verbal subject simply by virtue of

the fact that these subjects are properly governed. Like the connectedness framework, the CED must account for the lack of contrast between the examples in (118) by making appeal to some other principle of grammar in the case of post-verbal subjects; thus they would use the CED to rule out extraction from a preverbal subject and this other principle, whatever it may be, to rule out extraction from a post-verbal subject.

Finally, neither the connectedness framework nor the CED has much to say about the alternation found in retroactive nominals (repeated here):

(120)a. *John$_i$ could use [$_{NP}$ PRO$_i$ some conversation [$_{PP}$ with $\underline{t}_i$]]

b. John$_i$ could use [$_{NP}$ PRO$_i$ some conversing [$_{PP}$ with $\underline{t}_i$]]

In both examples in (120), the gap is governed and is a sub-part of a right-branch. Thus, all things being equal, the connectedness framework does not predict a contrast between the two examples. Presumably, the contrast itself would not be due to the ECP. Furthermore, the PP which contains the empty category is itself properly governed by the head noun since an entire PP may be extracted from an NP in English. Thus, the CED also fails to differentiate between the two examples.

The framework developed in this chapter, however, attributes the contrast in (120) to the same mechanism which accounts for Left-Branch effects, Subject Condition effects and adjunct-island

effects. This is accomplished simply because the Constraint on Thematic Domains makes no reference to relations between nodes in a tree, but, rather, it makes reference to thematic relations that holds between various constituents. This fact, in itself, brings up an interesting question. The current research on tree geometric relations like government and connectedness is quite reasonable given that the essential properties of tree are determined by principles of $\overline{X}$-theory. Following the set of assumptions developed in Stowell (1981), the null hypothesis is that tree structure is the hierarchical representation of the argument structure of individual lexical items; that is, the mapping between argument structure and hierarchical structure is a function. All things being equal, we would suppose that one could replace all reference that the theory makes to argument structure with statements about hierarchical structure and the results would be largely unaffected. In general, this expectation seems to be borne out by the facts; there is a great deal of overlap between the framework in this chapter, the connectedness framework and the CED. In the case of Null Subject Languages, the subject of S may be realized in two ways--preverbally or post-verbally. Thus, the mapping between argument structure and tree structure is not a function since the input (the subject of S) can be mapped onto one of two outputs (preverbal or post-verbal). In fact, for languages which have syntactic movement, D Structure is the only level where we would expect a "clean" mapping between argument structure and tree

structure since tree relations may change in the course of a derivation, subject to the constraints of Universal Grammar. Thus, the relationship between argument structure and tree structure may change in the course of a derivation. It is, therefore, a matter of some interest whether constraints on movement are to be stated solely in terms of relations between nodes in a tree or whether such constraints must also make reference to properties of argument structure. I take the main result of this chapter to be that the latter course is the correct one to take.

3.6 Summary

We began this chapter by considering the status of a very restricted class of nominals with respect to the ECP. In chapter 2, we saw that these nominals allow NP-movement of a PRO from a complement position into Spec of NP. Furthermore, this movement is quite restricted; in particular, the PRO can be moved out of a complement to the head noun only if the noun is an ing nominal and the governing verbal is of a restricted semantic class. We hypothesized that, in the unmarked case, nouns do not assign a thematic role to their complements. The governing verbal involved in this class of constructions has the property of licensing thematic role assignment by the ing nominal. Thematic role assignment by the head noun, in turn, makes the complement "transparent" to extraction, thus allowing NP-movement of a governed PRO to the Spec position of NP.

We explored whether or not thematic role assignment by a lexical element is a precondition to that item's counting as a proper governor or whether being assigned a thematic role affected an item's standing as a bounding node for subjacency. We found that neither approach could fully account for the range of data examined--Subject Condition effects, the opacity of adjuncts and certain NPs with respect to extraction, and so forth. We therefore factored these thematic effects into an independent boundary condition (the Condition on Thematic Domains) by first defining "thematically transparent domain" and then allowing extraction out of an XP to be possible only if that XP counts has thematically transparent. We found that this approach generalized across a variety of constructions, providing an analysis which unified the data concerning retroactive nominals with Subject Condition effects, and so on. Unlike subjacency, which is a condition on rules, this condition is stated as a condition on representations which applies, minimally, to S Structure representations.

Chapter 4
A Guide to Binding and Non-Overt Operators

As the name of this chapter implies, it is intended to be a short course on certain fundamental aspects of the Binding Theory and the treatment of non-overt operators. The approach to binding and non-overt operators developed in this chapter will provide the foundation to control that we will develop in chapter 5. Hence, anyone familiar with the binding theory as outlined in Chomsky (1981) need not spend much time on section 4.1 where this approach to binding is outlined. We differ from Chomsky (1981) only in the assumption that a pronoun may act as a bound variable just in case it is locally A-bound by a variable. In the next section, we will examine the framework of Generalized Binding as developed in, for example, Aoun (1985). Again, if the reader is already familiar with Generalized Binding, then this section may be skipped. Finally, in section 4.3, we will examine the

treatment of non-overt operators found in Aoun & Clark (1985), where non-overt operators are analyzed as anaphors. Any reader who is not familiar with the results of Aoun & Clark (1985) should pay particularly close attention to this section.

4.1 The Classical Binding Theory

We will begin by considering the Binding theory developed in Chomsky (1981). As an independent subcomponent of the grammar, the Binding theory is intended to constrain the possible coreference relations between NP positions in syntactic representations.[1] As we will see, the Binding theory also constrains the distribution of certain types of NPs.

4.1.1 Anaphors

Consider, first, sentences like those in example (1):

(1)a. John saw himself in the mirror.
b. John talked to himself.
c. John considers himself to be the best lawyer in the Bar Association.

In each of the above sentences, the reflexive, himself, is understood as being coreferential with the name, John; there is, in fact, no other way to assign a referent to the reflexive in

---

1 It should be kept in mind that, strictly speaking, the term "coreference relations" conflates two different conditions. One condition is that a particular type of NP be referentially dependent (i.e., picks up its reference) on another NP in some local domain. The other condition is that a particular type of NP be referentially independent (i.e., disjoint) from every NP in some domain. Conceptually, these two constraints are independent and there is no a priori reason why these two conditions should be elements of the same subcomponent of the grammar. It is, therefore, an empirical question whether or not the Binding theory should encompass both of these constraints.

the above sentences. It seems to be the case that a reflexive element may not refer to anything on its own; it must, rather, rely on an antecedent, some other NP which the reflexive can rely on to pick out a referent. We will assume that the inability to refer directly is the defining property of the set of anaphors. If this hypothesis is correct, then we have a natural account of the following examples:

(2)a. *it rained on himself
b. *it seemed to himself that it shouldn't be raining
c. *it struck himself that it was difficult to walk

In each of the examples of (2), the only available antecedent for the reflexive is the pleonastic element, <u>it</u>. If we assume that the pleonastic element does not have the ability to refer--in the sense required by the Binding theory--then we have a natural account of the ungrammaticality of the examples in (2). The reflexive in each of the examples in (2) must have an antecedent in order to refer, but none of the reflexives have an available referential element from which it can pick up a referent. Thus, each of the reflexives in (2) is non-referential. Notice that each of the anaphors in (2) is in a position that is associated with a slot in the argument structure of a lexical head. We now make the following assumption:

(3) In order for an NP to be associated with the argument structure of a head, that NP must refer.

Since, by the above argumentation, none of the reflexives in (2) may refer, it follows from (3) that none of them may satisfy the argument structure of any head. Hence, the reflexive elements in

(2) will never be associated with a θ-role and each of the examples violates the θ-criterion.

The requirement that anaphors have an antecedent is not in itself sufficient to characterize the distribution of anaphors. To illustrate this point, consider the examples in (4):

(4)a. *friends of John saw himself
b. *pictures of John disturb himself
c. *some stories about John were told to himself

In the examples in (4), John is a priori a potential antecedent for the anaphor; nevertheless, the sentences are not well-formed. We will assume that the antecedent of an element must c-command that element. In order to account for the facts in (4), we define binding as follows:

(5) Bind
X binds Y if and only if:
(i) X c-commands Y; and
(ii) X and Y have the same index.

In order for an NP to count as an antecedent for some referentially dependent element, that NP must bind the element in the sense given in (5).

We have so far placed two restrictions on anaphors. First, the anaphor must have an antecedent and, second, the antecedent must bind the anaphor. Although these constraints alone place strong restrictions on the distribution of anaphors, they are inadequate to account for the complete distribution of anaphoric elements; consider the examples in (6):

(6)a. *John believes Mary to have seen himself
b. *John claims that Georgia painted himself
c. *John dreamed that mirrors couldn´t reflect himself

In contrast with the examples in (2), the anaphor in each of the examples in (6) has a possible antecedent, John.[2] In principle, it is possible that the anaphor could pick up reference from this potential antecedent. Nevertheless, the examples in (6) are ill-formed. In order to account for this fact, we must assume that it is insufficient for the anaphor to have an antecedent; in order for a representation containing an anaphor to be well-formed, the anaphor must have an antecedent contained in some local domain. We will turn, now, to a characterization of this local domain.

In the examples in (6), the potential antecedent and the anaphor are separated by an S boundary:

(7)a. *John believes [Mary to have seen himself]
S
b. *John claims that [Georgia painted himself]
S
c. *John dreamed that [mirrors couldn´t reflect himself]
S

It is insufficient to require that the anaphor and its antecedent be clause-mates as the examples in (8) show:

(8)a. John considers [himself to be the brightest]
S

---

2 Notice that Georgia, Mary and mirrors cannot be taken as antecedents for the reflexive. Presumably, this is because himself is [+masculine] while the other NPs are not. This implies first that anaphors have ϕ features (i.e., features for gender, number and person) and that the ϕ features of an anaphor must agree with those of its antecedent.

b. John never believed that [a book about himself
S
would make the best-seller list]

The anaphors in the examples in (8) are parallel to the examples in (7) in that they are separated from their antecedents by an S boundary; nevertheless, the examples in (8) are grammatical, unlike those in (7). To further illustrate the point that the local domain that is relevant to binding is not coextensive with S, consider the following contrasts:

(9)a. John saw pictures of himself in the Museum of Contemporary Art.
b. *John saw Mary's pictures of himself in the Museum of Contemporary Art

(10)a. John resents rumors about himself.
b. *John resents the students' rumors about himself

The local domain for the anaphors in (9a) and (10a) must be the matrix S. The local domain for the anaphors in (9b) and (10b), however, must be the object NP.

Contrasts like those in (9) and (10) demonstrate that the proper definition of local domain must be stated in a category neutral way. The contrast in (10) shows, furthermore, that whether or not a category is a governing category for some element must be computed on the basis of structural relations. Thus, an NP is not a governing category simply by virtue of being an NP; its status as a governing category must be determined on the basis of its internal structure.

Notice that the object NP in (10b) differs from that in (10a) in that the NP in (10b) has a genitive NP. Following Stowell (1983), we will assume that <u>subject</u> is defined as the NP

that is an immediate constituent of XP where X ranges over the grammatical categories. Based on contrasts like that in (10), Chomsky (1981) proposes the following definition of governing category:

(11) Governing Category
B is the governing category for A if and only if B is the minimal category containing A, a governor of A, and a SUBJECT accessible to A.

We will put aside the definition of "accessible SUBJECT" for the moment and assume that SUBJECT is coextensive with the definition of subject given above. Let us consider the contrast in (10). The S Structure for (10a) is roughly:

(12) [John [resents [rumors [about himself]]]
S VP NP PP

The minimal category containing himself and a governor of himself is the PP complement of the noun, rumors. The PP, however, does not contain a SUBJECT and, so, it may not function as the governing category for the anaphor. The next minimal category containing the anaphor and its governor is the object NP, but, once again, the object NP does not contain a SUBJECT for the anaphor and cannot be the anaphor's governing category. The next minimal maximal category up is the VP which also does not contain a SUBJECT for the anaphor. Finally, we arrive at the matrix S. The S contains the anaphor, its governor and a SUBJECT for the anaphor, John; hence, the root S is the governing category for the anaphor.

Now, consider the structure for (10b) which is roughly:

(13) *[John [resents [[the students´] rumors [
S VP NP NP PP
about himself]]]

The minimal category containing the anaphor and its governor is the PP, but, as in example (12), the PP does not contain a SUBJECT for the anaphor. The object NP, however, does contain a SUBJECT--the genitive NP, the students´. Hence the object NP is the governing category for the anaphor.

Structurally, example (12) differs from example (13) in that the governing category for the anaphor in (12) is the root S while the governing category for the anaphor in (13) is the object NP. In both example (12) and example (13), however, the only possible antecedent for the anaphor is the subject of S. Hence, the anaphor in example (12) is bound in its governing category and the anaphor in example (13) is not bound (is free) in its governing category. We are now in a position to state condition A of the Binding theory:

(14) An anaphor must be bound in its governing category.

where bound is taken in the sense of (5) rather than the looser "has an antecedent in its governing category".

Let us, now, consider the contrast in (15):

(15)a. John believes [himself to be the brightest]
i S i
b. *John believes [himself Infl is the brightest]
i S i

We will assume that the governor for the anaphor in (15a) is the matrix verb, believe. Hence, the minimal category containing the anaphor, its governor and a SUBJECT is the root S. Since John is

the antecedent for the anaphor, the anaphor is bound in its governing category, as required by condition A of the Binding theory.

The situation is somewhat different in (15b). In (15b), the anaphor is governed within the embedded S by INFL, which assigns nominative Case to that position.[3] The minimal category containing the anaphor and its governor is, therefore, the embedded clause. Since the anaphor is in the subject position of the embedded clause, it is not obvious what element is acting as its SUBJECT.

The major difference between (15a) and (15b) is that (15a) is tensed whereas (15b) is nonfinite. We will assume that a tensed Infl carries agreement features--the features for person, number and, optionally gender. These agreement features are identical to the ϕ features associated with NPs in general. We can define SUBJECT, as opposed to subject, as follows:

(16) The SUBJECT of a category is the most prominent element in that category which bears ϕ features.

Since the Infl in (15b) is tensed, and therefore bears ϕ features, it will act as the SUBJECT for the anaphor in subject position of the embedded clause. It follows that the embedded clause in (15b) is the governing category for the anaphor. Since the anaphor is not bound in its governing category, condition A

---

[3] This would seem to contradict the directionality of government assumed in chapter 2. Following Bennis (1981) and Stowell (1983), I will assume that nominative Case assignment occurs from Comp and that the tense element of Infl moves to Comp.

of the Binding theory correctly rules that (15b) is ungrammatical.

There remains the task of explicating the notion of "accessible SUBJECT" given in the definition of governing category. To do this, consider the contrast in (17):

(17)a. *John$_i$ believes [$_S$ himself$_i$ Infl is a success]

b. John believes [$_S$ [$_{NP}$ stories about himself$_i$] Infl are a success]

Example (17a) is identical to example (15b) so we need not go into its derivation. Example (17b) is similar to (17a) and (15b) since the embedded clause is tensed; in (17b), however, the anaphor is a subpart of the subject NP. Despite the tensed Infl, the anaphor is well-formed under the interpretation where John is taken as the antecedent of the anaphor.

Let us assume that, by convention, two elements that agree are coindexed. Since AGR agrees with the subject NP, Infl and the subject are coindexed. We will define "accessibility" in terms of the following well-formedness condition:

(18) The i-within-i Constraint
*[$_A$ ...B...], where A and B bear the same index

(19) Accessible
A is accessible to B if and only if B is in the c-command domain of A and assignment to B of the index of A would not violate the i-within-i constraint.

The question now reduces to whether the tensed Infl of the embedded clause in (17b) is an accessible SUBJECT for the anaphor inside the subject position. Coindexing the anaphor with Infl

would give the following:

(20) ...[$_S$ [$_{NP}$ stories about himself$_i$]$_i$ Infl$_i$ ...

But the indexing in (20) violates (18); hence, the Infl of an embedded sentence cannot act as accessible SUBJECT for an anaphor properly contained with the subject NP. This means that the governing category for the anaphor in (17b) is the root S, not the embedded S. Since the anaphor is bound in its governing category, example (17b) does not violate condition A of the Binding theory.

Consider, finally, sentences like the following:

(21)a. *[$_{\bar{S}}$ who$_i$ [$_S$ were [$_{N''}$ stories himself$_i$] told by $\underline{t}_i$]]

b. [$_{\bar{S}}$ who$_i$ [$_S$ $\underline{t}_i$ told stories about himself$_i$]

In example (21a), the anaphor, <u>himself</u>, is bound by the wh-element, <u>who</u>. The example, nevertheless, is ungrammatical. Example (21a) should be contrasted with example (21b) where a wh-element is ultimately the antecedent for the anaphor. Notice that the anaphor in (21a) is bound from Comp while the anaphor in (21b) is bound by the variable in the subject position of the clause. The Comp position is never assigned a thematic role by any head and, as a result bears no grammatical function with respect to any head. Comp falls in the class of non-argument ($\bar{A}$) positions. The subject position may serve to satisfy the argument structure of the verb; it therefore falls into the class of argument (A) positions. The contrast in (21) follows naturally if we assume that condition A of the binding theory

((14) above) is interpreted as:

(22) An anaphor must be A-bound in its governing category.

Throughout the remainder of this section, we will interpret "bound" and "free" in the binding conditions as meaning "A-bound" and "A-free".

The binding theory would be of only limited interest if it only held for overt NPs. We will assume that empty categories also must obey the binding conditions. All things being equal, we will assume that for each condition of the binding theory, there is a corresponding empty category which is subject to that condition.

Consider the contrast in (23):

(23)a. $\underset{i}{\text{John}}$ seems [$\underset{S\ i}{\underline{t}}$ to have left]

b. *$\underset{i}{\text{John}}$ seems [$\underset{S\ i}{\underline{t}}$ Infl has left]

Both examples in (23) involve NP-movement, that is movement from a position that receives a θ-role (a θ-position) to a position which does not receive a θ-role (a $\overline{\theta}$-position). In fact, (23a) differs from (23b) only in the fact that the former is infinitival while the latter is tensed. Assuming that this is the only difference in structure, how are we to account for the difference in grammaticality between the two examples? In (23a), the empty category has no accessible SUBJECT within the embedded S; its governing category must be the matrix clause since the matrix clause contains the only SUBJECT accessible to the empty category. Since the matrix subject binds the empty category, we

may conclude that the empty category is A-bound in its governing category and must, therefore, be an anaphor. In (23b), on the other hand, the empty category has a tensed Infl as its accessible SUBJECT. The governing category for this empty category will be the embedded clause. Since its antecedent, John, is outside of the embedded clause, the empty category is A-free in its governing category. We have evidence, then, that the trace left by NP-movement is subject to condition A of the binding theory and is, by definition, an anaphor.

4.1.2 Pronominals

We saw, in the previous section, that there is a class of elements, anaphors, which are obligatorily referentially dependent. These elements must, furthermore, be bound in a local domain; an anaphor which has an antecedent outside of its governing category is not well-formed. In this section we will delineate some of the properties of pronominals.

Consider the coreference possibilities in the following examples:

(24) John$_i$ believes [$_S$ Bill$_j$ saw him$_{\{i,*j\}}$]

(25) [$_{N''}$ John$_i$'s relatives] annoy him$_i$

Although the matrix subject, John, in (24) may be taken as coreferent with the pronoun, him, the embedded subject must be disjoint from the pronoun. The governing category for the pronoun is the embedded S since the embedded S is the minimal category that contains pronoun, a governor of the pronoun (the

verb, saw) and a SUBJECT accessible to the pronoun. In example (25), however, the pronoun may be interpreted as coreferent with the genitive NP, John. Notice, however, that the genitive NP, although contained in the same governing category as the pronoun, does not c-command the pronoun; by definition, it may not bind the pronoun. It appears, then, that a pronoun must be disjoint in reference from every c-commanding NP in its governing category; condition B of the binding theory is formulated as follows:

(26) A pronominal must be free in its governing category.

As with anaphors, "free" is interpreted as "A-free".

Condition B, along with definitions of SUBJECT and governing category, predicts minimal pairs like the following:

(27)a. John_i believes $[_S$ he_i Infl is the brightest]

b. *John_i believes $[_S$ him_i to be the brightest]

The governing category of the pronoun, he, in (27a) is the embedded S, since the embedded S contains the pronoun and its governor, the tensed Infl, which also acts as the accessible SUBJECT for the pronoun. Hence, the pronoun and the matrix subject, John, may corefer freely. The embedded S in (27b), however, is infinitival and, hence, lacks a tensed Infl to act as the accessible SUBJECT for the pronoun, him. The governor for the pronoun is, furthermore, the matrix verb, believes. Since the minimal category containing the pronoun, its governor, and a SUBJECT accessible to the pronoun is the matrix S in (27b),

condition B correctly predicts that John and the pronoun must be disjoint.

Despite the similarity of formulation, conditions A and B of the binding theory place very different conditions on the elements they range over. Where condition A defines a minimal domain in which an anaphor must find its antecedent, condition B defines a minimal domain in which a pronominal must not find its antecedent. This difference can be accounted for, at least in part, by considering the referential properties of anaphors and pronominals. Let us assume, with Chomsky (1982), that elements refer only insofar as they are able to range over elements of domain D, where domain D consists of mental representations of individuals. Under our assumptions, anaphors have access to domain D only indirectly; that is, they may access domain D only via an antecedent.

Pronominals, by contrast, may access domain D either by means of an antecedent or independently. Evans (1980) distinguishes four different uses of pronouns. We can illustrate Evans's typology of pronouns with the following sets of sentences:[4]

(28)a. He left the room.
b. It was unclear to him how to leave the room.

(29)a. John claims that he left the room early.

[4] This section owes a great deal to the work found in Montalbetti (1984).

b. John's friends think that he must have left early.

(30)a. Every student thinks he is the brightest.
b. No one claims he left the room early.

(31)a. Very few senators admire Kennedy and they are very junior.
b. Harry owns some sheep and Bill vaccinates them in the spring.

The examples in (28) show that a pronoun may access domain D directly without relying on an antecedent somewhere in the S. Evans cites examples like those in (29) as coreferent pronouns; that is, the pronoun corefers with another NP in the sentence. Note that the pronoun in (29b) may be taken as coreferential with John although the latter does not c-command the former. It is, therefore, clear that Evans's use of "coreference" is not coextensive with the binding theory's "bound"; the binding theory is not a complete theory of coreference.[5] The examples in (30) illustrate "bound" in Evans's sense. To be explicit, consider the LF representations associated with the examples in (30):

(32)a. [every student $_{i}$ [$_{S}$ t $_{i}$ thinks [$_{S}$ he $_{i}$ is the brightest]]]

---

5 No one, to my knowledge, would ever claim that the binding theory is a complete theory of reference. The question is what status the binding theory has with respect to the theory of reference. One could claim that the binding theory is a subpart of the theory of reference. One could plausibly argue that the binding theory is a theory of indexation that is independent of the theory of reference. Finally, one could just as plausibly abandon the binding theory for another theory of coreference relations (or a set of interactive subtheories of grammar). For the moment, I will leave these questions open.

b. [no one [$\underline{t}$ claims [he left the room
S_i S_i S_i
early]]]

The examples in (32) show that the local binder for the pronoun must be taken as the variable left be the scope assignment rule, Quantifier Raising (QR), not the quantifier itself. Let us assume, following Jaeggli (1984b) and Montalbetti (1984) that a pronoun may act as a variable only if the pronoun is bound either by a variable left by application of QR to the quantified NP or another element in an A-position that is so bound; a pronoun may not be locally bound by the quantified NP itself at LF.[6] The reference of the pronoun will vary with the range of quantifier in the sense that instantiating an element of domain D for the variable will trigger a like instantiation of the pronoun.

Prima facie, the distinction between "coreferent" and "bound" may appear to be unclear. Evans makes it obvious that these two terms are not to be confused; coreference does not imply binding. The distinction is far from trivial as the ambiguity of the following indicates:

(33) Most college freshmen think they're smart.

---

6 This proposal is similar to the analysis of bound pronouns found in Higginbotham (1980). See also Brody (1984); for a somewhat different approach see Safir (1984b). We will define locally bound as:

(i) $\underline{A}$ locally binds $\underline{B}$ if and only if $\underline{A}$ binds $\underline{B}$ and there is no $\underline{C}$ such that $\underline{C}$ binds $\underline{B}$ and $\underline{C}$ does not bind $\underline{A}$.

Intuitively, the local binder of an element is the first binder up in the phrase marker.

On one reading, the pronoun, they, acts like a bound pronoun. That is, given the set of college freshmen, most members of that set will have the property of thinking that he is smart without necessarily having beliefs about the set of college freshmen as a whole. The coreferent reading, on the other hand, implies a belief about the set of college freshmen; that is, most college freshmen think that college freshmen are smart. We may contrast the ambiguity of (33) with (34):

(34) No one thinks they're smart.

Example (34) requires that no individual has the property of thinking that he is smart; no one picks out the empty set. Since the quantified NP in (34) has scope, the pronoun in (34) may be interpreted as bound. The set that the quantified NP picks out, however, is empty; thus we would not expect a coreferent reading for (34). My intuition is that this is correct.

While the distinction between bound and coreferent readings of pronouns is undoubtedly correct, it is not at all clear how best to codify this distinction in the classic binding theory of Chomsky (1981). The binding theory provides us with a single notational device for expressing coreference relations--indexation. If bound and coreferent imply identity of indices on NPs, then it is difficult to see how indexation serves to disambiguate between the two readings. Using identity of indexation to represent both the bound and coreferent readings of a sentence would seem to force us to conclude that the level of Logical Form does not completely disambiguate scope relations

since both readings will be identically represented at that level.[7]

We assumed, above, that in order for a pronoun to be interpreted as a bound variable, it must be A-bound by a true variable. We can take this as implying that, in determining the truth value of a proposition involving quantification. every time a fregean substitution is performed on the variable an identical substitution is performed on the bound pronoun. That is, the pronoun will always "pick out" the same element of domain D as the variable picks out. The possible range of reference of the variable (and, by virtue of binding, the pronoun) is ultimately restricted by the operator which binds the variable.

Loosely speaking, let us say that the range of the quantification defines a (possibly empty) set of entities. We can suppose that the pronoun may pick out the same set that the quantification defines. Assuming that this is the correct approach to interpreting the meaning of <u>coreferent</u>, then it is not the case that the pronoun and the variable bound by a quantifier pick out the same sorts of entities; the pronoun, rather, picks out the set of entities defined by the operator/variable pair. In other words, under a coreferent

---

7 The system of <u>linking</u> developed in Higginbotham (1983) and extended in Montalbetti (1984) provides an unambiguous means of representing the bound and coreferent readings of pronouns. See the cited references for discussion. For present purposes, I will continue to explore a version binding theory which represents coreference relations by means of indexation.

interpretation of a pronoun, the variable picks out elements of some set while the pronoun picks out the set itself. Under this interpretation of coreferent, there is no sense in which the variable binds the pronoun when the pronoun corefers with the operator/variable pair. It would appear, then, that the coreferent interpretation of pronouns involves contraindexation of the pronoun and the variable.

If the pronoun and the variable are contraindexed, however, then the coreferent representation of a pronoun is identical to the representation of the independent use of the pronoun. Prima facie, we are faced with the same problem as above; namely, applications of the rule of scope assignment (which, following May (1977), we will assume is Quantifier Raising (QR)) and rules of interpretation (RI) have failed to disambiguate LF representations. This may not be as bad as one might suppose. In particular, let us assume that indexation is the syntactic reflex of the semantics of reference. That is the assignment of indices occurs in the syntax and is, therefore, subject to purely syntactic constraints. Indices, however, do not fully interpret NPs; let us suppose that NPs are interpreted by some relation, call it h, which takes as its argument the index associated with an NP and returns some element of domain D. If two NPs have identical indices, then h will, of course, assign the same element of D to each NP:

(35) For all indices i, j: if $i=j$ then $h(i)=h(j)$

Consider, however, the case where *i* does not equal *j*. We must, of course, allow the possibility that the value of the assignment relation will differ at these two indices. But, crucially, the value of the assignment relation need not differ for distinct indices. Recall, for example, the famous case of *the morning star* and *the evening star* as in:

(36) The morning star is the same as the evening star.

As we will see below, the binding theory requires that *the morning star* and *the evening star* be contraindexed in (36). The example is true, however, only if the two NPs are assigned the same individual by the assignment relation. In order to provide a coherent account of the truth conditions of (36), we must allow the case where the assignment relation returns the same element of D for distinct indices. But this is exactly the mechanism required to allow the coreferent interpretation of pronouns. Thus, the assignment relation will allow operator/variable pairs to select a range of elements of D and a coreferent, but contraindexed, pronoun to select a set that consists of the elements selected by the operator/variable pair.

As always, some residual puzzles remain to be solved. Consider, for example, the following indexation:

(37) Bill$_i$ saw him$_j$.

Although *i* does not equal *j* in (37), our reasoning above admits the possibility that h(i) equals h(j); that is, the assignment relation may give *Bill* and *him* the same value. This is obviously

incorrect; Bill and him may not be interpreted as coreferent. We will assume with Montalbetti (UCLA colloquium; 1984) that this is not a problem for the binding theory but, rather, a problem either for the mechanics of the assignment relation or for the general theory of pragmatics.

Let us now turn to the fourth type of pronoun identified by Evans (1980). This is the so-called "e-type" pronoun exemplified by the examples in (31), repeated here:

(38)a. Very few senators admire Kennedy and they are very junior.
b. Harry owns some sheep and Bill vaccinates them in the spring.

It is a common mistake to identify the underlined pronouns in (38) as bound by the quantified NPs very few senators and some sheep in (38a) and (38b), respectively. That this is an error can be seen by changing the examples to ones where accidental coreference is impossible:

(39)a. *no senator admires Kennedy and he is very junior.
b. *Harry owns not a single sheep and Bill vaccinates it in the spring.

The examples in (39) are ungrammatical under the interpretation that takes the underlined NPs as being in a binding relation. The examples in (39) should be contrasted with those in (40):

(40)a. No student thinks he is the brightest.
b. No one claims he left the room early.

where the underlined NPs may be interpreted as being in a binding relation. Recall that, in the LF representations of the examples in (40), the pronoun will be locally bound by the trace left by application of QR to the quantified NP. By our assumption that a

pronoun may act as a bound variable only if it is locally bound by a variable, the pronouns in (40) may be interpreted as bound variables.

In the examples in (38) and (39), on the other hand, the variable left by applying QR to the quantified NPs in the left-hand conjunct will not bind the pronouns in the right-hand conjunct. The bound interpretation of these pronoun is, therefore, impossible. In general, however, it should be possible to interpret the pronouns and the quantified NPs as coreferential. This is the reading found in the examples in (38). By substituting the <u>no</u> for the quantifier, as in (39), we rule the possibility of the coreferential reading; hence the examples in (39) are anomolous. Our treatment of pronominals would thus class the e-type interpretation with the coreferent interpretation.

The assumption that a pronoun is a bound variable only if it is bound by a variable places an interesting typology on certain widely studied constructions. In particular, the "donkey sentences" will not be treated as involving bound pronouns:[8]

(41) every man $[_{\bar{S}}$ who$_i$ $[_S$ $\underline{t}_i$ owns [a donkey]$_j$]] beats it$_j$

---

8 Recent studies of the donkey sentences include Heim (1982) and Haik (1984). Both of these works treat the phenomenon as a form of binding, albeit binding of a rather unusual nature. See the cited works for details and chapter 5 for a discussion of Heim's analysis. May (1985) is also of great interest in this regard.

Under the assumption that the indefinite NP, a donkey, undergoes QR, the trace left by this movement will not bind the pronoun, it, since the trace and the pronoun do not stand in the c-command relation. Since the relationship between the quantified NP and the pronoun in (41) will not be mediated by a variable, the pronoun will be unable to act as a bound variable at LF. Hence, the proper account of the interpretation of sentences like (41) must follow from some component other than the binding theory as we have so far outlined it.

A similar story may be told for the Bach-Peters sentences:[9]

(42) [every pilot who shot at it$_j$]$_i$ hit [some MIG that chased him$_i$]$_j$

where the pronoun, it, is interpreted as "bound" by the NP, some MIG that chased him, and the pronoun, him, is interpreted as "bound" by the NP, every pilot who shot at it. If we assume that QR adjoins both of the quantified NPs to S in (42), the resulting structure will be:

(43) [$_S$ [some MIG that chased him$_i$]$_j$ [$_S$ [every pilot who shot at it$_j$]$_i$ [$_S$ t$_i$ hit t$_j$]]]

Neither of the pronouns, it or him, is bound by a variable; thus, neither pronoun will be treated as a bound variable. As in the

---

9 See Higginbotham & May (1981) for an account of the Bach Peters sentences. Again, the phenomenon is treated as binding, but only after the application at LF of an absorption rule which builds n-ary quantifiers.

case of the donkey sentences, we are forced to treat the interpretation of the Bach-Peters sentences as following from some component other than the binding theory.

The assumption that pronouns act as variables only when bound by a variable provides for a straightforward account of weak crossover structures such as:

(44) *[his$_i$ mother] loves [every man$_i$]

The genitive pronoun, his, in (44) neither c-commands nor is c-commanded by the position occupied by every man at S Structure. Application of QR to every man will give the structure:

(45) $[_S$ [every man]$_i$ $[_S$ [his$_i$ mother] loves t_i]]

As can be seen by inspection of (45), the pronoun, his, is not bound by a variable. Since the pronoun may not be directly bound by the quantified NP, the bound interpretation of the pronoun is impossible.

Notice that we predict that a coreferent interpretation should be possible in weak crossover structures given an appropriate quantified NP/pronoun pair. Thus, the following should at least marginally admit a coreferent interpretation:

(46) [their mothers] love all the men

where underlining indicates the coreference relation. My intuition is that this is correct, although we are far from providing a full account of coreference.

The assumption that pronouns may only be bound by an operator if that binding is mediated by an A-position places an

interesting typology on the possible relationships between an operator and a pronoun. Certain cases of apparent binding, like the donkey sentences and the Bach-Peters sentences must be taken as falling in the domain of some component other than the binding theory. Some cases of pronominal binding remain to be solved. Consider, for example, the non-overt element, PRO. We have assumed throughout that PRO is a pronominal anaphor; that is that PRO has the features both of a pronoun and of an anaphor[10] and is, as a result, subject to both conditions A and B of the binding theory. Since conditions A and B place mutually exclusive requirements on the elements they hold of, an element subject to both these conditions may not have a governing category and, hence, may not be governed (the PRO theorem).

The assumption that PRO has pronominal features predicts, under the framework adopted here, that PRO may not be directly bound by a quantifier. As Higginbotham (1980) observes, this prediction is not correct:

(47) [PRO$_i$ seeing his$_i$ father] pleased [every boy]$_i$

Application of QR to (47) will result in the following structure:

(48) [[every boy]$_i$ [[PRO$_i$ seeing his$_i$ father] pleased
S S
t$_i$]]

The trace left by application of QR to every boy in (48) does not

---

10 For some criticisms of this position, see Bouchard (1984), Lebeaux (1984) and Sportiche (1983).

c-command and, hence, cannot bind the PRO. Nevertheless, the PRO (and the genitive pronoun his) behave as if bound by the quantified NP. We can further test that this is binding and not coreference by using a negatively quantified NP in place of every boy:

(49) [PRO$_i$ seeing his$_i$ father] pleased [none of the boys]$_i$

Example (49), under the reading indicated by the given indexation, is not anomolous. Such sentences must, therefore, involve genuine binding of the PRO by the quantified NP. Higginbotham (1980) argues that there is a 'PRO gate': PRO acts as a gate which may exempt certain structures from weak crossover in that it exceptionally allows direct binding by an operator.[11] If PRO has pronominal characteristics, then the PRO gate facts are apparent counterexamples to the principle that pronominals may not be directly bound by an operator. For the moment, we will treat the PRO gate simply as a counterexample. We will, however, have occasion to return to these facts below.

4.1.3 R-Expressions

We turn now to the behavior with respect to the binding theory of elements that have inherent reference (unlike anaphors) and which, furthermore, may not be referentially dependent on any other element (unlike pronominals). This class of elements

---

[11] Safir (1984b) and Brody (1984) also discuss the PRO gate in some detail.

includes that are neither anaphoric nor pronominal. Consider, for example, the behavior of proper nouns:

(50)a. she$_i$ saw Mary$_{\{j, *i\}}$

b. he$_i$ thought that [$_S$ Bill$_{\{j, *i\}}$ left early]

c. he$_i$ claimed that [$_S$ she$_i$ said that [$_S$ Bill$_{\{j, *i\}}$ left early]]

d. after she$_i$ left, Mary$_i$ saw Bill

Example (50a) shows that an R-expression, in this case Mary, may not be locally A-bound. Examples (50b) and (50c) show that this prohibition against the A-binding of an R-expression is not merely local, but extends across governing categories. Finally, (50d) shows that precedence is not necessarily relevant to determining coreference between an R-expression and a pronominal, since the pronoun in (50d) precedes the R-expression it is coreferential with. Notice, however, that the pronoun does not c-command the R-expression; since c-command is a necessary condition for binding, the pronoun does not A-bind the R-expression and the example is grammatical. We can, therefore, formulate condition C of the binding theory as:

(51) R-expressions must be free.

As always, free and bound are to be read as A-free and A-bound for purposes of the binding theory.

Compare the examples in (50) with those in (52):

(52)a. *who$_i$ did he$_i$ say $\underline{t}_i$ left

b. *the man who$_i$ he$_i$ saw $\underline{t}_i$ ...

c. *who$_i$ did Mary say John$_i$ saw $\underline{t}_i$

In each of the examples in (52) the wh-element may not be coreferent with some element in an A-position; for example, <u>who</u> may not be construed as binding <u>he</u> in (52a). Notice that in each of the examples in (52), the overt element c-commands the variable left by wh-movement; thus, the variables in (52) are locally A-bound. If the trace of wh-movement has the properties of an R-expression (and is, as a result, subject to condition C), then the binding facts in (52) follow trivially from the binding theory (see Chomsky, 1981 and the references cited there for further discussion). Each of the variables in (52) is locally A-bound which violates the condition that R-expressions be A-free.

The examples in (52) should be contrasted with those in (53):

(53)a. John$_i$ is easy to please $\underline{t}_i$

b. this book$_i$ was bought to put $\underline{t}_i$ on the shelf

c. a shovel$_i$ is to dig with $\underline{t}_i$

Following Chomsky (1977), let us assume that the examples involve wh-movement of an abstract element, <u>Op</u>; the examples in (53) will then have the structure:

(54)a. John$_i$ is easy [Op$_i$ [PRO to please $\underline{t}_i$]]

b. this book$_i$ was bought [Op$_i$ [PRO to put $\underline{t}_i$ on the shelf]]

c. a shovel$_i$ is [Op$_i$ [PRO to dig with $\underline{t}_i$]]

The variables in each of the above examples appears to be bound by an A-position, the matrix subject in each of the above examples. Assuming this to be correct, the examples in (54) stand as a priori counterexamples to the assumption that variables are subject to condition C.

If we compare the examples in (54) with the ungrammatical examples in (53), a clear structural difference between the two sets of structures: The variables in (53) are locally bound by an element in an A-position while the variables in (54) are locally bound by an element in an $\bar{A}$-position. Thus, the relative structural position of the operator defines a domain in which the variable associated with that operator must be A-free. Following Chomsky (1981) we will assume that the following condition obtains:

(55) A variable must be A-free in the domain of its operator (i.e., its binder).

For our purposes, it will be irrelevant whether (55) is taken as an independent principle or as a way of relativizing condition C for variables.

4.1.4 Summary

We have so far been concerned with a fairly simple binding theory, repeated here as (56):

(56) Binding Theory
A. Anaphors must be bound in their governing category.
B. Pronominals must be free in their governing category.
C. R-expressions must be free (in all governing categories).

This version of the binding theory serves to constrain the relation of identity of indexing for elements occupying A-positions. With the possible exception of the condition in (55) and the constraint that operators may not locally $\bar{A}$-bind pronominals (see the discussion above), non-argument positions do not enter into this version of the binding theory.

The binding theory given in (56) is intended to constrain all NPs, whether overt or empty. Thus, condition A holds for anaphors and reflexives, both types of NPs being phonologically overt, and NP-traces, which are phonologically non-overt. Condition B constrains the distribution of pronouns (overt) and pro (non-overt).[12] Condition C holds for proper and common NPs (overt) and for variables (non-overt). Finally, the distribution of the pronominal anaphor, PRO, is regulated by conditions A and B.

Despite the wide-applicability of the binding theory in (56), it is still restricted to filtering relations that obtain between A-positions. It has little or nothing to say about coindexing relationships between an A-position and an $\bar{A}$-position or between two $\bar{A}$-positions. It is, therefore, or some interest to discover whether a binding theory similar to the one given in (56) could be extended across the two different types of

---

[12] We have not, as yet, had occasion to discuss the properties of the non-overt pronominal pro. We will defer such discussion for below.

positions, rather than holding only of A-positions. In the next section, we will examine such a binding theory.

4.2. A Generalized Binding Theory

In recent work,[13] Joseph Aoun has explored the idea of extending the binding theory to cover not only the possible coindexing relationships that may hold between A-positions, but also the possible coindexing relationships that may hold between A-positions and $\bar{A}$-positions. Such a binding theory predicts the existence, for example, of $\bar{A}$-anaphors in addition to the A-anaphors discussed above. In particular, one might imagine that the following possibilities obtain for antecedent/anaphor relations:

(57)	Antecedent	Anaphor
a.	A-position	A-position
b.	A-position	$\bar{A}$-position
c.	$\bar{A}$-position	A-position
d.	$\bar{A}$-position	$\bar{A}$-position

(57a) represents the case where the antecedent and the anaphor both occupy A-positions. This is, of course, the case covered by condition A of the standard binding theory outlined above.

(57b), where the anaphor is in an $\bar{A}$-position and the antecedent is in an A-position, represents a more interesting case. Let us suppose that condition A of the binding theory is

---

13 See, for example, Aoun (1981) and Aoun (1985) for extensive discussion

now formulated as:

(58) An X-anaphor must be X-bound in its governing category.

Thus an A-anaphor must be A-bound in its governing category while an $\bar{A}$-anaphor must be $\bar{A}$-bound in its governing category. We are interested in finding an element that must be A-bound although the element itself occurs in an $\bar{A}$-position. Consider the following set of sentences (from Aoun (forthcoming)) involving quantifier floating in Italian:

(59)a. [i miei amici]$_i$ hanno parlato tutti$_i$ dello stesso problema

´All my friends spoke of the same problem.´

b. *Mario$_i$ ha parlatto tutti$_j$ dello stesso problema

(´*Mario all spoke of the same problem´)

c. Mario$_i$ sostenne che [i miei amici]$_j$ parlarano tutti$_{\{j,\ *i\}}$ dello stesso problema

´Mario said that my friends all speak of the same problem.´

d. *[i miei amici]$_i$ sostenne che Mario parlo tutti$_i$ dello stesso problema.

(´*my friends said that Mario all speaks speaks of the same problem.´)

Example (59a) represents a well-formed instance of quantifier float. In example (59b), however, the quantifier, <u>tutti</u>, has no antecedent since the only available antecedent, <u>Mario</u>, is singular and <u>tutti</u> requires a plural antecedent; thus, it appears

that floated tutti has anaphoric properties insofar as it must have an antecedent somewhere in the sentence. Example (59c) shows, again, a well-formed instance of quantifier float, this time involving a bi-clausal structure. Consider, now, example (59d). The quantifier, tutti, has a NP which agrees with it for number, but this NP is in the matrix clause while tutti is in the embedded clause. Notice, furthermore, that the embedded clause contains a SUBJECT (i.e., either Infl or Mario). It appears, then, that the potential antecedent for tutti in (59d) is too far away to act as its actual antecedent. These distributional facts are remarkably similar to the familiar A-anaphor facts outlined in section 1. Hence, it seems reasonable to suppose that tutti is also an A-anaphor which must be bound in its governing category. Notice, furthermore, that the position that tutti occupies in each of the examples in (59) is not an A-position since it does not receive a θ-role and may not be subcategorized for. Hence, this position is arguably an $\bar{A}$-position. It would appear, then, that A-anaphors may occupy $\bar{A}$-positions, as the schema in (57b) would lead us to believe.

We turn now to schema (57c) which represents an element in an $\bar{A}$-position which locally binds an anaphor in an A-position. This relationship is exemplified by an element which we have already had occasion to deal with; namely, wh-trace which must occur in an A-positions and must be bound by an element in an $\bar{A}$-position. The standard Binding theory outlined in section 1 requires only that these elements not be locally bound by an

element in an A-position. As a quick inspection of the Binding theory (collected in (56), above) will verify, the Binding theory does not explicitly require that wh-trace be bound by an element in an $\bar{A}$-position. In a such a theory, the $\bar{A}$-binding of variables is guaranteed by a principle which lies outside the Binding theory per se. In general, it is assumed that unbound variables are ruled out, along with vacuous quantification, by the Bijection Principle (Koopman & Sportiche (1981)).

A generalized Binding theory would treat the operator/variable relationship as falling under condition A of the Binding theory. At the same time, we have seen above that variables appear to obey condition C of the Binding theory. At first glance, these principles would seem to be in conflict since condition A requires that an element be bound in its governing category while condition C requires an element to be free. How is it possible to satisfy both requirements at once. It is helpful to think of a generalized Binding theory as delineating two parallel systems of binding. One, the A-system, holds of relationships between A-positions only. The other, $\bar{A}$-system, adds $\bar{A}$-positions to the set of possible relationships. Thus, it is possible to obey condition A of the $\bar{A}$-system and condition C of the A-system at the same time.

Let us suppose, then, that variables are anaphors with respect to the $\bar{A}$-system of binding and R-expressions with respect to the A-system of binding. We will begin with variables in the

subject position of tensed clauses. We have already seen that the subject position of tensed clauses has an accessible SUBJECT; namely, the tense element of Infl acts as the SUBJECT. Notice, further, that Infl does not occupy an argument position since this position is never assigned a θ-role; thus it does not follow from the fact that tense and Infl are coindexed that Infl (or tense) may A-bind the subject position. We predict, however, that a variable in the subject position of a tensed clause does have a governing category.

In the A-system of binding outlined above, the governing category of an element in the subject position of a tensed clause would be S. Suppose that we tried to maintain this position in a generalized Binding theory. The governing category for a variable in subject position, as in (60), would be S:

(60) $[_{\bar{S}}$ who$_i$ $[_S$ $\underline{t}_i$ Infl saw John]]

Notice that the variable in the subject position of (60) is not $\bar{A}$-bound in S; its local $\bar{A}$-binder is in Comp which is not contained in S. Furthermore, there is no $\bar{A}$-position internal to S that could bind a variable in the subject position. It would seem, at first blush, that the generalized Binding theory rules out extraction from the subject position. This problem is rather simply solved if we allow the categories that the definition of governing category to range over to include $\bar{S}$ rather than S. This change in the categories available to the definition of governing category will not have any effect on the A-binding

system. Consider an anaphor in the subject position of a tensed clause, as in:

(61) $*[_{\bar{S}}$ Comp $[_{S}$ himself Infl bothers John]]

The category, $\bar{S}$, does not contain any A-positions not contained in S; hence, the change from S to $\bar{S}$ will not make available any position position that could possibly save the derivation of a structure similar to (61) in the relevant respects.

We have seen that a variable in the subject position of a tensed clause must be bound in the minimal $\bar{S}$ that contains given condition A of a generalized Binding theory. This requirement has the effect of forcing an wh-element, for instance, in the subject position of a tensed clause to move to the Comp that governs this position. This effect has a number of interesting ramifications; consider, for example, a standard complementizer-trace effect as in (62) (see Pesetsky, 1982):

(62) $*[_{\bar{S}}$ who$_i$ $[_{S}$ did John say $[_{\bar{S}}$ $\underline{t}_i$ that $[_{S}$ $\underline{t}_i{}'$ Infl left]]]]

Recall that an empty category in a Comp which also contains a complementizer, in this case, <u>that</u>, may not c-command out of that Comp. The trace in Comp in (62) does not c-command $\underline{t}'$; but if the trace in Comp does not c-command the variable in subject position of the embedded clause, the variable cannot be bound by the trace in Comp. A variable in the subject position of a tensed clause must be bound in the minimal $\bar{S}$ that contains it, as we have already seen. Thus, the trace in the subject position of

the embedded clause in (62) violates condition A of the $\bar{A}$-system since it is not $\bar{A}$-bound in its governing category. Notice that we did not need to make any appeal to the ECP in ruling out the derivation of (62); as we will see below, many standard ECP effects may be derived by appeal to the generalized Binding theory.

If a variable in the subject position of a tensed clause must be $\bar{A}$-bound in the minimal superordinate $\bar{S}$, how may we account for the distribution of variables in a position other than subject? In particular, can a generalized Binding theory account for the lack of complementizer-trace effects for variables that are, for example, in the object position.? Consider a representation like that in (63):

(63) $[_{\bar{S}}$ who$_i$ $[_S$ did John say $[_{\bar{S}}$ $\underline{t}_i$ that $[_S$ Bill Infl saw $\underline{t}_i'$]]]]

The variable in the object position of the embedded clause (i.e., $\underline{t}'$) is clearly not $\bar{A}$-bound in the minimal superordinate $\bar{S}$ since the trace in the embedded Comp position, $\underline{t}$, may not c-command out of Comp by virtue of the presence of the complementizer, <u>that</u>. At this point, the basic tension between the A-system of binding and the $\bar{A}$-system of binding becomes relevant. We argued above that variables must obey condition C and of the A-system of binding and condition A of the $\bar{A}$-system of binding at the same time; in other words, a variable must be A-free (with the proviso noted in section 1) and it must be $\bar{A}$-bound in some local domain.

The problem is one of calculating the local domain in which the variable must be $\bar{A}$-bound. Let us return to the representation in (63). By convention, an agreement rule coindexes the element in subject position with the (tensed) Infl that commands it. Thus, the indexing in (63) is as follows:

(64) $[_{\bar{S}}$ who$_i$ $[_S$ did John say $[_{\bar{S}}$ $\underline{t}_i$ that $[_S$ Bill$_j$ Infl$_j$ saw $\underline{t}_i$ ´]]]]

where Bill and the Infl node of the embedded clause bear the same index. Now, recall that the definition of governing category made crucial reference to an accessible SUBJECT. Roughly speaking, a SUBJECT is accessible to some element, A, if the coindexation of that SUBJECT and A is well-formed. Is the Infl of the embedded clause an accessible SUBJECT for an element in object position? We can change the indexing on (64) in such a way that Infl and the object are coindexed:

(64) *$[_{\bar{S}}$ who$_i$ $[_S$ did John say $[_{\bar{S}}$ $\underline{t}_i$ that $[_S$ Bill$_i$ Infl$_i$ saw $\underline{t}_i$ ´]]]]

Since the subject position is obligatorily coindexed with Infl, coindexing a variable in object position of a clause with Infl means that the variable will also be coindexed with the subject by the transitivity of the coindexing relation. Thus, the indexing shown in (64) inevitably leads to a violation of condition C of the A-system of binding since the variable in object position will be A-bound by the subject in the domain of the variable´s operator. Hence, the Infl of the embedded clause

is not an accessible SUBJECT for a variable in non-subject position since coindexation of the two elements inevitably leads to ill-formedness (we will see an exception to this below).

The line of reasoning above has a rather curious result. Let us continue the argument to consider whether or not the subordinate object has an accessible SUBJECT in the superordinate clause. Assuming that we coindex the Infl of the superordinate clause with the object of the embedded clause, the following representation will result:

(65) $*[_{\bar{S}}$ who$_i$ $[_S$ did John$_j$ Infl$_j$ say $[_{\bar{S}}$ $\underline{t}_i$ that $[_S$ Bill$_i$ Infl$_i$ saw $\underline{t}_j$ ´]]]]

Since the subject of the superordinate clause is obligatorily coindexed with the Infl of the superordinate clause, coindexing between this Infl and the subordinate object once again results in the subordinate object being A-bound, in violation of condition C of the A-binding system. Since the coindexing of the superordinate Infl with the subordinate object is not well-formed, the superordinate Infl is not an accessible SUBJECT for the embedded object. The superordinate clause cannot be the governing category for the embedded object since it does not contain a SUBJECT accessible to the object.

The above line of reasoning can be carried on ad infinitum for any depth of embedding. It would appear that variables which are not in the subject position of a clause never have an accessible SUBJECT since any possible coindexing of the variable

with a SUBJECT leads to a violation of condition C. But this means that, unless we make some stipulation, variables in a non-subject position will always lack a governing category. We could, of course, follow Chomsky (1981) in stipulating that elements that otherwise lack a governing category will have the root $\bar{S}$ as their governing category by default. There is no real need to do this, however; extraction from non-subject position are in general unbounded, ignoring, of course, subjacency effects.

Consider, now, a multiple wh-question like:

(66) Who ate what?

Extraction of the subject applies in the syntax. Since the matrix Comp is now occupied, we will assume, as is standard, that a rule of Comp indexing applies which copies the index of the element inside Comp onto the Comp node itself. The variable in subject position is now $\bar{A}$-bound in its governing category as required by condition A of the $\bar{A}$-system of binding. The S Structure of (66) is thus:

(67) $[_{\bar{S}}\ [_{Comp}\ \text{who}_{i}\ ]_{i}\ [_{S}\ \underline{t}_{i}\ \text{ate what}]]$

At LF, move alpha will apply to the wh-element in object position, adjoining it to the matrix Comp:

(67) $[_{\bar{S}}\ [_{Comp}\ \text{what}_{j}\ \text{who}_{i}\ ]_{i}\ [_{S}\ \underline{t}_{i}\ \text{ate}\ \underline{t}_{j}\ ]]$

Notice that the Comp node in (67) is branching; neither <u>who</u> nor <u>what</u> may c-command out of this Comp. Since Comp indexing has

applied at S Structure (as discussed in Aoun, forthcoming, and the references cited there), the variable in subject position will be bound by the Comp node itself. The variable in object position will not, strictly speaking, be bound by what. As we have seen, above, the variable in object position has no governing category; since this is the case, we may say that the variable in object position in (67) is trivially bound in its (non-existent) governing category. We also want to say that, in some sense, the variable in object position is bound by what because it must pick up a range from what. It would appear, then, that two senses of binding are required. One sense, which we may call strict binding, is relevant to the binding theory; another, looser, sense of binding is relevant to the assignment of scope. Assuming that this is the case, then the generalized binding theory does not eliminate the need for a principle that independently filters out unbound variables (e.g., the Bijection Principle) since some variables will be only loosely bound and not bound in the strict fashion required by the Binding theory.

We saw above that the generalized Binding theory may be used to filter out some representations that are also filtered out by the ECP; in particular, the generalized Binding theory covers the complementizer-trace effects. The generalized Binding theory may also be used to rule out other ECP violations. Consider the superiority effect in (68):

(68) *what did who see

In (68), the wh-element in object position has been moved to Comp

in the syntax. At S Structure, Comp Indexing will apply, copying the index of the wh-element onto the Comp node; the resulting representation is illustrated in (69):

(69) $*[_{\bar{S}}\ [_{Comp}\ \ \text{what}_{j}\ ]_{j}\ [_{S}\ \text{did who see}\ \underline{t}_{j}\ ]]$

Move alpha will apply to the wh-element in subject position during the mapping to LF. After <u>who</u> has been adjoined to Comp, the representation is:

(70) $*[_{\bar{S}}\ [_{Comp}\ \ \text{who}_{i}\ \ \text{what}_{j}\ ]_{j}\ [_{S}\ \text{did}\ \underline{t}_{i}\ \ \text{see}\ \underline{t}_{j}\ ]]$

<u>Who</u> does not strictly bind the variable in subject position in (70), but only loosely binds it. For purposes of the Binding theory, the variable in the subject position of (70) is simply not bound. Recall that variables in the subject position of a tensed clause do, in fact, have a governing category since Infl acts as an accessible SUBJECT. Example (70), therefore, is a violation of condition A of the $\bar{A}$-system of binding since the variable in subject position is free for purposes of the Binding theory.

It appears, then, that a generalized Binding theory may be used to capture many of the phenomena derived by the ECP. Let us suppose that we replace the ECP with a generalized Binding theory. If we are to rule out superiority violations like (70) by means of the generalized Binding theory described up to this point, then it must be the case that this Binding theory applies at LF since the illicit variable in examples like (70) are not

available until the level of LF representation. I will therefore assume that a generalized Binding theory constrains LF representations. For the moment I will leave open the question of whether the generalized Binding theory applies at any level (or levels) other than LF.

It is not, however, immediately obvious that a generalized Binding theory can handle all the effects attributed to superiority. In particular, consider sentences like the following:

(71)a. $*[_{\bar{S}}$ what $_i$ $[_S$ did John order $\underline{t}_i$ why]]

b. $[_{\bar{S}}$ why $_j$ $[_S$ did John order what $\underline{t}_j$]]

(72)a. $*[_{\bar{S}}$ what $_i$ $[_S$ did John build $\underline{t}_i$ how]]

b. $[_{\bar{S}}$ how $_j$ $[_S$ did John build what $\underline{t}_j$]]

A standard ECP account of the facts in (71) and (72) would claim that the positions occupied by <u>why</u> in (71a) and <u>how</u> in (72a) is not a position which is properly governed by a lexical head. In order for a variable in these positions to be properly governed, then, they must be antecedent governed. The adjunct positions in (71a) and (72a) may only be antecedent governed from Comp. Since Comp indexing applies a S Structure and the adjuncts in the ungrammatical examples have not moved to Comp during the syntax, LF movement of the adjuncts will result in ungrammaticality. In the well-formed examples, (71b) and (72b), the adjuncts have moved to Comp in the syntax. When Comp indexing applies at S Structure, the indices of <u>why</u> and <u>how</u> will be copied onto the

Comp node; as a result, the variables left by these adjuncts will be antecedent governed at LF. Examples (71b) and (72b) do not, unlike examples (71a) and (72a), violate the ECP. Notice that a variable in object position will never violate the ECP since this position is lexically governed by the verb.

In terms of a generalized Binding theory, the question becomes one of what governing category an adjunct trace has. Variables in A-positions other than subject have no governing category since they never have an accessible SUBJECT. This followed from the fact that variables are subject both to condition A of the $\bar{A}$-binding system and condition C of the A-binding system. As a result, variables in such positions need not be $\bar{A}$-bound in a local domain. It is, therefore, licit to leave wh-elements in non-subject position in situ since LF movement will not result in a violation of condition A of the $\bar{A}$-system.

The examples, above, show that the traces left by movement of the adjuncts <u>why</u> and <u>how</u> must be $\bar{A}$-bound in a local domain. It would appear, then, that the variable left by extraction of one of these elements has a governing category. But the only way that their variables may have a governing category is if they are both governed and have an accessible SUBJECT inside that domain. It must be the case, then, that coindexing <u>why</u> and <u>how</u> with Infl does not lead to a violation of condition C. In the case of variables in argument positions, condition C was violated because coindexing the variable with Infl resulted in A-binding of the

variable by the subject. Note that the subject also c-commands these adjuncts, as the following example illustrates:

(73) *he$_i$ left because John$_i$ was angry

In (73) the pronoun, he, in subject position c-commands and is coindexed with an R-expression inside an adjunct. Thus, (73) violates condition C of the A-binding system.

Given that the subject c-commands adjuncts, then coindexing a variable in an adjunct position will mean that the subject A-binds the variable. Aoun (forthcoming) argues, however, that adjuncts, like the manner and reason wh-elements how and why, are not referential in the same sense that arguments like who and what are. The latter two elements range over individuals in domain D while the latter two do not. If Aoun is correct, then there is no reason to expect that the variables left by extraction of these non-referential elements will be subject to condition C. Aoun concludes that the traces left by extraction of why and how are subject only to condition A of the $\bar{A}$-system and are completely indifferent to condition C of the A-binding system.

It is of some interest that, while locative and temporal adjuncts have deictic elements (e.g., here, there, then, etc.), manner and reason adjuncts do not. If we assume that these deictic elements indicate that such adjuncts may range over individuals then, by the above line of reasoning, variables left by the extraction of such adjuncts should be subject to condition

C of the A-system. We would not, therefore, expect syntactic movement of wh-elements in these positions to be obligatory; in short, wh-elements like where and when should not show the same superiority effects that why and how do. This prediction does seem to be borne out by the facts; compare the status of (71) and (72) with the following:

(74)a. $[_{\bar{S}}$ what $_i$ $[_S$ did John buy $\underline{t}_i$ where]]

b. $[_{\bar{S}}$ where $_i$ $[_S$ did John buy what $\underline{t}_i$]]

(75)a. $[_{\bar{S}}$ who $_i$ $[_S$ did John contact $\underline{t}_i$ when]]

b. $[_{\bar{S}}$ when $_i$ $[_S$ did John contact who(m) $_i$ $\underline{t}_i$]]

Given the assumption that temporal and locative adjuncts may refer to individuals, these elements are subject to condition C of the A-binding system. Furthermore, variables are always subject to condition A of the $\bar{A}$-binding system. Suppose that the variable in (75b) were to have Infl as an accessible SUBJECT; coindexing this variable with Infl should be well-formed:

(76) $[_{\bar{S}}$ where $_i$ $[_S$ did John $_i$ Infl $_i$ buy what $\underline{t}_i$]]

In (76), the subject, John, A-binds the variable left by where. Since temporals refer to individuals in domain D, however, they are subject to condition C and may not be A-bound in the domain of their operator. The indexing shown in (76) cannot be well-formed; the Infl of the matrix clause may not act as the accessible SUBJECT for the variable left by where. If this is so, then the variable left by where lacks a governing category

and is, therefore, trivially treated as $\bar{A}$-bound in its governing category.

Turning to the LF representation of (74a), which is shown in (77):

(77) $[_{\bar{S}}\ [_{Comp}\ \text{where}_j\ \text{what}_i\ ]\ [_{S}\ \text{did John buy}\ \underline{t}_i\ \underline{t}_j\ ]]$

we can see that the variable left by <u>where</u> is not bound from Comp for the purposes of the generalized Binding theory. Given the discussion above, however, we know that the variable in question lacks a governing category. Hence, the variable of <u>where</u> trivially satisfies the generalized Binding theory, even though the adjunct, <u>where</u>, was left in situ in the syntax.

At least two observations should be made about the generalized Binding account of superiority effects with adjuncts. First, this account is not purely structurally based. This is not to imply that structural relations do not come into play in this account. Clearly, structural notions like c-command and accessibility are fundamental to the analysis. It is equally true that the generalized Binding account of the differences between the examples in (71) and (72), on the one hand, and the examples in (73) and (74), on the other, does not make any crucial reference to differences in hierarchical structure. That is the following examples may have the same structure, abstracting away from individual lexical items:

(78)a. $[_{\bar{S}}\ \text{what}_i\ [_{S}\ \text{did John order}\ \underline{t}_i\ \text{where}]]$

b. $*[_{\bar{S}}\ \text{what}_i\ [_{S}\ \text{did John order}\ \underline{t}_i\ \text{why}]]$

The difference in grammaticality between the two examples arises from the difference in the way that where and why range over domain D; the crucial difference being that where ranges over individuals in D while why ranges over propositions. A proponent of a generalized Binding account of these superiority effects does not, therefore, need to appeal to structural differences between sentences like (78a) and (78b). He is able to discount these putative structural differences only at a certain cost, however. The price to be paid is that the theory of reference must play a more central role in his syntactic theory; the practitioner of generalized Binding must be willing to develop some theory of reference and to admit that this apparently purely semantic subsystem of grammar is directly reflected in syntactic levels of representation. The idea that the semantics has direct reflections in the syntax is not, however, a particularly revolutionary one; the Projection Principle, for example, requires that the lexical semantics of a head be reflected syntactically in the preservation of argument positions required by that head across syntactic levels of representation. It should come as no surprise, then, that the referential properties of individual items conditions the way in which these items behave with respect to the Binding theory.

The second observation is simply that we have tacitly been assuming throughout the above discussion on the binding of variables in adjunct positions that adjuncts occupy A-positions.

In other words, the relationship between where and its variable in (79) is exactly analogous to the relationship between an operator and an argument:

(79) $[_{S}$ where $_{i}$ $[_{S}$ did John buy the book $\underline{t}_{i}$]]

The validity of this assumption is open to question, however. Let us assume, as is standard, that A-positions are positions that are present to meet the lexical requirements of some head. This is the content of the requirement that A-positions are all and only those positions that may be assigned a thematic role. Adjuncts, however, are never present to satisfy the lexical requirement of a head; their presence is optional and their omission generally preserves the grammaticality of the structure in question, although it is not obvious the optionality alone is a sufficient condition for being an adjunct. Thus, the examples in (80) are both fully grammatical:

(80)a. John bought the book downtown.
b. John bought the book.

Given that adjuncts do not occupy A-positions in the normal sense in which the term A-position is taken, it is not at all clear that the behavior of variables in adjunct positions should be parallel to the behavior of variables in certified A-positions.

This puzzle is less worrisome that one might imagine, however. The goal of a generalized Binding theory has been to extend the Binding theory across possible positions. The relevant point of orientation in this system is the position of the antecedent. We have already seen that A-anaphors may occupy

$\bar{A}$-positions (see the discussion of Italian tutti, above). We would naturally expect $\bar{A}$-anaphors to be able to occupy $\bar{A}$-positions, as indicated in by the schema under (57d). If we can show that there is a genuine example of schema (57d), then the puzzle outlined above would appear to pose little immediate threat to the system of generalized Binding outlined in this section.[14]

Aoun (forthcoming) argues that intermediate traces in Comp are pure $\bar{A}$-anaphors which, themselves, occupy $\bar{A}$-positions. Consider a representation like:

(81) $[_{\bar{S}}$ who $_i$ $[_S$ John Infl say $[_{\bar{S}}$ $\underline{t}_i$ $[_S$ $\underline{t}_i'$ Infl bought the book]]]]

The trace in the embedded subject position, $\underline{t}'$, is the subject of a predication and, hence, bears a thematic role. Thus, this trace is subject to condition A of the $\bar{A}$-binding system and condition C of the A-binding system. Since the tensed Infl is an

---

14 We are still left with the question of what sort of position an "adjunct position" is. We argued above that adjuncts do not occupy A-positions by virtue of the definition of A-position. Does it then follow that adjuncts occupy A-bar-positions? If we insist that A-bar-positions are positions which may operator-bind an argument (in the sense that a wh-element in Comp binds a variable) then a distinction between adjuncts and A-bar-positions should be made. One possibility is to partition A-bar-positions into "operator positions" (i.e., positions that may contain an operator that binds an element in an A-position) and "adjunct positions". At this point, of course, the distinction is purely taxonomic and of virtually no explanatory value. Such a distinction would gain in explanatory value by entering into a system; the development of such a system would, of course, take me well out of the bounds of this dissertation.

accessible SUBJECT for the variable in the embedded S, that variable must find an antecedent inside the embedded $\overline{S}$ (by condition A). Thus, a representation like (81) is well-formed only if there is a trace in Comp to $\overline{A}$-bind the variable in the embedded subject position. Having established the obligatory presence of $\underline{t}$ in (81), we can now enquire into its properties. This trace does not, itself, bear a θ-role--at most it transmits the θ-role predicated of the embedded subject to the operator in the matrix Comp--nor does this trace have any inherent range to restrict the variable in the embedded subject position. The variable in (81) must receive its range from the wh-element in Comp. If it is true that the trace in Comp neither bears a θ-role nor has a range, then we may assume that the trace is, in some sense, referentially empty; it lacks any inherent capacity to range over elements of domain D. But if traces in Comp lack the ability to select elements of domain D, we would not expect traces in Comp to be subject to condition C of the Binding theory.

Traces in Comp must, nevertheless, be bound from an $\overline{A}$-position. It is not possible, for example, to leave a trace in Comp unbound as in (82):

(82) *John said $[_{\overline{S}}\ \underline{t}_i\ [_S\ \underline{t}_i{}'$ Infl left]]

Nor is it possible for the antecedent of a trace in Comp to occupy an A-position as in (83):

(83) *John$_i$ same $[_{\overline{S}}\ \underline{t}_i\ [_S\ \underline{t}_i{}'$ Infl left]]

We may safely conclude that a trace in Comp is an $\overline{A}$-anaphor.

Since a trace in Comp is not subject to condition C and since it is a pure $\overline{A}$-anaphor, it follows that coindexing the trace in Comp with the Infl of the superordinate clause, as in (84), will not trigger a violation of condition C of the Binding theory:

(84) $[_{\overline{S}}$ who$_i$ $[_S$ John$_j$ Infl$_j$ say $[_{\overline{S}}$ $\underline{t}_j$ $[_S$ Mary Infl saw $\underline{t}_i$]]]

Assuming that this approach is correct, then it follows that the Infl of the superordinate clause is an accessible SUBJECT for the trace in Comp. But, since the trace in Comp in the above examples has an accessible SUBJECT, it also has a governing category, namely, the superordinate $\overline{S}$. By condition A of the $\overline{A}$-binding system, the trace in Comp must be $\overline{A}$-bound in that governing category.

Let us, then, adopt the assumption that traces in Comp are pure $\overline{A}$-anaphors, not subject to condition C of the Binding theory. It follows that a trace in Comp must always be $\overline{A}$-bound by an element in the superordinate $\overline{S}$. Thus, whenever an element moves from Comp to Comp, it must move to the Comp of the superordinate $\overline{S}$. This makes a rather interesting predication: Even in a language where only $\overline{S}$, and not both S and $\overline{S}$, is the bounding node for subjacency, Comp to Comp movement must proceed to the immediately superordinate Comp. The following situation is, therefore, ruled out:

(85) $[_{\bar{S}}\ Op_1\ [_S \ldots [_{\bar{S}}\ wh_2\ [_S \ldots t_2 \ldots [_{\bar{S}}\ t_1\ [_S \ldots t_1 \ldots]]]]]$

even in a thematically based formulation of subjacency that does not count the most deeply embedded $\bar{S}$ as a bounding node. This is because the trace in the lowermost Comp is not $\bar{A}$-bound in its governing category, the immediately superordinate $\bar{S}$.

The discussion so far has given support for the idea that the Binding theory may be generalized across positions, as indicated by the diagram in (57). We have so far, however, only considered principle A of the $\bar{A}$-binding system. This has already given us a more articulated system of binding dependencies. We could, of course, assume that a generalized Binding theory extends to condition B of the A-binding system and, possibly, condition C. Extending the generalized Binding theory to condition B, for example, would imply the existence of categories that must be $\bar{A}$-free in their governing category, a possibility that may be instantiated by resumptive pronouns (see Sells, 1984 and the references cited there).[15] If we were to extend the Binding theory in this way, we would expect a number of predictions just as we found in the case of the extension of condition A discussed above. I refer interested readers to Finer

15 It should be noted that I have nothing to say about resumptive pronouns and I am assuming that they differ radically from the pronominals I will be looking at; to be specific, resumptive pronouns may, apparently, be locally A-bar-bound. Recall that I am assuming that pronominals may not be A-bar-bound; thus, I must assume that resumptives have very special properties.

(1984) for some discussion of this extension of condition B.

It is less clear that condition C may be generalized across positions in quite the way done for condition A. Such an extension would imply the existence of elements that must be $\bar{A}$-free in every governing category (although they may be A-bound). It seems, for example, that names like John have the property that they must be $\bar{A}$-free (in addition to A-free) in every governing category, as the following attests:

(86) *him$_i$ Mary persuaded t_i that Janet likes John$_i$

It is not clear, though, that the fact that names may not be $\bar{A}$-bound follows from a generalization condition C or from more general properties of the theory of reference. If the latter is the case, then condition C may follow as theorem of this theory of reference.

In the remainder of this chapter, we will mainly be involved with some of the ramifications of a generalized version of condition A. We will, therefore, leave further generalizations of the Binding theory as a problem for future research.

4.3 Non-overt operators and the relationship between binding and predication

We saw, in section 4.1.3, above, that some appeal had to be made to the existence of variables that are bound by phonologically empty operators. It was assumed that these non-overt operators occur in structures like the following:

(87)a. John$_i$ is easy [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to please t_i]]

b. John bought the book $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to read $\underline{t}_i$ to the children]]

c. [the man $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to see $\underline{t}_i$]]] is John

Example (87a) illustrates a case of "tough movement", example (87b) is a purpose clause, and example (87c) is an infinitival relative. As has been discussed in a variety of places (see, for example Chomsky (1977), Bach (1982) and Fiengo (1980), these constructions show a number of properties that are similar to wh-movement. There appear, for example, to be wh-island effects:

(88)a. *John$_i$ is easy $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to tell Mary $[_{\bar{S}}$ how $[_S$ PRO to please $\underline{t}_i$]]]]

b. *John made up the story $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to show Bill $[_{\bar{S}}$ how $[_S$ PRO to trick to the children with $\underline{t}_i$]]]]

c. *[the man $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to show Bill $[_{\bar{S}}$ where $[_S$ PRO to see $\underline{t}_i$]]]]] is John

The ungrammaticality of the examples in (88) does not follow from a strict boundedness condition between the overt element in an A-position and the variable itself: The variable may be very deeply embedded as the examples in (89) show:

(89)a. John$_i$ is easy $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to persuade Mary $[_S$ PRO to please $\underline{t}_i$]]]

b. John made up the story $[_{\bar{S}}$ Op$_i$ for $[_S$ Bill $[_S$ PRO to trick to the children with $\underline{t}_i$]]]

c. [the man $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to tell Bill $[_S$ PRO to see $\underline{t}_i$]]]] is John

The variable is well-formed in positions from which wh-movement is well-formed, no matter how deeply embedded the extraction site is.

Furthermore, recall that variables must be free in the domain of their operator. Assuming that the empty categories in the examples of (87) and (89) are variables, notice that each one is coreferential with an element in an A-position. If there were no operator to locally $\bar{A}$-bind these variables, we would expect the examples in (88) and (89) to be ungrammatical. We have, in fact, little choice but to accept these putative variables as genuine variables given the current typology of empty categories. They cannot be NP-trace since they are not locally A-bound. They cannot be pro because they are not locally identified. They cannot be PRO since they occur in governed, Case-marked positions. The only possibility that we are left with is variable.

We have seen that there is strong evidence, both empirical and theory internal, for the existence of non-overt operators. It has recently been proposed, by Contreras (1984), that non-overt operators are crucially involved in parasitic gap constructions. Parasitic gap constructions are exemplified in (90):

(90)a. ?which articles$_i$ did John file t_i without reading e_i

b. ?the man who$_i$ Mary called t_i an idiot as often as Jane called e_i a cretin

where t represents the true gap and e represents the parasitic gap. The analysis of parasitic gaps given in Chomsky (1982) assumed that the parasitic gap position is empty a D Structure. At D Structure, this gap would necessarily have the properties of a pronominal since it is not locally A-bound or $\bar{A}$-bound. Thus, the sentences in (90) would have a D Structure of the following form:

(91)a. $[_{\bar{S}}$ $[_{S}$ John Infl file which articles [without $[_{S}$ PRO reading pro]]]]

b. $[_{NP}$ the man $[_{\bar{S}}$ $[_{S}$ Mary Infl called who an idiot [as often as $[_{S}$ Jane called pro a cretin]]]]]]]

Let us suppose that wh-movement applies to some element in the mapping from D Structure to S Structure, as in (92):

(92)a. $[_{\bar{S}}$ which articles $_{i}$ $[_{S}$ John Infl file t_{i} [without $[_{S}$ PRO reading pro]]]]

b. $[_{NP}$ the man $[_{\bar{S}}$ who $_{i}$ $[_{S}$ Mary Infl called t_{i} an idiot [as often as $[_{S}$ Jane called pro a cretin]]]]]]]

Let us suppose, following Chomsky (1982), that an indexing procedure applies at S Structure which assigns an index to all A-positions. If this indexing procedure assigns an index to the pros which is disjoint from the index on the wh-element/variable pair, then the pro will be both $\bar{A}$-free and A-free. Since the pro is $\bar{A}$-free, it cannot be treated as a variable. Since the pro is A-free, it may not be treated as an NP-trace. Thus, the pro must

either remain a pure pronominal or be treated as PRO. Since the empty category is not locally identified by some agreement-like element, we cannot treat it as pro. Hence, the empty category must be PRO:

(93)a. $[_{\bar{S}}$ which articles $[_{S_i}$ John Infl file $\underline{t}_i$ [without $[_{S_j}$ PRO reading PRO]]]]

b. $[_{NP}$ the man $[_{\bar{S}_i}$ who $[_{S_j}$ Mary Infl called $\underline{t}_i$ an idiot [as often as $[_{S_j}$ Jane called PRO a cretin]]]]]]]

But this PRO occurs in a governed (and Case-marked position). The representation in (93) will, therefore, be marked as ungrammatical.

Suppose, on the other hand that the empty category had been freely coindexed with the operator:

(94)a. $[_{\bar{S}}$ which articles $[_{S_i}$ John Infl file $\underline{t}_i$ [without $[_{S_i}$ PRO reading $\underline{e}_i$]]]]

b. $[_{NP}$ the man $[_{\bar{S}_i}$ who $[_{S_i}$ Mary Infl called $\underline{t}_i$ an idiot [as often as $[_{S_i}$ Jane called $\underline{e}_i$ a cretin]]]]]]]

The empty category, $\underline{e}$, in the examples in (94) is now $\bar{A}$-bound and can be treated as a variable. This variable did not, however, arise from movement; it is treated as a variable through free assignment of indices at S Structure. In fact, a variable in the position of the parasitic gap in the above structures could not have arisen from movement since movement from the parasitic gap position to the matrix Comp would violate subjacency:

(95)a. *[$_{\bar{S}}$ which books$_i$ [$_S$ John Infl file the articles [without [$_S$ PRO reading $\underline{e}_i$]]]]

b. *[$_{NP}$ the man [$_{\bar{S}}$ who$_i$ [$_S$ Mary Infl called John an idiot [as often as [$_S$ Jane called $\underline{e}_i$ a cretin]]]]]]]

Thus, parasitic gap structures appeared to provide motivation for the "contextual definition of empty categories" (i.e., the idea that the properties of empty categories could be deduced from independent facts about the environment they occur, implying that empty categories have no inherent features) and for the idea that subjacency is a constraint on rules and not a constraint on representations.

Chomsky's analysis of parasitic gaps has the virtue of deriving the following facts, as the reader may verify by consulting the analysis in Chomsky (1982):

(96)a. A parasitic gap is licensed by a variable that does not c-command it.

b. A parasitic gap cannot be licensed by an element in an A-position even if that element is a wh-element in situ.

As pointed out in Kayne (1984), Chomsky's analysis predicts that a parasitic gap may be generated anywhere just so long as it is not c-commanded by the variable that licenses it. As Kayne and Contreras have shown, this is not correct; Kayne notes contrasts like the following:

(97)a. ?the person that John described $\underline{t}$ without examining any pictures of $\underline{e}$

b. *the person that John described $\underline{t}$ without any pictures of $\underline{e}$ being on file

c. ?the paper that we should destroy t before someone steals a copy of e

d. *the paper that we should destroy t before a copy of e gets stolen by someone

Kayne derives the above contrasts by appeal to his reformulation of the ECP (see the discussion in chapter 3). Contreras, on the other hand, argues that the above facts follow if the parasitic gap arises through movement of a non-overt operator. In particular, example (97a) would have a D Structure like:

(98) [the person [that [John described WH without [PRO examining any pictures of Op]]]]]
NP S S S S

where WH stands for the relative operator, and Op stands for the non-overt operator. In the syntax, "move alpha" would apply both to the relative operator, WH, and to the non-overt operator, Op:

(99) [the person [that WH [John described t
NP S i S i
without [Op [PRO examining any pictures of
S j S
t]]]]]
j

Below, we will return to the question of how the non-overt operator assigns the same range as the "overt" operator.[16] For the moment, we will simply assume that some mechanism independently guarantees this identity.

---

[16] Overt is in quotation marks since the relative clause operator may be deleted. I assume that this deletion is of a wholly different nature than the non-overt operator phenomenon. Even if non-overt operators involve deletion (and note that this deletion is apparently obligatory) the concerns about the distribution of non-overt operators discussed below would still apply.

If the non-overt operator approach is correct, then a number of otherwise fairly puzzling facts follow immediately. The contrast noted by Kayne (the examples in (97), above) follow either from the Constraint on Thematic Domains developed in chapter 3 or from Chomsky's recent reformulation of subjacency. Subjacency should, furthermore, derive facts like the following:

(100)a. *the woman who$_i$ John saw $\underline{t}_i$ without wondering who$_j$ met $\underline{e}_i$

b. *the paper which$_i$ John studied $\underline{t}_i$ after Bill talked to him without reading $\underline{e}_i$

c. *the story which$_i$ John believed $\underline{t}_i$ without realizing the fact that Bill made $\underline{e}_i$ up

Example (100a) involves extraction of the non-overt operator out of a wh-island and is, therefore, ruled out by subjacency. Example (100b) involves extraction of the non-overt operator out of an adjunct and is ruled out either by subjacency or the constraint in chapter 2. Example (100c) involves extraction of the non-overt operator out of a complex NP and is, similarly, ruled out.

We have seen that there is good reason for believing that non-overt operators are involved in a number of fairly diverse constructions. Given the existence of non-overt operators, consider the following representations:

(101)a. $[_{\bar{S}}$ Op$_i$ $[_S$ John hit $\underline{t}_i$]]

b. $[_{\bar{S}}$ Op$_i$ $[_S$ Mary believes $[_{\bar{S}}$ $\underline{t}_i$ $[_S$ $\underline{t}_i$' saw Bill]]]]

c. $[_{\bar{S}}\ [_{S}$ Mary knows $[_{\bar{S}}$ Op$_i$ $[_{S}$ $\underline{t}_i$ saw Bill]]]]

which would be associated with the following strings:

(102)a. *John hit
b. *Mary believes saw Bill
c. *Mary knows saw Bill

While examples (101a) and (101b) could be eliminated by the ECP, granted the assumption that non-overt operators must be properly governed, example (101c) cannot be so eliminated since the non-overt operator is in a Comp which is, presumably, properly governed by the verb <u>know</u>. Consider, for example, (103) which involves successive cyclic movement:

(103) $[_{\bar{S}}$ who$_i$ $[_{S}$ does Mary know $[_{\bar{S}}$ $\underline{t}_i$ $[_{S}$ $\underline{t}_i$ saw Bill]]]]

Example (103) illustrates the fact that <u>know</u> may properly govern into Comp. Stating the problem slightly differently, why is it that we do not get apparent violations of the Projection Principle like those illustrated in (102)?

The fault does not lie with the variables in the examples in (101). Each of the variables in (101) is properly governed, $\bar{A}$-bound in its governing category and the non-overt operator is not extracted out of a thematically opaque domain. We can safely assume that the ungrammaticality of the above examples does not follow from the violation of a core principle on the part of the variable. This leaves the non-overt operator as our culprit. Some principle, or set of principles acts to constrain the distribution of non-overt operators.

In a recent paper (Aoun & Clark, 1985), J. Aoun and I

argued that non-overt operators have anaphoric properties. The guiding assumption was that, in a sense, non-overt operators are truly empty in that they simply lack a range. This means that a non-overt operator has no inherent restrictions. Consider, for example, a wh-element like who. It is reasonable to assume that such an overt wh-operator has lexical specifications; among these lexical specifications are features that give a clue as to the kind of elements of domain D this operator may range over. Consider the following examples:

(104)a. the rock who John threw...
b. who did John fill the tub with?
c. who elapsed?

In example (104a), the relative clause is predicated of the rock. But the lexical features of the rock are in conflict with the features that specify the potential range of the relative operator who; hence, the example is odd. Under one reading, example (104b) presupposes that John filled the tub using some person; since people are not generally the sort of material one uses to fill a tub, this reading is again odd and the comitative reading is the only natural one. The oddness of (104b) is most easily captured by assuming that who may only range over the set of humans. The verb, elapse, in (104c) selects nouns that measure time as its subject; given that who is restricted to range over humans, the example is again odd.

The behavior of non-overt operators contrasts markedly from that of overt operators insofar as a non-overt operator may range

over just about anything:

(105)a. We found a person $[_{\bar{S}_i}$ Op$_i$ $[_S$ PRO to hire $\underline{t}_i$]]

b. [an hour $[_{\bar{S}_i}$ Op$_i$ $[_S$ PRO to waste $\underline{t}_i$]]] would be a nice luxury.

c. John gave Mary a rock $[_{\bar{S}_i}$ Op$_i$ $[_S$ PRO to throw $\underline{t}_i$ at Bill

d. We need some water $[_{\bar{S}_i}$ Op$_i$ $[_S$ PRO to fill this tub with $\underline{t}_i$]]

As the above examples illustrate, non-overt operators may range over just about anything: people, units of time, rocks, water, etc. It is a plausible operating assumption that non-overt operators have no inherent range; they must, rather, pick up their range from some other element.[17]

Aoun & Clark assume that non-overt operators must be identified; that is, non-overt operators must pick up their range from some other element. Furthermore, this identification procedure applies at S Structure. Let us take parasitic gap constructions as the core case of non-overt operators. While this strategy may appear unusual, parasitic gaps are accepted by many speakers despite the plausible assumption that not much information about parasitic gaps exist in the primary linguistic

---

17 One could, of course, assume that there are several non-overt operators, each with its own lexical entry and its own inherent range. This approach is completely uninteresting insofar as it merely stipulates the properties of non-overt operators rather than attempting to derive these properties from independent principles.

data available to the language learner. Given the paucity of available data and the striking degree of agreement of judgements that speakers of English exhibit with respect to parasitic gaps, we are more or less forced to adopt the position that properties of parasitic gap constructions follow from principles of core grammar.

We had two interesting properties of parasitic gap constructions (see (96), above). First, a parasitic gap is licensed by a variable that (apparently) does not c-command it. Thus, an example like:

(106) Who_i did John snap [$_{NP}$ photos of $\underline{t}_i$] without warning $\underline{e}_i$?

is well-formed although the licensing variable almost certainly does not c-command out of the NP which contains it. Second, a wh-element that remains in-situ at S Structure is never capable of licensing a parasitic gap:

(107) *why did John snap [$_{NP}$ photos of who(m)_i] without warning $\underline{e}_i$?

Let us assume, as seems rational, that these two properties of parasitic gap constructions are not unrelated; that is, the fact that a parasitic gap must be licensed by a variable which does not c-command it and the fact that wh-elements in situ cannot license parasitic gaps are really the same fact. Notice further that the operator which binds the licensing variable must

c-command the parasitic gap (and, so we would assume, the non-overt operator which binds the parasitic gap):

(108) *[$_{NP}$ the book [that Mary filed $\underline{t}_i$]] was reviewed (by John) without reading $\underline{e}_i$

In (108), the relative operator may not c-command out of the relative clause and the parasitic gap is contained in an adjunct on the matrix S; hence, the relative operator cannot c-command the non-overt operator. We assume, then, that the operator which binds the licensing variable must c-command the non-overt operator, although the licensing variable may not c-command the non-overt operator.

We, therefore, have some very good evidence that the non-overt operator involved in parasitic gap structures has anaphoric properties. First, we have seen that the non-overt operator has no inherent range; rather, it must pick up a range from some other element. Second, we have seen that the non-overt operator must be c-commanded by the element from which it receives its range. This second requirement looks very much like a standard binding relationship, so we will hypothesize that the non-overt operator must be bound.

The assumption that binding is involved in the licensing of non-overt operators immediately derives one of the basic properties of parasitic gap constructions. We have evidence of this from the fact that elements in situ may not license a parasitic gap. By assumption, however, a wh-element in situ will

move to Comp at LF; thus, example (107) will have an LF representation like the following:

(109) $[_{\bar{S}}$ [why who] [did John snap [photos of $\underline{t}_i$]
$_{Comp}$ $_j$ $_i$ $_j$ $_{NP}$
$\underline{t}_j$ without $[_{\bar{S}}$ Op $_i$ [$_S$ PRO warning $\underline{e}_i$]]]]

Since the only wh-element in Comp at S Structure is <u>why</u>, Comp will bear the index of <u>why</u>. As argued above, then, the wh-element that moves to Comp at LF, <u>who</u>, is not capable of binding out of Comp for the purposes of the Binding theory; hence, the non-overt operator in (109) is unbound even at LF. Given the correctness of our assumption that non-overt operators are anaphors, the non-overt operator in (109) will violate condition A of the Binding theory.

But we are operating under the assumption that the Binding theory really represents two systems of binding--the A-system and the $\bar{A}$-system. So now we can ask which version of condition A do non-overt operators fall under. Are non-overt operators A-anaphors or $\bar{A}$-anaphors? Once again, we have already seen the answer to this question: a parasitic gap is not well-formed if the licensing variable c-commands it. This fact follows immediately if we assume that non-overt operators are pure $\bar{A}$-anaphors. Given that non-overt operators are pure $\bar{A}$-anaphors, they may not be locally A-bound. Consider examples like:

(110)a. *$[_{\bar{S}}$ which papers [$\underline{t}_i$ were filed [without $[_{\bar{S}}$ Op $_i$
$_i$ $_S$
[PRO reading $\underline{t}_i$]]]]]
$_S$

b. *[$_{\bar{S}}$ who$_i$ [$_S$ $\underline{t}_i$ left [without [$_{\bar{S}}$ Op$_i$ [$_S$ John's seeing $\underline{t}_i$]]]]]

In each of the examples of (110) there is a variable in the subject position of the matrix S. This variable will c-command the non-overt operator in the Comp of the adjunct $\bar{S}$. Given these conditions, the variable in the subject position of the matrix S will locally A-bind the non-overt operator. If this is correct, then the non-overt operator is not locally $\bar{A}$-bound which violates condition A of the $\bar{A}$-system of binding. We have seen, then, that the two major properties of parasitic gap constructions noted above follow immediately from the assumption that non-overt operators obey condition A of the $\bar{A}$-system of binding.

Our assumption that non-overt operators are $\bar{A}$-anaphors makes an additional prediction, if we assume that they are pure $\bar{A}$-anaphors subject only to condition A of the $\bar{A}$-system of binding. Let us consider, for a moment, the minimal domain in which such an anaphor must be bound:

(111) [$_{\bar{S}}$ who$_i$ [$_S$ John Infl admire $\underline{t}_i$ [without [$_{\bar{S}}$ Op$_i$ [$_S$ PRO ever having met $\underline{t}_i$]]]]]

The non-overt operator is in the Comp of the adjunct clause and, so, may not have the adjunct $\bar{S}$ as its governing category. The question is whether or not the governing category of the non-overt operator is the superordinate $\bar{S}$. But we already have evidence that the governing category for anaphors in Comp is the superordinate $\bar{S}$; consider the contrast in (112):

(112)a. John_i wondered [$_{\bar{S}}$ [which pictures of himself]$_{NP\ i\ j}$ [Sue liked $\underline{t}_j$]]$_S$

b. *John_i said [$_{\bar{S}}$ that [Mary wondered [$_{\bar{S}}$ [which pictures of himself]$_{NP\ i\ j}$ [Sue liked $\underline{t}_j$]]]]$_S$

Example (112a) shows that an anaphor in an embedded Comp may have an antecedent in the superordinate S. Example (112b) shows that the anaphor in Comp may look no higher than the superordinate $\bar{S}$ for its antecedent. Similarly, we have already seen that traces in Comp must be bound in the superordinate $\bar{S}$ (see the discussion of examples (82) through (84), above). We will assume then that the governing category for pure anaphors in Comp is the superordinate clause.

If this is correct, then we have a very strict locality condition on the distribution of non-overt operators. Consider examples like those in (113):

(113)a. *[$_{\bar{S}}$ which famous professor [did Mary persuade $\underline{t}_i$]$_{S\ i}$ [$_{\bar{S}}$ that [John invented the binding theory [$_{\bar{S}}$ without Op_i [PRO consulting $\underline{t}_i$]]]]]]$_S$

b. *[$_{NP}$ the girl [$_{\bar{S}}$ who [I talked to $\underline{t}_i$ [$_{\bar{S}}$ without [PRO knowing [$_{\bar{S}}$ that [John was a very happy man [$_{\bar{S}}$ before Op_i [he met $\underline{t}_i$]]]]]]]]$_S$

In example (113a), the adjunct containing the non-overt operator is itself contained in an embedded clause; by our assumption, the governing category of the non-overt operator must be the embedded

$\bar{S}$. The antecedent for the non-overt operator is in the matrix clause, however. The non-overt operator is free in its governing category and the example is ungrammatical. In (113b), the non-overt operator is contained in an adjunct (headed by before) that is itself contained in an adjunct (headed by without). The $\bar{S}$ headed by without must be the governing category for the non-overt operator, under our operating assumption. The antecedent, however, is outside of the non-overt operator's governing category, which means, of course, that the non-overt operator is free in its governing category. As expected, the example is ungrammatical.

Chomsky (lecture; 1984) accounts for examples like those in (113) by requiring that the chain formed by the non-overt operator and the gap that it is coindexed with must be connected to the chain formed by the licensing variable and its $\bar{A}$-binder (this is, in some ways, reminiscent of Kayne, 1984 and Longobardi, 1985). The two chains would "compose" and this process of "chain composition" would be subject to the modified version of subjacency alluded to in chapter 2. The locality requirement on non-overt operators would, then, follow from subjacency, which would act both as a constraint on rules (in the case of extraction) and as a constraint on representations (in the case of chain composition).

The subjacency account would appear to make predictions which are virtually identical to the binding account developed in Aoun & Clark. The two analyses differ, however, with respect to

languages where $\bar{S}$, but not S, counts as a bounding node for subjacency. As Sportiche (1981) has demonstrated, French is a language where S does not count as a bounding node for subjacency. Thus, in the following sentence the wh-element <u>quel livre</u> ("which book") has been moved in one step to the matrix COMP from its base-generated position; the same point has been demonstrated for Italian sentences parallel to (114) by Rizzi (1982):

(114) $[_{\bar{S}}$ quel livre $[_{i\ S}$ Jean sait $[_{\bar{S}}$ à qui $[_{j\ S}$ PRO offrir $\underline{t}_i$ $\underline{t}_j]]]]$
"Which book does Jean know to whom to offer?"

Note that the wh-quantified, <u>quel livre</u>, must cross two S boundaries in one step to get to Comp in (114). Thus, S does not count as a bounding node in French.

With this in mind, consider the contrast between (115a) and(115b), on the one hand, and (115c) on the other:

(115)a. $[_{\bar{S}}$ quel livre $[_{i\ S}$ Jean a offert $\underline{t}_i$ à Pierre sans $[_{\bar{S}\ i\ S}$ Op $[$ PRO avoir mis $\underline{t}_i$ sur la table$]]]]$
"Which book did Jean offer to Pierre without having put on the table?"

b. $[_{\bar{S}}$ quel livre $[_{i\ S}$ Jean sait $[_{\bar{S}}$ $\underline{t}_i$ que $[$ tu as offert $\underline{t}_i$ à Pierre sans $[_{\bar{S}\ i\ S}$ Op $[$ PRO avoir mis $\underline{t}_i$ sur la table$]]]]]$
"Which book does Jean know you offered to Pierre without having put on the table?"

c. *[$_{\bar{S}}$ quel livre$_i$ [$_S$ Jean a su [$_{\bar{S}}$ qui [$_S$ PRO offrir $\underline{t}_i$
sans [$_{\bar{S}}$ Op$_i$ [$_S$ PRO avoir mis $\underline{t}_i$ sur la table]]]]]]
*"Which book did Jean know to whom to offer without having put on the table"

Example (115a) is a standard parasitic gap structure. The wh-quantified element, <u>quel livre</u>, is in the matrix Comp. Since the adjunct is on the matrix clause, the governing category of the non-overt operator is also the matrix $\bar{S}$. Since the wh-quantified NP in the matrix Comp $\bar{A}$-binds the non-overt operator in its governing category, the sentence is well-formed. In example (115b), the adjunct is on an embedded clause; this embedded clause is also the d-structure locus for the wh-quantified NP. Since the movement of the wh-quantified NP may be successive cyclic, we may assume the existence of a trace in the Comp of the embedded clause. This trace in the intermediate Comp will $\bar{A}$-bind the non-overt operator in the Comp of the adjunct clause. Since the governing category of the non-overt operator is, by assumption, the embedded complement clause, the non-overt operator is $\bar{A}$-bound in its governing category. Notice that the variable associated with the intermediate trace in Comp is also subjacent to the non-overt operator. Example (115b) is compatible both with the chain composition analysis of Chomsky and the binding account of Aoun & Clark.

Example (115c), however, is compatible only with the binding account of the distribution of non-overt operators. Example (115c) differs from example (115b) only in the presence of a wh-

element in the Comp of the embedded clause; this wh-element indicates that an intermediate trace in Comp left by movement of quel livre to the matrix Comp will not c-command out of Comp. Crucially, the licensing variable is subjacent to the non-overt operator, just as in example (115b); because of this fact, the chain composition analysis would predict that example (115c) is grammatical. Since an intermediate trace in Comp cannot bind anything due to the presence of qui, the non-overt operator is $\bar{A}$-free in its governing category. The binding analysis of non-overt operators will correctly predict that example (115c) is ungrammatical. Suppose, however, that the non-overt operator is coindexed with qui; since the variable bound by qui is in the subject position of the embedded clause, the non-overt operator will be locally A-bound by this variable. Again, this will result in a violation of the binding theory since the non-overt operator is an $\bar{A}$-anaphor. Thus, there is no way to save example (115c) under the binding account.

We have seen some evidence that would lead us to select the binding theory account of non-overt anaphors over a chain composition account. Nevertheless, there are a number of apparent counterexamples:

(116) a. John_i is easy [$_{\bar{S}}$ Op_i [$_S$ PRO to please $\underline{t}_i$]]

b. the book$_i$ was bought [$_{\bar{S}}$ Op_i [$_S$ PRO to read $\underline{t}_i$ to the children]]

c. a shovel$_i$ is [$_{\bar{S}}$ Op_i [$_S$ PRO to dig with $\underline{t}_i$]]

Example (116a) is a tough movement construction, example (116b) is a purposive and example (116c) has an infinitival clause in the predicate position of the matrix clause. In each of the examples, the subject of the matrix c-commands a non-overt operator. Under the assumption that the subject in each of the above sentences is coindexed with the non-overt operator, we are apparently forced to admit that the non-overt operator in each of the above examples is A-bound, a state of affairs which, as argued above, is not allowed under the binding analysis of non-overt operators. One could, of course, argue that the structures in (116) all involve predication. Following Chomsky (1982), we could stipulate that the subject and the predicate in each of the above examples do not share identical indices until some late level of representation, for example, LF´. But this will not help us out of our dilemma; although we may be able to demonstrate, given this basis, that the subject in the above examples does not A-bind the non-overt operator, we have not provided an element that will bind the non-overt operator. Supposing that non-overt operators are pure $\bar{A}$-anaphors, we must be able to supply an $\bar{A}$-binder for each of the non-overt operators in (116).

We have seen that non-overt operators have no implicit range; the only way a non-overt operator can assign a range to the variable which it binds is if it can pick up a range from some other element. In the case of parasitic gaps, the non-overt operator picks up a range from some operator which locally binds

it. Let us suppose that the structures in (116) involve a form of predication. That is, in (116a), the $\bar{S}$ complement to easy is predicated of the matrix subject, John; in (116b), the purpose clause, to read to the children, is predicated of object, the book; and in (116c), to dig with is predicated of the subject, a shovel.

Following Williams (1980), let us assume that predication involves the copying of the index of the subject onto the predicate that applies to this subject. Hornstein (1984) has supplied some evidence that this view of predication is correct in that the predicate obeys the i-within-i condition (see the discussion above). Consider the following:

(117) Bill is [his cook]
i NP {j,*i}

In (117), the genitive pronoun, his, may only be taken as disjoint from the matrix subject Bill. There is, however, no general prohibition against genitive pronouns being coreferential with a c-commanding NP; thus, the following is perfectly well-formed:

(118) Bill loves [his cook]
i NP i

The difference between the two examples is that his cook is directly predicated of Bill in (117), but not (118). If we assume that predication involves coindexing the predicate with its subject, then the facts follow trivially. Applying predication to (117) will yield:

(119) Bill is [his cook]
i NP {j,*i} i

If the genitive pronoun has the same index as the subject, the structure will violate the i-within-i condition, since his will have the same index as the phrase which contains it. Since his cook in (118) is not predicated of the subject, no i-within-i violation will arise; in fact, since his cook in (118) is a bona fide R-expression (as opposed to a predicate), it must be disjoint from Bill, by virtue of condition C of the A-binding system. Hence, the phrase containing his will never come to be coindexed with the subject, leaving the genitive pronoun free to be (optionally) coindexed with Bill.

There is good reason for saying that a predicate and its subject are coindexed. Formally, this coindexation appears to be identical to binding; we can, however, question whether or not we want to interpret the coindexation of a predicate and its subject as an instance of binding. The position I have taken in this chapter is that binding is, strictly speaking, the syntactic reflection of referential properties of items; although we cannot identify the indices on NPs with the elements they refer to, the indices are, nonetheless, interpreted (see the discussion above). A simple version of this approach would treat each element with an index as though it refers to some element of domain D. Thus, a predicate would refer to some element of D; in particular, it would refer to the same element as the subject refers to. Geach (1960) argues that this view of predicates as referring expressions is much too naive to be practicable. For example, he

views it as nonsense to ask what cat cat refers to in Jemima is a cat or which dog dog refers to in Jemima isn't a dog. While his arguments are too involved to develop here, I will view his point as taken. I will therefore assume that the relationship between a predicate and its subject is never one of binding, since, in any case, it makes little sense to say that the predicate picks up its reference from the subject.

Let us, further, take as given that the features of a node are shared by all the projections of that node. In particular, an index assigned to a projection will percolate down to the head of that projection. For example, the representations in (116) are structures of predication, so the index on the subject will be copied onto the predicate, as in (120):

(120)a. John_i is easy $[_{\bar{S}}\ \text{Op}_j\ [_S\ \text{PRO to please}\ \underline{t}_j\ ]]_i$

b. the book_i was bought $[_{\bar{S}}\ \text{Op}_j\ [_S\ \text{PRO to read}\ \underline{t}_j$ to the children$]]_i$

c. a shovel_i is $[_{\bar{S}}\ \text{Op}_j\ [_S\ \text{PRO to dig with}\ \underline{t}_j\ ]]_i$

The index assigned to each of the $\bar{S}$'s in (120) under predication will percolate down to Comp. Hence, the non-overt operator (and the trace it binds) will be reindexed by this rule as in (121):

(121)a. John_i is easy $[_{\bar{S}}\ \text{Op}_i\ [_S\ \text{PRO to please}\ \underline{t}_i\ ]]_i$

b. the book_i was bought $[_{\bar{S}}\ \text{Op}_i\ [_S\ \text{PRO to read}\ \underline{t}_i$ to the children$]]_i$

c. a shovel_i is $[_{\bar{S}}\ \text{Op}_i\ [_S\ \text{PRO to dig with}\ \underline{t}_i\ ]]_i$

Since the non-overt operator occupies the head of $\bar{S}$, no i-within-i violation is generated (heads may always be coindexed with their projections) and the range of the non-overt operator is delimited by that of the subject of the predication. Notice that the "range" of the subject is sometimes rather trivial, as in (121a) where the proper noun, John, refers to a unique individual.

Having established that the relationship between the non-overt operator and the element that delimits its range is not one of binding (strictly speaking), we must still explicate what constraint holds of non-overt operators in cases of predication. Let us reconsider the approach to the assignment of indices found in Chomsky (1982). Chomsky observes that indices may only be assigned to A-positions at S Structure. There is a very good reason for this constraint. Suppose a wh-element were base generated in Comp and could freely receive an index. We could freely generate a D Structure representation like:

(122) $[_{\bar{S}}$ who $[_{S}$ John believe $[_{NP}$ the claim $[_{\bar{S}}$ that $[_{S}$ Mary saw $[_{NP}$ $\underline{e}]]]]]]$

By free assignment of indices, the representation in (122) could be mapped to an S Structure like:

(123) $[_{\bar{S}}$ who$_i$ $[_{S}$ John believe $[_{NP}$ the claim $[_{\bar{S}}$ that $[_{S}$ Mary saw $[_{NP}$ $\underline{e}]_i]]]]]$

The derivation of example (123) would not violate subjacency since subjacency is a constraint on "move alpha". We would then

be able to freely generate strings which violate subjacency merely by base generating a wh-element in a Comp, an empty category in some (non-subjacent) argument position and relating the two by means of free assignment of indices. Implicit in this analysis is that elements must bear an index, either through inherent reference (and occupying an A-position) or by being bound by an element with an index. We can state the following as a corollary to this principle:

(124) Non-overt operators must be identified.

"Identification" is intended to mean that the element bears an index. We may assume that this identification requirement holds of S Structure representations.[18]

The solution to this problem is to prevent elements in non-argument positions from freely receiving indices. The only way that an element in a non-argument position may receive an index is by virtue of being associated with an A-position under "move

---

[18] Longobardi (1985) points out the existence in Italian of sentences where the adverbial clause has "crossed over" the antecedent:

(i) ?Senza conoscere e prima bene, non so propio quale altra ragazza Gianni sarebbe disposto a sposare e.

'Without knowing well beforehand, I really don't know which other girl Gianni would be ready to marry.'

It is likely that these sentences involve reconstruction such that the adverbial clause is put back in situ at some level of representation. Nevertheless, they may indicate that the identification requirement does not hold of S Structure representations, but of representations of some later level, perhaps LF.

alpha". At S Structure, the variable will receive an index and this index will be copied onto the operator by convention. Herein lies a framework by which we may rethink the problem of the distribution of non-overt operators.

Let us suppose that R-expressions and pronominals freely receive an index at S Structure since they may refer independently. Anaphors receive an index by virtue of being related to an element which may refer independently. The output of indexation is, of course, subject to a generalized binding theory. Thus a variable, which is an $\bar{A}$-anaphor, receives an index by virtue of occupying an A-position and by virtue of being bound by an element with an inherent range. But non-overt operators have no inherent range. The only way a non-overt operator and the variable which it binds can receive an index is by being associated with an element that independently receives an index. We have seen that there are two ways in which a non-overt operator may be associated with an element that bears an index. First, the non-overt operator may enter into a binding relation with the other element. In this case, the binder must occupy an $\bar{A}$-position and the relationship between the non-overt operator and its binder is subject to condition A of the $\bar{A}$-binding system. Second, the non-overt operator may occur in the head position of a predicate in a predication structure. In this case, the index on the subject will be copied onto the predicate and percolate down to the non-overt operator. Since the

relationship between a predicate and its subject is local, this second manner in which a non-overt operator may be identified is quite restrictive.

We have seen that the identification requirement on non-overt operators places a strong requirement on the distribution of such operators. Furthermore, this treatment provides a unified analysis of parasitic gap constructions, purposive clauses, tough movement constructions and infinitival relatives. We are now in a position to apply the results of this section in the analysis of control, with particular attention being paid to the so-called PRO gate.

Chapter 5
An Approach to Control

In this chapter, we will compare the analysis of non-overt operators in Aoun & Clark (1985) with the theory of control. It will be shown that PRO can be treated as a non-overt operator. In brief, I will demonstrate that obligatory control can be treated as the identification, at S Structure, of a non-overt operator while so-called arbitrary control is the identification, at LF, of a non-overt operator. In section 5.1 we will bring up a number of empirical issues that a theory of control must answer to. I will rely on a careful examination of the theory of control found in Manzini (1983). We will also look at the problem of the PRO gate. In section 5.2 we will replace PRO with non-overt operators, based on the results of Epstein (1984) and Lebeaux (1984); in particular, we will assume that PRO simply does not exist. In addition, we will consider cases of "double

control" (cf. Chomsky, 1980) and some further S Structure conditions on the distribution of non-overt operators. In section 5.3 we will consider "arbitrary" control, in particular we will focus on control by implicit arguments and by various unselective operators and adverbs of quantification. The analysis in this section will support the hypothesis that LF is an extremely rich level of representation and the hypothesis that LF is interpreted, in the sense of Keenan & Faltz (1985).

5.1 Control and the PRO Gate

Control theory is involved with the referential possibilities of a special designated element, PRO. In Chomsky (1981), PRO is taken as having the feature specification [+pronominal, +anaphor]. As pointed out, above, this places PRO under conditions A and B of the A-binding system. Since no element could possibly satisfy both conditions A and B of the binding theory--to do so would mean that such an element would be both free and bound in a local domain--PRO may not having a governing category. This is taken as implying that PRO may not be governed.[1] The only A-position which is not governed is the subject of a non-finite clause. Thus, PRO is restricted to occurring in the subject position of infinitivals and gerunds:

---

1 Strictly speaking, an element could escape the governing category requirement by not having an accessible SUBJECT in any domain that contains it. To my knowledge, this possibility has been ignored in the literature. Notice that variables in object position lack a governing category, under the assumptions of the type of generalized binding theory discussed above. In a sense, I will exploit this fact below.

(1)a. John persuaded Bill [PRO to study counterpoint]
b. [PRO to study counterpoint] would be a hardship beyond endurance.
c. [PRO studying counterpoint] is probably worthwhile.

The goal of control theory is to explain which NP PRO may corefer with, if any. Thus, in (1a) PRO is taken as coreferring with Bill but not John while in (1b) and (1c), PRO is taken as being arbitrary in reference.

Manzini (1983) outlines a number of the important properties of PRO that a control theory must account for. First, when PRO occurs in a sentential complement, PRO must be coreferential with the subject of with the object of the superordinate clause:

(2)a. John asked Bill [PRO to paint himself]
i j j j
b. *John asked Bill [PRO to paint himself]
i j i i
c. *John asked Bill [PRO to paint oneself]
i j arb
(3)a. John promised Bill [PRO to paint himself]
i j i i
b. *John promised Bill [PRO to paint himself]
i j j j
c. *John promised Bill [PRO to paint oneself]
i j arb

The special index arb is taken as designating an arbitrary PRO. In (2a), the PRO obligatorily corefers with the object, Bill, and may not be taken either as coreferring with the subject, John, or as free in reference. In (3a), the PRO must be taken as coreferential with the subject, John, and any construal under which the PRO corefers with the object or is free is ungrammatical. Facts like those in (2) and (3) have led to the identification of the feature [+anaphor] with PRO since,

intuitively, the PRO in these constructions refers only by virtue of being related with some other NP that either refers independently or is bound.

The exception to the generalization that PRO in a sentential complement is anaphoric appears to be certain verbs of saying, as in (4):

(4) John {whispered, shouted, said, etc.} [PRO to behave oneself]

The verbs in these cases have some rather special properties. For example, many, but not all, of these verbs are non-bridge verbs:

(5)a. *who$_i$ did John whisper [$_{\bar{S}}$ $\underline{t}_i$ [$_S$ Mary saw $\underline{t}_i$]]

b. *what kind of poison$_i$ did Mary shout [$_{\bar{S}}$ $\underline{t}_i$ [$_S$ Bill drank $\underline{t}_i$]]

In a sense, these verb, furthermore, do not denote a relation between entities and propositions (compare verbs of belief) but rather, they set up a relation between entities and utterances. Thus, one can whisper, say, shout, etc. an order:

(6)a. John shouted to leave.
b. Bill whispered to stand up straight.
c. The general said to be quiet.

This is despite the fact that orders do not bear a truth-value. Since it is plausible that verbs of saying fall into a special class, it should come as no surprise that they have special control properties. We will not be overly concerned with the control properties of this class of structures.

The contrasts, above, are generally interpreted as

demonstrating that control of PRO involves some sort of lexical specification on the part of the verb that subcategorizes for the non-finite complement, although it seems reasonable to invoke some form of the Minimal Distance Principle of Rosenbaum (1967). Thus, minimal control contrasts like the following:

(7)a. John_{i} promised Bill_{j} [PRO$_{\{i,*j\}}$ to leave]

b. John_{i} persuaded Bill_{j} [PRO$_{\{j,*i\}}$ to leave]

may be accounted for by appealing to the differences in the lexical structure of <u>promise</u> and <u>persuade</u>; <u>promise</u> designates its external argument as the controller of PRO while <u>persuade</u> designates its internal argument as the controller. Structurally, however, (7a) and (7b) are identical. Thus, obligatory control is not based on purely structural factors, unlike determining the range of possible antecedents for a particular anaphor. Elements of lexical semantics must be taken into account in determining the controller of PRO.

Notice further that the PRO is outside of the government domain of the superordinate verb which determines its controller. This is a further unusual property of obligatory control since a the verb must determine a property of PRO although the verb may never govern it. Under current assumptions, this latter point is taken as a virtual theorem of the binding theory since a PRO may never be governed.

The relationship of obligatory control which is by hypothesis specified by a verb is also sensitive to properties of

Comp. In particular, a [+wh] Comp blocks obligatory control:

(8)a. John$_i$ asked [PRO$_i$ to shave himself$_i$]

b. *John$_i$ asked [PRO$_{arb}$ to shave oneself]

c. John$_i$ asked [$_{\bar{S}}$ how [$_S$ PRO$_i$ to shave himself$_i$]]

d. John$_i$ asked [$_{\bar{S}}$ how [$_S$ PRO$_{arb}$ to shave oneself]]

The verb, ask, when used intransitively, specifies the subject as the controller of PRO, as shown in (8a) and (8b), where the presence of the impersonal reflexive element, oneself, results in ungrammaticality. When a wh-element occurs in Comp, as in (8c) and (8d), control is not obligatory, the PRO subject may be indexed with arb, and the impersonal reflexive is permitted.

A further property of control that must be accounted for by a control theory is that control is generally optional when the clause containing a PRO subject occurs as a sentential subject:

(9)a. [PRO$_i$ to behave oneself in public] is desirable

b. [PRO$_i$ to kill himself$_i$] would bother Bill.

c. Mary$_i$ thinks that [PRO$_i$ to make a film about herself$_i$] would be profitable

It should be noted that judgements regarding control into a sentential subject seem to be fairly elastic. For many of the speakers I consulted, control appears to be obligatory in (9b) and strongly preferred in (9c). I will accept these judgements, although they run counter to judgements occasionally found in the literature.

To summarize, we have found three features of PRO that a control theory must account:

(10)a. A PRO subject of a subcategorized sentential complement must be controlled by an element in the superordinate clause.
b. A head may specify the controller of PRO.
c. A PRO subject of a sentential subject need not be controlled; its controller, furthermore, need not be in the immediately superordinate clause (cf. (9c).

Manzini (1983) accounts for the above generalizations by treating controlled PRO as an anaphor. Since this anaphor occurs in a ungoverned position, it will not have a governing category and, prima facie, it has no minimal domain in which it must be bound. To resolve this, Manzini defines "domain governing category":

(11) A is a domain governing category for B iff
a. A is a governing category for the c-domain of B, and
b. A contains a SUBJECT accessible to B.

(12) A is the c-domain of B iff A is the minimal maximal category dominating B.

The definitions in (11) and (12) have a very uniform effect. The c-domain of PRO will be the S that immediately dominates the PRO. The minimal category which contains the c-domain of PRO and a SUBJECT accessible to PRO will, in the case of subcategorized sentential complements, be the immediately superordinate clause. If the c-domain of PRO occurs in subject position (i.e., a non-finite sentential subject) then the superordinate clause will not contain a SUBJECT accessible to PRO. This is because the sentential subject and the Infl of the superordinate clause are

obligatorily coindexed; coindexing the PRO with the Infl will lead to a violation of the i-within-i condition. Hence, the domain governing category of such a PRO will have to be higher than the superordinate clause (if there is anything higher).

To make this more concrete, consider an example like:

(13) $\text{John}_i\ \text{Infl}_i$ tried $[_{\bar{S}}\ [_S$ PRO to shave himself]]

The c-domain of PRO is the S that contains it--the embedded clause. We will assume that the Comp of the embedded clause is not present since it is not independently required as, for example, a landing site. If this is so, then the embedded S is governed by the matrix verb, <u>try</u> since the Comp dominating the embedded clause does not branch. The governing category for the c-domain of PRO will be the matrix S since the matrix S is the minimal maximal category containing a governor for the embedded clause and a SUBJECT accessible to the embedded clause; in this case, the matrix Infl is the accessible SUBJECT for the embedded clause. The matrix Infl is also an accessible SUBJECT for the PRO inside the embedded clause. Hence, the matrix clause is the domain governing category for PRO.

We can contrast example (13) with the following:

(14) John_i Infl wondered $[_{\bar{S}}$ how $[_S$ PRO to solve the problem]]

The c-domain of PRO is, again, the embedded clause. But this time, Manzini assumes that the embedded clause is not governed,

due to the presence of how in the embedded Comp.[2] If the embedded S is not governed, then it cannot have a governing category. But if the c-domain of PRO has no governing category, then PRO lacks a domain governing category. A similar result may be derived for sentential subjects, since the embedded clause is, under Manzini's assumptions, ungoverned due to the presence of Comp. Hence, a PRO in a sentential subject or in an $\bar{S}$ with a phonologically overt Comp will lack a domain governing category.

In general, then, if PRO has a domain governing category, its domain governing category will be the superordinate S; otherwise, PRO will simply lack a domain governing category. Recall that one of the features of PRO is that a PRO in the subject position of a subcategorized complement is controlled by an element in the superordinate clause. We can restate this as a PRO which has a domain governing category must be bound in that domain governing category. Manzini states this as:

(15) An anaphor without a governing category is bound in its domain governing category.

---

2 The assumption that how in (14) does not govern the embedded S seems to be at odds with the framework developed in Lasnik & Saito (1984). They argue that how must be able to govern out of Comp since its trace, which occupies an adjunct position, is only properly governed if it is antecedent governed. If how is capable of governing its own trace, why is it incapable of governing the S? The obvious answer is, of course, that how is not coindexed with S and proper government relies on some form of coindexation. I argued in the previous chapter, on the basis of proper government of the subject of small clause and Exceptional Case Marking structures, that coindexation was not a necessary condition for proper government. I will leave the problem unresolved here, for reasons which will become apparent.

In the case of obligatorily controlled PRO, we have seen that it is always bound in its domain governing category, the superordinate clause. In case the PRO lacks a domain governing category--the cases of clauses with an element in Comp, subordinate nonfinite clauses in subject position, and so on--the constraint on PRO being bound in its domain governing category is trivially satisfied since PRO lacks a domain governing category to be bound in.

Manzini's account has a number of rather interesting ramifications. We will primarily be interested in Comp-to-Comp movement, which, in the sense in which this term is normally taken, is not possible. Consider the following set of sentences:

(16)a. John_i decided [PRO_i to shave himself_i]

b. *John_i decided [PRO_{arb} to shave oneself]

c. How_i did John_i decide [$\underline{t}_i$ [PRO_i to shave himself_i $\underline{t}_i$.]]

d. *how_i did John_i decide [$\underline{t}_i$ [PRO_{arb} to shave himself_i $\underline{t}_i$]]

The contrast between (16a) and (16b) show that <u>decide</u> is an obligatory control verb since it is impossible to have an impersonal reflexive in the embedded clause. It must be the case, then, that the c-domain of PRO (the embedded S) has the matrix S as its governing category. This implies that the matrix verb, <u>decide</u>, governs the embedded S. It was argued above that the matrix verb could do this only if $\bar{S}$ does not branch. Hence,

there can be no Comp. Now consider (16c). If the wh-element moved to the intermediate Comp on its way to its ultimate landing site, then the embedded $\overline{S}$ must branch; but then we would not expect obligatory control. Nevertheless, (16d) shows that control is still obligatory. Thus, for this version of control theory to work, Comp-to-Comp movement must be impossible. Intermediate traces must, rather, be adjoined to S (or, perhaps $\overline{S}$, depending on your theory of adjunction and government) so that the $\overline{S}$ will not branch in obligatory control structures. In order to make this serviceable, we need a constraint like the following:

(17) An element in Comp may not undergo "move alpha"; an element adjoined to S in the syntax must undergo "move alpha".

or some other constraint that will guarantee that intermediate traces are never in Comp.[3]

Thus, under Manzini's theory of control, we must rework some commonly held assumptions about the nature of successive cyclic movement and the role of Comp in the theory of grammar. Furthermore, we must add the notion of "domain governing category" to the binding theory. This definition ensures that some superordinate category will be the governing category for PRO. Notice that, in the previous section, we saw a similar

---

3 One could, I suppose, eliminate Comp altogether, and simply filter out multiple adjunctions to S at S Structure somehow. This proposal would have a number of consequences (and problems) that would take us far afield.

state of affairs with regard to non-overt operators. We will return to this parallelism in greater detail below (section 5.2).

Let us turn now to the question of the PRO gate which we discussed briefly, above. Earlier, we introduced the convention that a pronoun could be interpreted as a bound variable only if it is locally A-bound by a variable (or locally A-bound by an element that is ultimately bound by a variable). Crucially, a pronoun could not be locally $\bar{A}$-bound. This convention derived the weak crossover effects as illustrated in:

(18) $[\text{his}_{\{i,*j\}}\ \text{mother}]$ loves every man_j

The quantified NP, <u>every man</u>, is adjoined to the matrix S in the LF representation of (18):

(19) $[_S\ [\text{every man}]_j\ [_S\ [\text{his}_{\{i,*j\}}\ \text{mother}]\ \text{loves}\ \underline{t}_j\ ]]$

If the genitive pronoun, <u>his</u>, bears the same index as the quantified NP, it will be locally $\bar{A}$-bound at LF, in violation of our interpretive convention. Hence, the genitive pronoun must be interpreted as disjoint from the quantified NP.

If we invert the surface order of the two NPs in (18), then the genitive pronoun may be bound by the quantified NP:

(20) every man_j loves $[\text{his}_{\{i,j\}}\ \text{mother}]$

This is so because in the LF representation of (20), the genitive pronoun will be locally A-bound by the variable left by QR, rather than being locally $\bar{A}$-bound by the quantified NP itself:

(21) $[_S\ [\text{every man}]_j\ [_S\ \underline{t}_j\ \text{loves}\ [\text{his}_{\{i,j\}}\ \text{mother}]]]$

Since the genitive pronoun is (optionally) locally A-bound by a variable it may be interpreted as a bound pronoun.

Consider a sentences like:

(22)a. [PRO$_i$ kissing [his$_i$ mother]] bothers every boy$_i$

b. [$_S$ [every boy]$_i$ [$_S$ [PRO$_i$ kissing [his$_i$ mother]] bothers $\underline{t}_i$]]

c. [$_{\bar{S}}$ who$_i$ [$_S$ does [PRO$_i$ kissing [his$_i$ mother]] bother $\underline{t}_i$]]

In both of the above sentences, the PRO may receive a "bound pronoun" interpretation. The examples in (22) are, in some sense, troubling. In a theory where PRO bears the features [+pronominal, +anaphor], its pronominal character should prevent its interpretation as a bound variable unless it is locally A-bound by a variable. In (22a), and its associated LF, (22b), the PRO is never locally A-bound by a variable; in the LF representation, in fact, the PRO is locally $\bar{A}$-bound. In (22c), the PRO is locally $\bar{A}$-bound at S Structure. It was facts like those in (22) that led Higginbotham (1980) to formulate the PRO gate which stipulates that pronominals, except for PRO, are interpreted as bound variables only if locally A-bound by a variable; PRO may be locally $\bar{A}$-bound despite its specification as a pronominal.

The examples in (22) have another disturbing implication. Chomsky (1982) proposes that the features of empty categories may be determined by their syntactic environment and, therefore,

these is no reason to suppose that empty categories have any inherent features. He proposed a sort of algorithm, the functional determination of empty categories, which could assign features to empty categories on the basis of purely syntactic considerations. Thus, an empty category which occupies an A-position and is locally $\overline{A}$-bound must be a variable and is assigned the features [-pronominal, -anaphor]; otherwise, the empty category is assigned the feature [+anaphor]. If the empty category is free or locally A-bound by an antecedent which bears an independent thematic role, it is assigned the feature [+pronominal], thus making it [+pronominal, +anaphor] or, in other words, PRO. The other possibility is that the empty category is bound by an antecedent in a $\overline{\theta}$-position. In this case, the empty category is assigned [-pronominal] making it a pure anaphor, or, in other words, NP-trace. Notice that this algorithm lacks a way of assigning features to an empty pronominal, pro. It is not difficult to find ways to incorporate pro into the above algorithm by, for example, stating that an empty category that is identified in the appropriate way is a pure pronominal.

This approach, despite whatever shortcomings it may have, has a certain appeal. In a way, the functional determination of empty categories represents a codification of a very elegent theory of empty categories. There is a rather large hole in this approach to empty categories; specifically, the PRO gate runs

counter to the functional determination. Let us look more closely at an example:

(23) $[_{\bar{S}}$ who$_i$ $[_S$ does $[_{PRO_i}$ kissing $[_{his_i}$ mother$]]$ bother $\underline{t}_i]]$

The empty category in the subject position of gerund occupies an A-position and, furthermore, it is locally $\bar{A}$-bound by <u>who</u>; hence, the functional determination would have us assign this empty category the features [-pronominal, -anaphor]. In other words, the functional determination of empty categories systematically misidentifies PRO as a variable in certain structural environments. It should be noted that PRO is identified as a variable so far as the semantic interpretation is concerned, as we will see, however, it is not obvious that the syntax can accept this identification.

From a sufficiently abstract viewpoint, this misidentification of PRO may not appear to be such a terrible tragedy. Consider, however, a theory in which variables and traces are subject to some form of the ECP. The PRO in (23) will violate whatever version of the ECP you care to select simply because the PRO occupies an ungoverned position and every version of the ECP presuppose some form of government of the empty category by some element.[4]

---

[4] If the ECP is completely replaced by some version of a generalized binding theory, then the PRO in (23) may be taken as a variable. I hope to show that even a proponent of this sort of theory would not want to treat the PRO as a variable, all other

There is a further problem with treating the PRO in (23) as a variable. Specifically, it is at least highly marked, if not impossible, for a single operator to locally bind two variables (cf. Koopman & Sportiche, 1983). Until recently, a proposed case of an operator locally binding two variables was in parasitic gap constructions. As we saw above, however, the idea that a single operator bound the two variables in parasitic gap structures was replaced by the proposal that the parasitic gap is locally $\bar{A}$-bound by a non-overt operator. If this is correct, then constructions like (23) represent the only case where a single operator locally $\bar{A}$-binds two variables. Furthermore, if we admit the possibility of a single operator locally binding two variables, then we must find an independent means of filtering out ungrammatical examples like:

(24) *who$_i$ did John describe $\underline{t}_i$ without any pictures of $\underline{t}_i$ being on file

The defunct parasitic gap analysis of Chomsky (1982) is now back since an empty category may be generated inside the subject position of the adjunct clause only to be reanalyzed as a variable at S Structure. Recall that we cannot use a "connectedness" style ECP to rule out (24) without also ruling out (23).

It would appear, then, that we must assume that some, if not

---

things being held equal.

all, empty categories have inherent features. This assumption will prevent the problems we have seen with locally $\bar{A}$-bound PRO since PRO will have inherent features which it cannot change in the course of a derivation. While this approach will solve the technical problem of PRO being converted to a variable, it leaves unanswered a rather interesting problem. Let us first put the question in a general form, and then add some theoretical considerations that make the question more germane. What is the relationship between a PRO in an A-position and an operator in an $\bar{A}$-position? Why is it that a PRO may be treated as a variable with respect to binding and interpretation, but as a PRO with respect to the conditions which license other empty categories?

Koopman & Sportiche (1983) have proposed that, in general, there is a biuniqueness relationship between $\bar{A}$-binders and bound elements in A-positions. They propose the Bijection Principle which requires that operators locally bind one and only one variable. The most famous counterexample to this proposal, parasitic gaps, is a red herring since in these constructions the overt wh-element locally binds only one variable, the true gap; the other element bound by the overt wh-element is a non-overt operator. Non-overt operators are distinct from a true variables since they occupy $\bar{A}$-positions while a variable, by definition, occupies an A-position.

We have already seen that pronouns may not be locally $\bar{A}$-bound from the weak crossover effects discussed above. R-expressions, like names, must be both A- and $\bar{A}$-free. This leaves

anaphors as elements that may, in some sense, be locally $\bar{A}$-bound. It would appear to be a theorem of the binding theory that A-anaphors could not be locally $\bar{A}$-bound from condition A of the A-binding system. Even so-called "psych verbs" do not allow this construction. Consider the sentences in (25):

(25)a. [stories about himself$_i$] annoy John$_i$.

b. [pictures of himself$_i$] please Bill$_i$.

c. [a reflection of himself$_i$ in the mirror$_i$] would startle Renfield$_i$.

In the above sentences, an anaphor inside a subject is coindexed with, but apparently not c-commanded by, an NP in object position. Most speakers seem to dislike sentences like the following:

(26)a. *?who$_i$ did [stories about himself$_i$] annoy $\underline{t}_i$

b. *?which politician$_i$ do [pictures of himself$_i$] please $\underline{t}_i$

c. *?which psychotic$_i$ did [a reflection of himself$_i$ in the mirror] startle $\underline{t}_i$

A plausible account of these facts is that an overt anaphor is locally $\bar{A}$-bound at S Structure. This fact seems to follow from the Bijection Principle since an operator in each of the examples in (26) locally $\bar{A}$-binds two argument positions.

As things stand, however, locally $\bar{A}$-bound PRO (as in the examples in (22)) is a true counterexample to the Bijection

Principle, since an operator must at some level locally $\bar{A}$-bind two argument positions. As things stand, we must simply modify the Bijection Principle with the stipulation that an operator may locally $\bar{A}$-bind PRO in addition to some other argument position. Such a stipulation would, of course, represent a genuine explanatory lacuna in the theory.

Notice, however, that such a theory contains a number of quite unexpected "coincidences". First, given that Manzini's theory of obligatory control is essentially correct, the domain governing category of an obligatorily controlled PRO is the clause that is superordinate to the one containing the PRO. The analysis of non-overt operators given in Aoun & Clark (1985) (see above) requires that non-overt operators must be "identified" (either bound or receive an index under predication) in the superordinate clause. Thus, PRO and non-overt operators seem to have the same properties with respect to the local domain in which they must be bound. Second, PRO may be locally $\bar{A}$-bound in violation of the Bijection Principle; we have seen in the analysis of non-overt operators that they too may be locally $\bar{A}$-bound by an operator that also locally binds an element in an argument position.

One way to derive these parallelisms between PRO and non-overt operators is to identify the two. Such an approach would collapse control phenomena with phenomena that currently fall under the analysis of parasitic gaps; in particular, control, parasitic gaps, purposive clauses, infinitival relatives and

tough movement would be treated as fundamentally the same sort of phenomena. In the next section, we will begin to explore this hypothesis.

5.2 Non-overt operators and control

The idea that the control of PRO has some connection with $\bar{A}$-binding was, to my knowledge, first discussed in Epstein (1984) and Lebeaux (1984), both of whom discuss the similarity between the interpretation of PRO_{arb} and a universally quantified argument. Thus, Epstein points out that a sentence like:

(27) It is fun $[_{\bar{S}}$ $[_{S}$ PRO_{arb} to play baseball]]

seems to have an interpretation similar to:

(28) [for all x [it is fun [x to play baseball]]]

Epstein argues that the arbitrary PRO, itself, could not be treated as an operator since that would allow a reading like the following:

(29) it is fun [for all x [x to play baseball]]

where fun has scope over the the universal quantifier. He argues, instead, that there is a non-overt argument in (27) which acts as the controller of the "arbitrary" PRO. Notice that fun may take an overt benefactive which may be taken as an obligatory controller for the PRO:

(30) It is fun for $John_i$ $[_{\bar{S}}$ $[_{S}$ PRO_i to play baseball]]

He identifies this non-overt argument with pro; thus, (27) has a structure parallel to:

(31) $[_{\bar{S}}\ [_{S}$ it is fun $\underline{pro}_{i}$ $[_{\bar{S}}$ PRO$_{i}$ to play baseball]]]]

The pro in (31) is subject to quantifier raising and interpretation as a universal quantifier, thus yielding an LF where the universal quantifier has scope over the matrix predicate, fun, as desired:

(32) $[_{\bar{S}}\ [_{S}$ pro$_{i}$ $[_{S}$ it is fun $\underline{t}_{i}$ $[_{\bar{S}}\ [_{S}$ PRO$_{i}$ to play baseball]]]]]

Similarly, Lebeaux proposes that a non-overt operator that will $\bar{A}$-bind the arbitrary PRO is inserted at LF. He points out that it is not always plausible to induce an "implicit" argument (as in the above example):[5]

(33)a. John likes stories about PRO bettering oneself.
b. PRO to know him is PRO to love him.

The interpretation of the above examples does not seem to admit another non-overt argument. Lebeaux argues, then, that a non-overt operator must be inserted into the LF representations of the above sentences rather than originating in a phonetically null argument position. Nevertheless, both Epstein and Lebeaux agree that arbitrary PRO behaves, ultimately, as though it were bound by some operator, possibly a universal quantifier, at LF.

The proposal to be developed in this section is somewhat

---

5 "Implicit" is, perhaps, not quite the right word since Epstein requires the "implicit" argument to be syntactically realized as a pro. Others, for example Keyser & Roeper (1984), place no such requirement on implicit arguments which they take as present in a head's argument structure, but not necessarily syntactically realized. See, especially, the next section.

different. In general, we will take PRO, whether it is obligatorily controlled or arbitrary, to be a special case of a non-overt operator which undergoes movement to a non-argument position. A number of questions are immediately apparent. To begin with, does the movement of this putative non-overt operator take place in the mapping from D Structure to S Structure or in the mapping from S Structure to LF? What position does the non-overt operator move to (i.e., is it constrained to move only to Comp or may it be adjoined to some other category)? Finally, there is a correlation between PRO and the subject position of non-finite clauses; can an analysis of PRO that identifies it with a non-overt operator still capture this perceived correlation? The best way to answer these questions is to consider the various control structures case by case and develop an account of these structures under our current assumptions.

We may begin with some cases of undisputed obligatory control as in:

(34)a. John$_i$ tried PRO$_i$ to leave

b. Bill$_i$ wanted PRO$_i$ to kick himself$_i$

c. Mary$_i$ promised PRO$_i$ to talk to Bill

Under our assumptions, these examples would have the representations at some level:

(35)a. [$_{\bar{S}}$ [$_S$ John$_i$ tried [$_Y$ Op$_x$ [$_S$ $\underline{t}_x$ to leave]]]]

b. [$_{\bar{S}}$ [$_S$ Bill$_i$ wanted [$_Y$ Op$_x$ [$_S$ $\underline{t}_x$ to kick himself$_x$]]]]

c. $[_{\bar{S}}\ [_{S}$ Mary$_{i}$ promised $[_{Y}$ Op$_{x}$ $[_{S}$ $\underline{t}_{x}$ to talk to Bill]]]]

We want to know what category "Y" is in each of the above examples. If we identify "Y" with $\bar{S}$, then we will assume that the non-overt operator is in Comp. Further, we want to know if the above examples are S Structure representations or representations at some other level, perhaps LF.

Let us assume, following, for example, Chomsky (1977) and Stowell (1981), that a head that subcategorizes for an $\bar{S}$ may subcategorize for certain features of the Comp of that $\bar{S}$. For example, <u>wonder</u> subcategorizes for a Comp marked [+WH]:

(36)a. John wondered $[_{\bar{S}}$ what$_{j}$ $[_{S}$ he should sell $\underline{t}_{j}$]]

b. *John wondered $[_{\bar{S}}$ that $[_{S}$ he should sell his car]]

Example (36b) would be ruled out because the embedded Comp does not dominate a [+WH] element and, so, cannot be marked [+WH]; the requirement placed on the lower Comp by <u>wonder</u> will not be satisfied. Other verbs, like <u>hope</u>, subcategorize for a [-WH] Comp:

(37)a. Bill hoped $[_{\bar{S}}$ that $[_{S}$ John would buy a new corvette]]

b. *Bill hope $[_{\bar{S}}$ what$_{i}$ $[_{S}$ John would buy $\underline{t}_{i}$]]

Similarly, verbs may specify whether or not the Comp is related to a tensed Infl or a non-finite Infl:

(38)a. John knows [that he is too tall]

b. *John knows [to be too tall]

(39)a. John tried [to enter the race]

b. *John tried [that he would enter the race]

although some verbs place no such restriction on their complement and may take either a finite or non-finite complement:

(40)a. John thought [that Bill is a fool]

b. John thought [Bill to be a fool]

Given that verbs may place restrictions on the content of the Comp of an $\bar{S}$ complement, we need only say that obligatory control verbs subcategorize for a Comp that contains a non-overt operator; technically, this is complement analogous to the class of verbs (e.g., <u>wonder</u>) that subcategorize for a wh-element in the Comp of their sentential complements. Continuing the analogy with verbs like <u>wonder</u>, we can say that a Comp that contains non-overt operator is marked with the feature [+N(on-overt) O(perator)], possibly related in some way to the feature [$\underline{+}$WH], since the two are in complementary distribution. Thus, obligatory control verbs subcategorize for a [+NO] Comp. We may instantiate "Y" in the structures in (35) as $\bar{S}$ and say that the non-overt operator is in Comp:

(41)a. $[_{\bar{S}}\ [_{S}\ \text{John}_{i}\ \text{tried}\ [_{\bar{S}}\ \text{Op}_{x}\ [_{S}\ \underline{t}_{x}\ \text{to leave}]]]]$

b. $[_{\bar{S}}\ [_{S}\ \text{Bill}_{i}\ \text{wanted}\ [_{\bar{S}}\ \text{Op}_{x}\ [_{S}\ \underline{t}_{x}\ \text{to kick himself}_{x}\ ]]]]$

c. $[_{\bar{S}}\ [_{S}\ \text{Mary}_{i}\ \text{promised}\ [_{\bar{S}}\ \text{Op}_{x}\ [_{S}\ \underline{t}_{x}\ \text{to talk to Bill}]]]]$

We turn now to the question of what level of representation

at which the non-overt operator is in Comp. Lasnik & Saito (1984) have argued that restrictions placed on Comp must be satisfied at S Structure. Their argument is based on examples like those in (36). That is, we know that wonder requires a [+WH] Comp since it must have an interrogative complement. Consider the following contrast:

(42)a. John wondered $[_{\bar{S}}$ what$_i$ $[_S$ Bill bought t_i]]

b. *John wondered $[_{\bar{S}}$ $[_S$ Bill bought what]]

The examples in (42) differ minimally in that wh-movement has applied to what in the syntax in (42a) while what has remained in situ in (42b). We assume, as is standard, that wh-movement must apply to wh-elements left in situ in the syntax in the mapping from S Structure to LF. Thus, the LF representation of (42b) will be identical that of (42a). Given that the two structures are indistinguishable at LF, it must be the case that they are differentiated at some level prior to LF. Assuming that the two structures have identical D Structures, then it must be the case that the requirement that the sentential complement to wonder have a [+WH] Comp must be checked at S Structure. Generalizing from this, we can say that such requirements are uniformly checked at S Structure in English.[6] We will assume, then, that

6 As Lasnik & Saito (1984) demonstrate, these requirements are not checked at S Structure universally. In particular, Chinese and Japanese lack syntactic movement so that all wh-elements are in situ at S Structure in these languages. Nevertheless, these languages have verbs which obligatorily take embedded interrogatives (see Huang (1982) and Lasnik & Saito (1984) for

obligatory control verbs subcategorize for a [+NO] Comp and that this restriction is checked at S Structure. Thus, the representations in (42) are S Structure representations.

In preceding sections, it was argued, following Aoun & Clark (1985), that non-overt operators must be identified (that is, receive an index) at S Structure and that this identification took the form either of $\bar{A}$-binding or of predication. In the latter case, the index of the subject of predication was copied onto the $\bar{S}$ containing the non-overt operator. This index, in turn, percolated down to the non-overt operator under the standard conventions of $\bar{X}$ theory. We have also seen that this identification is strictly local. The non-overt operator must find its identifier in the immediately superordinate clause.

Notice that the configurations in (42) collapse with the configurations which Williams (1980) identified as predication structures, although the theory of predication we are assuming is much closer to that of Chierchia (1985)--interested readers should consult the latter for details. Among these structures is:

(43) NP V X

where the NP is the subject of the predication and X is the predicate. This schema represents sentences like the following:

---

details). In these languages, the requirement must be checked at LF rather than at S Structure. This raises interesting possibilities for the analysis of PRO developed in this chapter.

(44)a. [John] [looks] [ill]
NP V AP
b. [Bill] [became] [a logical positivist]
NP V NP
c. [Frank] [grew] [into a sorry example of
NP V PP
mankind]

It is a lexical property of verbs like <u>look</u>, <u>become</u> and <u>grow</u> that their complement is predicated of the subject of S. In order to satisfy this requirement, the complement must come to be coindexed with the subject of S:

(45)a. [John] [looks] [ill]
NP i V AP i
b. [Bill] [became] [a logical positivist]
NP i V NP i
c. [Frank] [grew] [into a sorry example of
NP i V PP
mankind]
i

Our proposal is that control verbs are another instantiation of this schema, where X is taken as equal to $\overline{S}$, somewhat along the lines of Williams (1980). It is a lexical property of these verbs that sentential complement must be coindexed with the subject of S (or some other element specified by the verb).

To take a concrete example, let us consider:

(46) John tried to enter the race.

Example (46), under our analysis, will have the following D Structure:

(47) [$_{\overline{S}}$ [$_{S}$ John tried [$_{\overline{S}}$ [$_{S}$ Op to enter the race]]]]

Since <u>try</u> subcategorizes for a [+NO] Comp, the non-overt operator must move to Comp in the syntax:

(48) $[_{\bar{S}}\ [_{S}$ John tried $[_{\bar{S}}\ [_{Comp[+NO]}$ Op$_x$] $[_{S}$ $\underline{t}_x$ to enter the race]]]]

In addition, <u>try</u> specifies that its $\bar{S}$ complement must be predicated of the subject of S. Thus, the following S Structure results from this indexation plus the normal constraints on percolation:

(49) $[_{\bar{S}}\ [_{S}$ John$_i$ tried $[_{\bar{S}}\ [_{Comp[+NO]}$ Op$_i$] $[_{S}$ $\underline{t}_i$ to enter the race]]]]$_i$

Thus, the non-overt operator in (49) obeys our constraint that non-overt operators be identified at S Structure. This identification obeys the constraint on locality discussed above; the non-overt operator is associated with an element in the superordinate clause.

Let us take another example of control. In particular, let us consider the case of "psych" verbs:

(50)a. [PRO$_i$ shaving himself$_i$] bothers John

b. ??[PRO$_{arb}$ shaving oneself] bothers John

Psych verbs present a problem to theories which treat controlled PRO as an anaphor since the controlled PRO is never c-commanded by its antecedent. Most generally, such theories have treated PRO in these cases as arbitrary; this is often interpreted as meaning that the PRO has pronominal properties. In other words, the PRO may freely range over individuals in domain D. Thus, accidental coreference is possible between the PRO inside the sentential subject and an object NP.

There is a shortcoming in this account, however. If PRO may accidentally corefer with any NP by virtue of ranging freely over elements in domain D, then we would expect that PRO could be taken as coreferential with any NP in the sentence. J. Aoun (personal communication) has pointed out that this is not the case:

(51)a. *[PRO$_i$ shaving himself$_i$] bothers [John$_i$'s mother$_j$]

b. *[PRO$_i$ feeding themselves$_i$] worries [the man in charge of [the lions$_i$]]$_j$

While we will return to cases of nonobligatory control below, we can develop an account of controlled PRO in the above examples. Notice that psyche verbs, like obligatory control verbs discussed above, establish a relation between individuals and propositions. Like obligatory control verbs, we can say that psyche verbs require a predication relation between the subject of S and an object; in other words, psyche verbs require coindexing between subject and object.[7] Unlike obligatory control verbs, the embedded sentence is in subject position rather than object

---

7 Intuitively, the subject of the predication in the case of psych verbs seems to be the object of the verb. This fact indicates that the notion "subject of a predication" does not necessarily correspond with the notion "grammatical subject". "Subject of a predication" more closely corresponds to the notion of "subject" developed in, for example, Geach (1962). It is a matter of some interest that the two notions have a rather close correlation, although they may diverge in instances like psyche verb constructions. A complete exploration of the matter would take us well off the topic of this chapter and will be left as a matter for future research.

position; as a result, the verb does not govern, and so cannot place restrictions on, the embedded S. The D Structure of an example like (52a) would be (52b):

(52)a. to shave himself$_i$ would bother John$_i$

b. $[_{\bar{S}}$ $[_{S}$ Op to shave himself]] would bother John

In the syntax, the non-overt operator could optionally move to Comp, as in:

(53) $[_{\bar{S}}$ Op$_x$ $[_{S}$ $\underline{t}_x$ to shave himself]] would bother John

Obligatory indexation of the subject and object, along with the usual percolation and reindexation conventions, would yield:

(54) $[_{\bar{S}}$ Op$_i$ $[_{S}$ $\underline{t}_i$ to shave himself$_i$]]$_i$ would bother John$_i$

In example (54), the non-overt operator in the Comp of the sentential subject ultimately receives an index under predication and is, as a result, identified as required by our account of the distribution of non-overt operators. Notice that we do not generate examples like those in (51) where the PRO is coindexed with an NP that is not a sister to the verb. This is because the indexation specified by psych verbs must hold between the subject and some complement to the verb; given the usual assumptions about government and complementation, the verb could not specify that a non-complement be coindexed with the subject of S. Chierchia (1985) should be consulted for some discussion of the notion of predication that underlies this discussion.

Another environment where control of PRO seems to be

obligatory is in certain adjuncts:

(55)a. John$_i$ felt old after [PRO$_i$ seeing himself$_i$ in the mirror]

b. *John$_i$ felt old after [PRO$_{arb}$ seeing oneself in the mirror]

The contrast in grammaticality between (55a) and (55b) demonstrates that PRO in these examples may not be taken to be arbitrary. The necessity for PRO to be controlled in examples such as these even over-rides the inherently [+human] interpretation that PRO normally receives:

(56)a. It never snows after [PRO raining] in this part of the world

b. *[PRO raining] would be a disaster now

Example (56b) is anomalous precisely because PRO$_{arb}$ is normally taken as ranging over humans; thus, the restrictions placed on the subject position of weather verbs is violated in this example. Nevertheless, when the weather verb occurs in an adjunct, it may have a PRO subject as demonstrated in example (56a).

Notice, further, that by hypothesis the controller of a PRO in an adjunct may not be lexically specified. This is because an adjunct, by definition, does not satisfy lexical requirements; an adjunct may never be subcategorized for by a lexical head. Nevertheless, there are strong restrictions on the controller of an adjunct PRO:

(57)a. John kissed Mary after [PRO seeing himself in the mirror]
i j i i

b. *John kissed Mary after [PRO seeing herself in the mirror]
i j j j

The examples in (57) show that there is a strong tendency to take the subject of the matrix S as the controller of PRO. If some other NP is to be taken as coreferential with the subject of the adjunct clause, an overt pronominal must be used:

(58) John kissed Mary after [she saw herself in the mirror]
i j j j

The relationship between the PRO in an adjunct and the subject of a sentence is local as can be seen by considering the readings of the following sentence:

(59) John thought that Bill died after PRO seeing himself in the mirror
i j {i,j} {i,j}

If the PRO in the above sentence is taken as coreferential with <u>John</u>, then the entire adjunct clause must be taken as modifying the root S. In other words, the interpretation should be that after John saw himself in the mirror, the thought that Bill died struck him (i.e., John). If, on the other hand, the PRO is taken as being controlled by <u>Bill</u>, then the adjunct must be taken as modifying the embedded S. In this case, the interpretation should be that John thought that Bill died after Bill saw himself in the mirror. It is impossible, however, to get a "crossed" interpretation. That is, if <u>John</u> is taken as the antecedent of PRO, then the adjunct clause cannot be interpreted as modifying

the embedded clause. Thus, example (59) may not be interpreted as meaning that John thought that Bill's death occurred after John saw himself in the mirror. Similarly, if Bill is taken as the antecedent for the PRO then the adjunct may not be interpreted as modifying the root S. In other words, it cannot be the case that John's thinking that Bill had died occurs after Bill saw himself in the mirror. These facts support the contention (found in Williams, 1980, and Chomsky & Lasnik, 1977) that the controller of PRO must be in the superordinate clause, just as in the cases of obligatory control examined above.

The derivation of the above cases should, by now, be apparent. The D Structure of such a case will be like:

(60) $[_{\bar{S}}\ [_{S}$ John felt old [after $[_{\bar{S}}\ [_{S}$ Op seeing himself in the mirror]]]]]

"Move alpha" may apply in the syntax to the non-overt operator, yielding:

(61) $[_{\bar{S}}\ [_{S}$ John felt old [after $[_{\bar{S}}$ Op$_{x}$ $[_{S}$ t$_{x}$ seeing himself in the mirror]]]]]

The adjunct clause will enter into a predication relation with the matrix subject (note that I am assuming that assignment of indices under predication occurs in the syntax and is therefore an S Structure phenomenon). The conventions regarding predication, percolation and reindexing apply yielding the following:

(62) $[_{\bar{S}}\ [_{S}$ John$_i$ felt old [after $[_{\bar{S}}$ Op$_i$ $[_{S}$ $\underline{t}_i$ seeing himself$_i$ in the mirror]]]]]

The non-overt operator in (62) is identified, under predication, with the matrix subject, as required. Since the predication process is restricted to the S containing the subject of the predication, the "crossed" interpretations, discussed above, will not be derived.

Next, we turn to cases of infinitival relatives such as:

(63) $[_{NP}$ the right person [PRO to fix the sink]] could not be found.

The D Structure of an example like (63) is:

(64) $[_{NP}$ the right person$_j$ $[_{\bar{S}}$ $[_{S}$ Op to fix the sink]]

"Move alpha" may apply to the non-overt operator to give:

(65) $[_{NP}$ the right person$_j$ $[_{\bar{S}}$ Op$_i$ $[_{S}$ $\underline{t}_i$ to fix the sink]]

whereupon the $\bar{S}$ may be predicated of the head of the NP causing the relative to become coindexed with the head. The index on the relative will percolate down to the non-overt operator in Comp giving (66) as the S Structure representation:

(66) $[_{NP}$ the right person$_j$ $[_{\bar{S}}$ Op$_j$ $[_{S}$ $\underline{t}_j$ to fix the sink]]$_j$

The non-overt operator is identified under predication at S Structure, as required by our account. We will extend this account of infinitival relatives below.

The account of control developed to this point collapses obligatory control with other cases of identifying non-overt

operators. This idea, in fact, can be traced back to Chomsky (1980), where purposive clauses are treated as having the following S Structure:

(67)a. John bought Bill a dog to play with
b. John bought Bill a dog $[_{\bar{S}}$ PRO$_j$ $[_S$ PRO$_i$ to play with $\underline{t}_j]]$

The purposive clause in (67) contains two PROs. One PRO is, of course, in the subject position of the non-finite clause, and the other occurs in Comp and $\bar{A}$-binds the variable in the object of the preposition, <u>with</u>. Both of the PROs in (67) were then subject to a rule of control. Chomsky's theory of control consisted of the following principles:

(68) $\ldots V \ldots [_{\bar{S}}$ COMP $\ldots$PRO$\ldots]$
where V = [-F] and V and $\bar{S}$ c-command one another.

(69) NP is a <u>controller</u> for V in (68) if
a. NP is an indexed NP properly related to V;
b. if V = [+S(ubject) C(ontrol)] then NP is the subject of V.

Some comments on (68) and (69) are perhaps in order. The feature [±F] is used to indicate whether or not a particular verb is a control verb; it is, in some ways, quite analogous to the feature [±NO] in this chapter. The notion "properly related to V" is intended to include argument positions like subject, object and indirect object. We could, thus, express the intended relation in terms of the thematic grid on the verb. Finally, the feature [±SC] stipulates that if a verb is [+SC], then its subject will control the PRO. This stipulation is necessary since a minimal

distance principle is assumed to be involved with the rule of control in the unmarked case. Chomsky then gives a rule of control:

(70) In (68),
a. if Comp ≠ null and V has no controller, then PRO is assigned arb;
b. PRO is assigned the index of the nearest controller.

Clause (a) in (70) is intended to account for examples like:

(71) John won't tell [$_{\bar{S}}$ how [$_{S}$ PRO$_{arb}$ to solve the problem]]

where PRO is taken as ranging over the set of humans rather than as being controlled by John. We will return to these examples below. Clause (b) is the minimal distance principle alluded to above. To see how this principle works, let us return to the application of control to the double PRO cases:

(72) John$_1$ bought Bill$_2$ a dog$_3$ [$_{\bar{S}}$ PRO$_j$ [$_S$ PRO$_i$ to play with t$_j$]]

The indexing and control algorithm goes from the top of the tree down so we are first concerned with the PRO in Comp. Since a dog is properly related to the verb buy, it may count as a controller. Clause (b) requires PRO to be assigned the index of the nearest possible controller. In this instance, the nearest possible controller is a dog:

(73) John$_1$ bought Bill$_2$ a dog$_3$ [$_{\bar{S}}$ PRO$_3$ [$_S$ PRO$_i$ to play with t$_3$]]

The algorithm now applies to the next PRO, which is in the subject position of the purposive. Prima facie, the indexing algorithm appears to identify a dog as the controller of this PRO since it is the nearest NP that can count as a controller. In current terms, coindexing the PRO in the subject position of the purposive would induce a violation of condition C, since the PRO would locally A-bind the variable in the object position of the preposition.[8] Since this indexing would violate a core principle, we may take it as given that a dog is not a possible controller for this PRO. This leaves Bill as the nearest possible controller for the PRO. The indexing algorithm thus predicts the following indexing:

(74) John$_1$ bought Bill$_2$ a dog$_3$ [$_{\bar{S}}$ PRO$_3$ [$_S$ PRO$_2$ to play with $\underline{t}_3$]]

It should be noted that the example just discussed is the "dative shifted" version of:

---

8 Chomsky (1980) did not have recourse to the Binding Theory in ruling out a dog as a possible controller. Instead, he stipulates a bijective relation between controllers and controlled elements. Once an NP has been identified as a possible controller, it is no longer available as a controller for some other element. The effect of this stipulation, along with the top-down ordering of the algorithm and the minimal distance principle, is to arrange possible controllers in a push-down stack where only the top-most NP on the stack is an available controller and where each controller is popped off the stack when it has been identified as a controller. The feature [±SC] is necessary to get around this push-down stack arrangement in examples like:

(i) John$_i$ promised Bill$_j$ [PRO$_{\{i,*j\}}$ to leave]

(75) John $_1$ bought a dog $_3$ for Bill $_2$ [$_{\bar{S}}$ PRO $_3$ [$_{\bar{S}}$ PRO $_2$ to play with $\underline{t}_3$]]

The control properties of (75) are identical to those of (74) despite the difference in word order. This fact implies that some care must be taken in the definition of "nearest controller" in (70). In brief, Chomsky defines "nearest controller" to be the nearest NP such that mutual c-command holds between that NP and the $\bar{S}$ containing the PRO.[9] Failing that, the nearest controller is defined to be the closest NP (presumably in terms of the string) that is c-commanded by the $\bar{S}$. Thus, the nearest NP for control of the PRO in Comp in (75) is a dog, despite the fact that, purely in terms of the ordering of the string of words, Bill appears to be the nearest controller. This is because Bill is the object of the preposition for and so does not c-command the $\bar{S}$ containing the PRO.

The above discussion of control in the "On Binding" framework of Chomsky (1980) is intended to give something of a historical precedent for the position taken in this chapter; namely, that control of a PRO is best treated as a special case of the analysis of non-overt operator constructions like purposive clauses. There are some important differences between the two treatments which bear noting. First, Chomsky assimilates

9 Notice the similarity in this part of the definition of "nearest controller" and the definition of a predication structure given in Williams (1980) where mutual c-command between the subject and the predicate is assumed.

the binding of non-overt operators to a special subsystem of grammar, control theory. The project in this chapter is to assimilate control theory, as far as possible, to a generalized system of binding. This latter approach, itself, has some precedent in recent work (see, for example, Bouchard (1984), Lebeaux (1984), Manzini (1983) and Sportiche (1983), to name but a few).

Second, there some empirical problems with the approach taken in Chomsky (1980). For our purposes, it is sufficient to focus on one such problem. We have seen that Chomsky blocks control of PRO in only a very restricted set of environments; crucially, the subject position of a purposive was assumed not to be an environment for arbitrary PRO. This last assumption was necessary in order to induce the putatively obligatory nature of double control in these structures. Let us consider control into a purposive like the following:

(76) John_1 bought a dog_2 [$_{\bar{S}}$ PRO_i [$_S$ PRO_j to play with $\underline{t}_i$]]

As we have seen, the PRO in Comp in (76) will be controlled by the nearest controller, <u>a dog</u>. We must now determine the controller of the PRO in the subject position of the purposive; since <u>a dog</u> is already a controller, we are left with only one possible nearest controller, the subject of the matrix clause. After application of the control algorithm with respect to the subject position of the purposive, we predict the following indexing:

(77) John$_1$ bought a dog $[_{\bar{S}}$ PRO$_2$ $[_S$ PRO$_1$ to play with $\underline{t}_2]]$

This indexing seems to accord with the naive interpretation of example (76). This example should be contrasted, however, with an example like:

(78) Armageddon Inc. makes {bombs/them} to drop on communists.

Example (78) will have the following S Structure:

(79) Armageddon Inc.$_1$ makes bombs $[_{\bar{S}}$ PRO$_i$ $[_S$ PRO$_j$ to drop $\underline{t}_i$ on communists]]

We have already exemplified how the PRO in Comp comes to be controlled by the object of the verb, <u>make</u>, under Chomsky's theory. At issue here is the problem of the control properties of the PRO in the subject position of the purposive clause. Under the "On Binding" assumptions, the PRO will be obligatorily controlled by the subject of the matrix S, <u>Armageddon Inc.</u>, as shown in (80):

(80) Armageddon Inc.$_1$ makes bombs $[_{\bar{S}}$ PRO$_2$ $[_S$ PRO$_1$ to drop $\underline{t}_2$ on communists]]

Recall that we are assuming that indices are eventually interpreted by a function that maps indices onto elements of Domain D and that this function is constrained to preserve identity of indices; if two indices are identical in the syntax then they must be mapped onto the same element of Domain D by the interpretive function. This assumption about the link between

syntactic indices and the interpretation of reference implies that (80) is interpreted as true under some interpretation if and only if the same individual that makes bombs also drops the bombs on communists. Surely, example (80) implies nothing of the sort; the sentence in (80) is consistent with a state of affairs where Armageddon Inc. makes bombs and someone else drops these bombs on communists.

To press the point somewhat further, consider the following examples:

(81)a. John bought these books to read to the children.
b. The school board bought these books to read to the children.

In the absence of any prior context, many speakers are inclined to say that in order for (81a) to be true, John must both buy the books and read them to the children; in other words, John is apparently taken as the controller of the PRO in the subject position of the purposive clause in (81a). Example (81b) differs from (81a) only in the NP that occupies the subject position of the matrix clause; we may assume that there is no syntactically relevant structural difference between (81a) and (81b). Nevertheless, speakers I have consulted have consistently allowed an interpretation where the school board buys the books and some other party reads them to the children. Thus, identifying the PRO subject of a purposive with a position of obligatory control runs the risk of destroying any connection at all between the syntactic notion of "identity of indices" and the semantic notion of "identity of reference"(but see the discussion of "arbitrary"

control in the next section).

Some comment is in order on the tendency to take the subject of S as controlling the subject of the purposive clause. A complete discussion of this proclivity will, of necessity, foreshadow some of the analysis that will be developed below, so comments here will be of a somewhat tentative nature. Taking the examples at hand, notice that the verb buy allows the appearance of a benefactive argument; this benefactive may appear as the object of the preposition for, or it may appear immediately after the verb in "dative shift" constructions:

(82)a. John bought Bill these books to read to the children.
b. The school board bought these books for the teachers to read to the children.

This benefactive argument need not be realized in the syntax, as the examples in (81) illustrate. Notice that realization of this argument by an empty category is not possible in English under the assumption that this empty category is a pure pronominal (pro) since English apparently lacks pro, as can be seen from the ungrammaticality of:

(83) *pro left

and similar examples. Let us suppose, however, that at Logical Form (or, perhaps, same later level of representation which is semantically interpreted), elements that are in a verb's thematic grid must be realized. In particular, let us assume the existence of the following principle:

(84) Completeness
Predicates must be closed.

We interpret this completeness principle as requiring that implicit arguments must be realized at some level of representation and that, furthermore, these arguments must be bound. Such a principle would require that arguments that have been suppressed in the syntax, for example, the external thematic role in passive constructions, must be realized at some level of representation. For the sake of discussion, I will assume that this level of representation is LF. A passive S Structure as in (85a) will be associated with an LF representation as in (85b):

(85)a. $[_{\underline{S}}$ $[_{S}$ $[_{NP}$ the thief$]_{i}$ was seen $\underline{t}_{i}$]]

b. $[_{\underline{S}}$ $[_{S}$ [e] was $[_{VP}$ (Ex) $[_{VP}$ seen $[_{NP}$ the thief]]
$\underline{x}$]]]

where '(Ex)' is an existential quantifier. Notice that I further assume that NP traces may be covered in the sense that the head of an A-chain may be returned to its D Structure θ-position. Notice, finally, that the LF in (85b) contains an additional external argument position. Assuming that Case marking is an S Structure phenomenon, this induced argument need not receive Case since it is not syntactically realized until LF. Thus, the position may be adjoined to VP and no insertion of a Case marker need be assumed The operator introduced by this process does not induce any scope ambiguities, as can be seen in examples like:

(86) Every man is loved.

Example (86) is not ambiguous between a reading where some one

individual loves every man and a reading where every man is the object of somebody's affections (not necessarily the same someone for each man). I will therefore assume that the existential quantifier is simply adjoined to some projection of the element that assigns the thematic role being realized. In the case of external thematic roles, this category is, I will assume for the moment, VP. Thus, the purposive clause examples in (81) will have LF representations similar to those in (87) (leaving aside some details for later discussion):

(87)a. $[_{\bar{S}}\ [_{S}$ John $[_{VP}$ (Ex) $[_{VP}$ bought $[_{NP}$ these books$]_i$ $\underline{x}]$
$[_{\bar{S}}$ Op$_i$ $[_{S}$ Op to read $\underline{t}_i$ to the children$]]]]]$

b. $[_{\bar{S}}\ [_{S}$ The school board $[_{VP}$ (Ex) $[_{VP}$ bought
$[_{NP}$ these books$]_i$ $\underline{x}$ $[_{\bar{S}}$ Op$_i$ $[_{S}$ Op to read $\underline{t}_i$ to the
children$]]]]]]$

Finally, we need only assume that the empty element in the subject position of the purposive undergoes QR in the mapping to LF representation and may then be identified by the existential operator induced from the argument structure of the predicate of the superordinate clause as in (88):

(88) $[_{\bar{S}}\ [_{S}$ The school board $[_{VP}$ (Ex) $[_{VP}$ bought
$[_{NP}$ these books$]_i$ $\underline{x}$ $[_{\bar{S}}$ Op$_i$ $[_{S}$ Op$_x$ $[_{S}$ $\underline{t}_x$ to read $\underline{t}_i$ to
the children$]]]]]]$

A rough paraphrase of the above representation is the school board has the property of there being some individual(s) such that the school board bought these books for the individual(s) and the individual(s) read(s) these books to the children.

Returning to the double control cases, this type of sentence has the following representation:

(89) $[_{\bar{S}}\ [_{S}\ \text{John}\ [_{VP}\ (\text{Ex})\ [_{VP}\ \text{bought}\ [_{NP}\ \text{these books}]_{i}\ \underline{x}]$
$[_{\bar{S}}\ \text{Op}_{i}\ [_{S}\ \text{Op}_{x}\ [_{S}\ \underline{t}_{x}\ \text{to read}\ \underline{t}_{i}\ \text{to the children}]]]]]]$

We can paraphrase the above representation as: John has the property of buying books for someone such that that someone reads the books to the children. We assume that the existential quantifier may range over the set of individuals who may benefit from the buying of books (presumably the set of humans). This set will, of course, possibly include John himself; thus, John is an individual who may benefit from the book buying activity. The preference to take John as instantiating the benefactive argument could be stated as a pragmatic preference; we have seen, in example (81) above, that taking the subject as the benefactive is not simply a matter of logic, although, as we have demonstrated, this reading is a logical possibility. We will have occasion to extend this analysis of double control, below, so I will delay any further discussion of these examples.

The result of the preceding discussion is the so-called double control readings in purposives do not follow from standard control theory in the sense in which this theory is normally interpreted. This result should not be viewed as a purely negative one since it is in accord with some facts about PRO that have been noted previously. That is, when the Comp of the minimal $\bar{S}$ containing a PRO is not null, then the PRO is subject

to "arbitrary interpretation":

(90)a. John_i asked $[_{\bar{S}}$ $[_S$ PRO$_i$ to shave himself$_i$]]

b. *John_i asked $[_{\bar{S}}$ $[_S$ PRO$_{arb}$ to shave oneself]]

c. John_i asked $[_{\bar{S}}$ how $[_S$ PRO$_{arb}$ to shave oneself]]

Examples similar to those in (90) have been taken as implying that control is blocked whenever Comp is not empty (see, for example, Manzini's treatment of control, Manzini (1983)). Abstracting somewhat, we can identify the following as an environment for arbitrary control:

(91) NP may not control PRO in the environment:

$$...[_S\ ...NP...[_{\bar{S}}\ \text{Comp}\ [_S\ \text{PRO}...]]...]...$$

where Comp is non-null.

If the non-overt operator analysis of purposives is correct, then a purposive clause instantiates the schema in the following way:

(92) $$...[_S\ ...NP...[_{\bar{S}}\ \text{Op}_i\ [_S\ \text{PRO}...]]...]...$$

The Comp of the minimal $\bar{S}$ containing the PRO is non-null since it is occupied by the non-overt operator. Thus, we would not expect control to apply to the subject of a purposive clause, given the fundamental correctness of the schema in (91).

Let us consider the fate of an empty element (=PRO) under the analysis of PRO as a non-overt operator developed in this chapter. We begin with the following D Structure:

(93) John bought these books $[_{\bar{S}}$ $[_S$ $[_{NP}$ e] to read $[_{NP}$ e] to the children]]

where $[_{NP}\ e]$ is taken to be an empty element at D Structure that will ultimately be analyzed as a non-overt operator. As a first step, let us suppose that the empty element in the object position of <u>read</u> moves to the Comp of the purposive clause:

(94) John bought these books $[_{\bar{S}}\ [_{NP}\ e]_i\ [_S\ [_{NP}\ e]$ to read $\underline{t}_i$ to the children]]

Now, suppose that the empty element in the subject position of the purposive clause undergoes "move alpha". From the results of the previous chapter, we are guaranteed that this empty element may only move to the Comp of the minimal $\bar{S}$ containing it, since this $\bar{S}$ does not satisfy any lexical requirements (see chapter 2). Hence, the only possible S Structure, given movement of the subject, is:

(95) John bought these books $[_{\bar{S}}\ [_{NP}\ e]_i\ [_{NP}\ e]_j\ [_S\ \underline{t}_j$ to read $\underline{t}_i$ to the children]]

The representation in (95) is a classical example of a violation of superiority. If superiority effects are derived by means of the ECP, or some other form of head government, at LF, it is clear that (95) violates this condition since the trace in the subject position of the purposive clause is not properly governed due to the fact that its antecedent is in a branching Comp. Thus, we will assume that the only possible S Structure for this example is:

(96) John bought these books $[_{\bar{S}}\ [_{NP}\ e]_i\ [_S\ [_{NP}\ e]$ to read $\underline{t}_i$ to the children]]

We will return to the question of what prevents to empty element in the object position of read from remaining in situ in the syntax while the element in subject position moves to Comp. For the moment, let us note that the "arbitrary" PRO in (96) is an element that remains in situ in the syntax.

Notice that an empty element, $[_{NP}\ e]$, is left in an A-position in (96). Let us suppose that this empty element has no inherent interpretation and is, furthermore, unbound at S Structure. Since this element is unbound, it has no way of receiving an interpretation from some antecedent. We will take it as a given that an element that does not refer or, in the case of variables, does not receive a range over elements in domain D must be taken as a expletive. Thus, the empty element in the subject position of the purposive clause in (96), all things held equal, counts as an expletive. The subject position of the purposive, however, receives a thematic role under predication from the VP. We will take the following as a well-formedness constraint on LF representations:

(97) Non-referential elements may not bear a thematic role.

A constraint like that in (97) has long been assumed in government-binding theory, see, for example, Chomsky (1981). If (97) is correct, then the representation in (96) is not well-formed; the empty element in the subject position of the purposive clause has no inherent reference, and so must be an

expletive, but it bears a thematic role, and so may not be an expletive. The only way to circumvent this obvious contradiction is for the empty element to pick up a range from some other element.

Let us suppose that QR may optionally apply to the empty element in the subject position of the purposive as in (98):

(98) John bought these books $[_{\bar{S}}$ [e] $[_{NP_i}$ [$[_{S}$ e] $[_{NP_j}$ $[_{S}$ $\underline{t}_j$ to read $\underline{t}_i$ to the children]]

The empty element now occupies an $\bar{A}$-position and may be treated as a non-overt operator; meaning that this element is now treated as an $\bar{A}$-anaphor which may pick up a range from some other operator in S. Recall that we are also assuming that implicit arguments of the verb must be realized at LF (see the discussion of (84), above). The implicit benefactive argument of <u>buy</u> must, therefore, be realized and bound by an existential quantifier:

(99) $[_{\bar{S}}$ $[_{S}$ John $[_{VP}$ (Ex) $[_{VP}$ bought $[_{NP}$ these books] $\underline{x}_i$]
$[_{\bar{S}}$ $[_{S}$ e] $[_{NP_i}$ $[_{S}$ e] $[_{NP_j}$ $[_{S}$ $\underline{t}_j$ to read $\underline{t}_i$ to the children]]]]]

Let us, now, consider the governing category and possible $\bar{A}$-antecedents of the non-overt operator created in the mapping from S Structure to LF. Notice that the purposive clause contains no accessible SUBJECT for this operator. The nearest accessible subject is the Infl of the superordinate clause, as we would expect given our earlier discussion of non-overt operators. The representation in (99) contains two elements in $\bar{A}$-positions which

could act as a possible antecedent for the new non-overt operator. First, there is the non-overt operator in the Comp of the purposive. This element has received a range under predication of the purposive clause on <u>these</u> <u>books</u>. Let us suppose that the new non-overt operator is coindexed with this element:

(100) [$_{\bar{S}}$ [$_{S}$ John [$_{VP}$ (Ex) [$_{VP}$ bought [$_{NP}$ these books]$_{i}$ $\underline{x}$]
[$_{\bar{S}}$ [$_{NP}$ e]$_{i}$ [$_{S}$ [$_{NP}$ e]$_{i}$ [$_{S}$ $\underline{t}_{i}$ to read $\underline{t}_{i}$ to the
children]]]]]

Notice that reindexing the non-overt operator forces the variable that it binds to be reindexed simply by convention. The variable in the subject position of the purposive now bears the same index as the variable in the object position of the purposive. As a result, the representation in (100) violates condition C of the (A-system) of binding. Since the binding theory applies at least at LF (if not also at S Structure), example (99) is filtered out. We conclude that the non-overt operator in Comp is not a possible $\bar{A}$-antecedent for the non-overt operator adjoined to S.

There is another potential antecedent for the non-overt operator in the purposive. Namely, the existential operator induced from the argument structure of <u>buy</u> during the mapping from S Structure to LF. Let us suppose that the non-overt operator is coindexed with this existential operator:

(101) [$_{\bar{S}}$ [$_{S}$ John [$_{VP}$ (Ex) [$_{VP}$ bought [$_{NP}$ these books]$_{i}$ $\underline{x}$]
[$_{\bar{S}}$ [$_{NP}$ e]$_{i}$ [$_{S}$ [$_{NP}$ e]$_{x}$ [$_{S}$ $\underline{t}_{x}$ to read $\underline{t}_{i}$ to the
children]]]]]

No violation of condition C of the binding theory occurs in representation (101) since every variable is A-free in the domain of its operator. Furthermore, every non-overt operator in (101) has picked up a range, as is indirectly required by convention (97), above; the non-overt operator in the Comp of the purposive clause receives a range under predication with the object of <u>bought</u> and the non-overt operator adjoined to the S node of the purposive clause receives a range by virtue of being $\bar{A}$-bound by the existential operator.

The above example should be contrasted with an example where an implicit argument may not be induced. Consider the contrast in (102):

(102)a. John_i walked the dog $[_{\bar{S}}$ Op $_i$ $[_S$ $\underline{t}_i$ to exercise it]]

b. *John_i walked the dog $_j$ $[_{\bar{S}}$ Op $_j$ $[_S$ Op to exercise $\underline{t}_j$]]

In example (102a) there is only one non-overt operator which originates as the subject of the purposive at D Structure. This non-overt operator will move to Comp in the syntax and will be identified under predication with the subject of the matrix S. Let us, now, consider the derivation of (102b); this example will have the following D Structure:

(103) John_i walked the dog $_j$ $[_{\bar{S}}$ $[_S$ Op to exercise Op]]

The non-overt operator in object position moves to Comp in the syntax and predication applies between the purposive clause and the subject of the matrix S:

(104) John_i walked the dog_j $[_{\bar{S}}$ Op$_i$ $[_S$ Op to exercise $\underline{t}_i$]]

QR applies in the mapping from S Structure to LF to the empty element in the subject position of the purposive clause, yielding the following structure:

(105) John_i walked the dog_j $[_{\bar{S}}$ Op$_i$ $[_S$ Op$_k$ $[_S$ $\underline{t}_k$ to exercise $\underline{t}_i$]]

As shown, the representation in (105) is not well-formed since the variable in the subject position of the purposive is not assigned a range by the non-overt operator that binds it. Since the non-overt operator is an $\bar{A}$-anaphor, it may not be coindexed with the subject of the matrix S, John, since condition A of the $\bar{A}$-system of binding would be violated. Furthermore, if indexing obtains between the subject of S and the non-overt operator in question condition C of the binding theory will be violated since the variable in the subject position of the purposive will A-bind the variable in object position.

Similarly, suppose that the non-overt operator is coindexed with the object of walk. If the object of walk c-commands the purposive clause, then it will A-bind the non-overt operator in question, which violates condition A of the $\bar{A}$-system of binding; if the object of walk does not c-command the purposive clause, then the non-overt operator in question is free and, hence, lacks a range at LF and the representation is indirectly ruled out by principle (97). Since there is no additional argument that can be induced from the argument structure of walk, there is no $\bar{A}$-element available to bind the non-overt operator. Hence, the

non-overt operator may not provide a range to the variable in the subject position of the purposive clause. As a result, the subject position of the purposive must be interpreted as non-referential, in the sense of (97), and may not bear a thematic role. Since the subject position of the purposive does, in fact, receive a thematic role, the representation in (105) cannot be well-formed, since an expletive element bears a thematic role. Thus, under any assumptions about its indexing, the representation in (105) cannot be saved, which is the desired result.

We will now turn to the question of what principle or principles restrict the distribution of non-overt operators which originate in non-subject position and the related question of what principle forces a non-overt operator originating in non-subject position to move in the syntax as opposed to undergoing QR in the mapping to LF.

Turning first to the former question, it should be noted that non-overt operators originating in non-subject position have a very restricted distribution. Let us take the example of a non-overt operator in object position of a verb at D Structure and consider the environments where the output is well-formed:

(106)a. $[_{NP}$ a book $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to read $\underline{t}_i$]]] would be nice

b. $[_{NP}$ a well-qualified person $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to hire $\underline{t}_i$]]] will not be easy to find around here

c. [some papers [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to send $\underline{t}_i$ to John]]]
NP
are on your desk

(107)a. This book is [hard [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to read $\underline{t}_i$]]]
AdjP

b. A well-qualified person is [easy [$_{\bar{S}}$ Op$_i$ [$_S$ PRO
AdjP
to find $\underline{t}_i$]]]

c. These papers will be [easy [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to
AdvP
send $\underline{t}_i$ to John]]]

(108)a. John bought a book [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to read $\underline{t}_i$]]

b. John located a well-qualified person [$_{\bar{S}}$ Op$_i$ [$_S$ PRO
to hire $\underline{t}_i$

c. Bill stacked some papers [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to send $\underline{t}_i$ to
John]]

(109)a. Which book$_i$ did John buy $\underline{t}_i$ without [$_{\bar{S}}$ Op$_i$ [$_S$ PRO
reading $\underline{t}_i$]]

b. Who$_i$ did the committee interview $\underline{t}_i$ without [$_{\bar{S}}$ Op$_i$
[$_S$ PRO hiring $\underline{t}_i$]]

c. Which papers$_i$ did Bill put $\underline{t}_i$ in an envelope
without [$_{\bar{S}}$ Op$_i$ [$_S$ PRO sending $\underline{t}_i$ to Bill]]

(110)a. A book is [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to read $\underline{t}_i$]]

b. These papers are [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to send $\underline{t}_i$ to John]]

c. A shovel is [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to dig with $\underline{t}_i$]]

The examples in (106) represent the case in which a non-overt operator may originate in object position and move to the Comp of an $\bar{S}$ that modifies a noun. The examples in (107) are cases in which a non-overt operator moves to the Comp of a clause that is

a complement to an adjective. The examples in (108) and (109) are cases in which the non-overt operator moves to the Comp of a clause that is in an adjunct position. Finally, the examples in (110) are cases in which the non-overt operator moves to the Comp of a clause that is in a copular predicate structure.

The examples in (106-110) should be contrasted with the following structures:

(111)a. *[$_{\bar{S}}$ Op$_i$ [$_S$ PRO to read $\underline{t}_i$]] is a problem with this book

b. *[$_{\bar{S}}$ Op$_i$ [$_S$ PRO to hire $\underline{t}_i$]] is not so simple

c. *[$_{\bar{S}}$ Op$_i$ [$_S$ PRO to send $\underline{t}_i$ to John]] may be difficult

(112)a. *the man wants [$_{\bar{S}}$ Op$_i$ [$_S$ John {will/to} hit $\underline{t}_i$]]

b. *this person tried [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to hire $\underline{t}_i$]]

c. *these students asked [$_{\bar{S}}$ Op$_i$ [$_S$ PRO to send $\underline{t}_i$ to John]]

(113)a. *the man worried [$_{PP}$ about [$_{\bar{S}}$ Op$_i$ [$_S$ John´s hitting $\underline{t}_i$]]]

b. *the applicants were concerned [$_{PP}$ over [$_{\bar{S}}$ Op$_i$ [$_S$ the committee´s hiring $\underline{t}_i$]]]

In the examples of (111-113), the clause containing the non-overt operator in question as a subject, an object or an object of a preposition. Notice that each of the structures in (111-113) may take a non-overt operator in Comp, just so long as that non-overt operator originates in the subject position of a non-finite

clause:

(114)a. $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to read this book]] is a problem

b. $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to hire the right person]] is not so simple

c. $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to send these papers to John]] may be difficult

(115)a. the man wants $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to hit John]]

b. the committee tried $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to hire the most qualified person

c. these students asked $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to be sent $\underline{t}_i$ to John]]

(116)a. the man worried $[_{PP}$ about $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ hitting John]]]

b. the committee was concerned $[_{PP}$ over $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ hiring the best applicant]]]

We can place some organization on the above facts if we consider them in light of Case theory. Let us take the following simple form of Case theory as basically correct:

(117)a. A tensed Infl assigns nominative Case to the position it governs;

b. V assigns accusative Case to the position it governs

c. P assigns oblique Case to the position it governs

Notice that nouns and adjectives do not assign Case and, furthermore, adjunct positions are not Case marked. In the examples in (106-110), the $\bar{S}$ containing the non-overt operator which originates in non-subject position is not in a Case marked

position, while the $\overline{S}$ containing a similar non-overt operator in the examples in (111-113) are all in Case marked positions. We will assume that the Case assigned to an $\overline{S}$ percolates down to the head of $\overline{S}$, namely, Comp. The entire set of data in (106-116) can be schematized in the following way:

(118)a. ... Op ... t ...
[-Case] [-Case]

b. ... Op ... t ...
[+Case] [-Case]

c. ... Op ... t ...
[-Case] [+Case]

d. *... Op ... t ...
[+Case] [+Case]

The schema in (118a) represents movement of a non-overt operator from a position that is not Case-marked to the Comp of an $\overline{S}$ that is not Case-marked; thus, it corresponds to control of a subject in a clause that is in an adjunct position:

(119)a. John$_i$ left [$_{\overline{S}}$ Op$_i$ [$_{\overline{S}}$ t$_i$ to go to the store]]

b. Bill$_i$ ate the fondue [$_{\overline{S}}$ Op$_i$ [$_{\overline{S}}$ t$_i$ to please Mary]]

c. John$_i$ felt strange [$_{\overline{S}}$ Op$_i$ [$_{\overline{S}}$ t$_i$ drinking all the beer]]

The schema in (118b) represents the movement of a non-overt operator from a position that is not Case-marked to the Comp of an $\overline{S}$ which is assigned Case. This case corresponds to control of the subject of a clause in an A-position or one that follows a preposition:

(120)a. $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ having his picture taken]] annoys John$_i$

b. John$_i$ wanted $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ to buy a new car]]

c. John$_i$ became enraged $[_{PP}$ after $[_{\bar{S}}$ Op$_i$ $[_S$ $\underline{t}_i$ reading the letter from the IRS]]]

The schema in (118c) represents movement of a non-overt operator from a Case-marked position to the Comp of a clause that is not in a Case-marked position. This schema represents tough movement structures, infinitival relatives, purposive clauses, parasitic gaps, some sentential predicates and, possibly, topicalization:

(121)a. John$_i$ is $[_{AdjP}$ easy $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to please $\underline{t}_i$]]]

b. [NP the one $[_{\bar{S}}$ Op$_i$ $[_S$ PRO see $\underline{t}_i$]]] is John

c. John bought the car$_j$ $[_{\bar{S}}$ Op$_j$ $[_S$ PRO to fool around with $\underline{t}_j$]]

d. Which book$_i$ did Bill buy $\underline{t}_i$ without $[_{\bar{S}}$ Op$_i$ $[_S$ PRO reading $\underline{t}_i$

e. A shovel$_i$ is $[_{\bar{S}}$ Op$_i$ $[_S$ PRO to dig with $\underline{t}_i$]]

f. [this book$_i$ $[_{\bar{S}}$ Op$_i$ $[_S$ I've read $\underline{t}_i$ a hundred times]]]

Notice that in the cases represented in (118a-c), there has been at most one Case-marked position in the link between the non-overt operator and the variable which it binds. The schema represented in (118d) differs from the previous one in that the link between the non-overt operator and its variable involves two Case-marked positions: The variable, itself, receives Case and the non-overt operator lands in the Comp of a clause that is

Case-marked. This schema represents the ungrammatical structures in examples (111-113). It is interesting to observe that if the element in Comp is lexicalized, the structures become grammatical (provided, of course, that we control for the lexical specifications of some of the verbs):

(122)a. $[_{\bar{S}}$ what $_i$ $[_S$ PRO to read $\underline{t}_i$]] is always a problem

b. $[_{\bar{S}}$ who $_i$ $[_S$ PRO to hire $\underline{t}_i$]] is not so simple

c. $[_{\bar{S}}$ what $_i$ $[_S$ PRO to send $\underline{t}_i$ to John]] may present some difficulties

(123)a. the man wants $[_{\bar{S}}$ what $_i$ $[_S$ John has $\underline{t}_i$]]

b. this person knows $[_{\bar{S}}$ who $_i$ $[_S$ PRO to hire $\underline{t}_i$]]

c. the students asked $[_{\bar{S}}$ what $_i$ $[_S$ PRO to send $\underline{t}_i$ to John]]

(124)a. the man worried $[_{PP}$ about $[_{\bar{S}}$ what $_i$ $[_S$ John said $\underline{t}_i$]]]

b. the applicants were concerned $[_{PP}$ over $[_{\bar{S}}$ who $_i$ $[_S$ the committee would hire $\underline{t}_i$]]]

Given the complementary distribution between non-overt operators and overt operators in the environments in (118), let us assume that the overt operators may be inserted as a phonological spell-out of a non-overt operator. We must now specify the conditions under which this spell-out occurs.

To come to some understanding of the above paradigm, let us

take a definition of the notion "chain" (from Chomsky, 1981):[10]

(125) $C = (A_1, \ldots, A_n)$ is a chain if and only if

a. A_1 is an NP

b. A_i locally A-BINDS A_{i+1}

c. for $i > 1$, A_i is a non-pronominal empty category, or A_i is A-free

d. C is maximal, i.e., is not a proper subsequence of a chain meeting (a-c)

Notice that the above definition only holds for binding of A-positions since clause (b) stipulates that a chain may only be formed if successive links in the chain are A-bound by their predecessors. In general, it would seem that a chain may only bear one Case-marked position. Consider a structure like:

(126)a. It seems to John that Bill is sick.

b. *John$_i$ seems to $\underline{t}_i$ that Bill is sick

c. Bill$_j$ seems to John $\underline{t}_j$ to be sick

(127)a. It struck John that Bill was sick.

b. *John$_i$ struck $\underline{t}_i$ that Bill was sick

c. John$_i$ was struck $\underline{t}_i$ that Bill was sick

The subject position in (126a) and (127a) is a $\overline{\theta}$-position and,

---

10 Note that this definition bears some similarity to the definition of θ-chain in Safir (1985) where, among other things, the notion of BIND is eliminated. "BIND" in the definition refers to both subscripting and superscripting, while "bind" covers subscripting only. The notion of "BIND" was required to cover certain cases of rightward movement where the trace left by movement c-commands the landing site. See Safir (1985) for discussion.

so, should be available as a landing site for a moved NP, a hypothesis which is confirmed by the raising structure in (126c) and the passive in (127c). Notice that an NP originating in a Case-marked position may not land in subject position, as shown in (126b), where the NP receives oblique Case from the preposition, and (127b), where the NP receives accusative Case from the verb. It is reasonable to assume that the relationship between Case-marking and A-chains is biunique: An A-chain may not contain two Case-marked positions.

Let us generalize the definition of chain to include both A-binding and $\bar{A}$-binding:

(128) $C = (A_1, \ldots, A_n)$ is an X-chain if and only if

a. A_1 is an NP

b. A_i locally X-binds A_{i+1}

c. C is maximal, i.e., is not a proper subsequence of a chain meeting (a-b)

We can say that a non-overt operator must be phonologically spelled-out if it occurs in a Case-marked Comp and heads an $\bar{A}$-chain which bears Case. Assuming that an element that head a Case-marked $\bar{A}$-chain inherits (and, thus, bears) the Case of that chain, we can notate the restriction as follows:

(129) $*[_{\substack{\text{Comp} \\ [+\text{Case}]}}\ \underset{[+\text{Case}]}{e}\ ]$ where e is a non-overt operator

The scheme in (118) can be renotated as follows:

```
(130)a.  ...  Op     ...   t   ...
           [-Case]     [-Case]
              ↑_________|

         [              e      ]
          Comp     [-Case]
         [-Case]

     b.  ...  Op     ...   t   ...
           [+Case]     [-Case]
              ↑_________|

         [              e      ]
          Comp     [-Case]
         [+Case]

     c.  ...  Op     ...   t   ...
           [-Case]     [+Case]
              ↑_________|

         [              e      ]
          Comp     [+Case]
         [-Case]

     d.  *...  Op     ...   t   ...
           [+Case]     [+Case]
              ↑_________ |

         *[             e      ]
           Comp     [+Case]
          [+Case]
```

We would, therefore, expect that a non-overt operator can head a Case-less $\overline{A}$-chain (as in (130a) and (130b)); hence, the subject of an infinitive may always be bound be a non-overt operator even if the clause containing this structure occurs in a Case-marked position. If the non-overt operator heads a Case-marked $\overline{A}$-chain, then it must occur in a Caseless Comp; we would thus expect that such a structure could occur only when governed by a noun or adjective or when it occurs in an adjunct position.

We can attribute the above restriction on the possible landing sites for non-overt operators to the general tendency of

phonologically null elements to avoid bearing Case. As a rule, the only null elements that may bear Case are variables and, possibly, pro.[11] Other non-overt elements are systematically excluded from Case-marked positions. This insight underlies Bouchard's (1984) Principle of Lexicalization which requires that Case-bearing empty categories be spelled-out phonologically. In order to make the Principle of Lexicalization work, Bouchard must assume that wh-elements, for example, carry their Case with them when they move to Comp. This is because a Case-marked variable, under his assumptions, would have to be phonologically realized by virtue of its Case feature. This assumption is quite

---

11 The ability of pro to bear Case is, in my estimation, controversial. Assume, for a moment, that pro is restricted to the subject position of Null Subject languages and the A-position bound by clitics. In the former case, following Rizzi (1982), we might say that nominative Case is assigned optionally; in the latter case, following Aoun (1979) and Borer (1984), we might say that the clitic absorbs the Case normally assigned to the A-position occupied by pro. In these cases, it may plausibly be argued that pro does not bear Case. This leaves the problem of the empty category found in A-positions in Chinese and Japanese (for discussion, see Huang, 1982 and Huang, 1984); this empty category is not an A-anaphor, not bound by a clitic and, apparently, not A-bar-bound. In short, it appears to have the properties of pro, except that it appears in a Case-marked position (see, also, the discussion at the beginning of chapter 2). One hypothesis is that it is another instance of a variable bound by a non-overt operator (see Huang, 1984). The question then arises as to what provides this non-overt operator with a range, since this "pro" may occur in the root clause, for example. A possible response (from Huang, 1984) is that the non-overt operator may be identified with an abstract topic position. the difference between Chinese, on the one hand, and English, on the other, is that Chinese allows non-overt operators to be identified by an abstract topic position while English does not. Many details of such a proposal remain to be worked out, a problem that I will leave to future research.

plausible given a Case Filter that applies to all phonologically overt NPs, regardless of their position; a wh-element might well have to carry its Case with it to Comp in order to pass the Case Filter. Suppose, following Safir (1985), that the Case Filter is relativized to apply to A-positions; if this is so, then there is no need for an overt wh-element to bear Case when it occurs in an $\bar{A}$-position in order to pass the Case Filter. Adopting this modification of the Case Filter along with the assumption that variables may have a Case feature (despite the Principle of Lexicalization), we can assume that elements moved to $\bar{A}$-positions are Caseless, although they may head a Case-marked $\bar{A}$-chain.

Our generalization is that a non-overt operator which heads a Case-marked $\bar{A}$-chain may not appear in a Case-marked Comp. Given that the non-overt operator is phonologically null, by definition, the intuitive reason behind this restriction is that null elements avoid Case. The pairing of a Case-marked Comp with a Case-marked $\bar{A}$-chain is simply more than a phonologically null element can tolerate; in this environment, the non-overt operator must be lexicalized, as we have seen (examples (122-124), above). This fact sheds some light on another problem which we noted above.

We have been assuming, throughout the above discussion, that a non-overt operator in a Case-marked position obligatorily moves to Comp in the syntax, taking precedence over a non-overt operator in the subject position of an infinitive. This

assumption accounted for the interpretation of examples like:

(131) John_i is [$_{\text{AdjP}}$ easy [$_{\bar{S}}$ Op_i [$_S$ [e] to please $\underline{t}_i$]]]

where John obligatorily defines the range of the non-overt operator in Comp under predication. We must, however, account for why the following S Structure is impossible:

(132) John_i is [$_{\text{AdjP}}$ easy [$_{\bar{S}}$ Op_i [$_S$ $\underline{t}_i$ to please [e]]]]

In example (132), John again defines the range of the non-overt operator in Comp. The difference between (131) and (132) is that the non-overt operator in Comp binds an empty category in the subject position of the complement clause. In the mapping of the S Structure in (132) to LF, we are assuming that the syntactically unrealized dative argument of easy is induced and bound by a universal quantifier. Furthermore, the empty element in the object position of the complement clause would undergo QR and be interpreted as a non-overt operator:

(133) John_i is [$_{\text{AdjP}}$ (Ax) [$_{\text{AdjP}}$ easy $\underline{x}$ [$_{\bar{S}}$ Op_i [$_S$ [$_{\bar{S}}$ Op_x $\underline{t}_i$ to please $\underline{t}_x$]]]

Note that the new non-overt operator in (133) is an $\bar{A}$-anaphor and must receive a range from some $\bar{A}$-binder. The only possible $\bar{A}$-binder is the universal quantifier induced from the argument structure of easy. As shown in (133), the non-overt operator takes on the index of the universal quantifier. Presumably, such a structure would have the sensible interpretation that it is easy for John to please people; hence, the structure in (133)

cannot be ruled out purely on the basis of having an anomalous interpretation.

Thus, non-overt operators show an "anti-superiority" effect; that is, a non-overt operator in a non-subject position systematically moves to Comp in preference over a non-overt operator in subject position.[12] Notice that the subject position in the relevant examples does not receive nominative Case since it is not governed by a tensed Infl. The non-subject positions, on the other hand, are all environments of Case assignment (i.e., the non-overt operator occurs as object of a verb or a

---

12 Superiority effects are generally taken as following from some version of the ECP. The effects are generally noted with wh-elements that move to Comp; the interaction of this fact with Comp indexing derives the appropriate facts. We are assuming that non-overt operators need not move to Comp and, in the mapping to LF, may be adjoined to some maximal projection. Hence, we would not expect the ECP to play a role in this object/subject asymmetry.

May (1984) rules out LF representations like that in (133) on the basis of a path theoretic version of the ECP. The non-overt operator that binds the variable in object position intervenes between the non-overt operator in Comp and the subject position. This form of the ECP derives facts like the narrow scope on <u>everything</u> in:

(i) who$_i$ <u>t</u> bought everything for Max?

In (i), <u>everything</u> must adjoin to VP in order to avoid a violation of the ECP. Suppose we adjoin the non-overt operator to VP in (133). This will not help us rule out the representation in (133) since this non-overt operator still does not have an accessible SUBJECT in the complement clause. Notice that coindexing the non-overt operator adjoined to VP with the non-overt operator in Comp is impossible since the former non-overt operator would be A-bound by the variable in subject position; hence, the non-overt operator would still have the superordinate S as its governing category and our problem with how to rule out (133) remains.

preposition at D Structure). We have already seen, however, that non-overt operators are resistant to Case assignment; in general, the only empty category that may bear Case is a variable (see the discussion on pro in footnote 11). The assumption throughout this work has been that Case assignment is a property of S Structure so the reason for the anti-superiority effect of non-overt operators is now apparent: An empty category that cannot be interpreted as a variable (for example, the empty category that will be interpreted as a non-overt operator) may not directly bear Case; if such an element appears in a position that will receive Case, it must evacuate that position in order to avoid receiving the Case feature. The S Structure in (134) is, therefore, not well-formed:

(134) $\text{John}_i \text{ is } [_{\text{AdjP}} \text{ easy } [_{\bar{S}} \text{ Op}_i \; [_{\bar{S}} \; \underline{t}_i \text{ to please [e]}]]]$

The empty category in the object position of the complement clause cannot be interpreted as a variable since it is not $\bar{A}$-bound. Since it is not a variable, this empty category may not bear a Case feature. It is, however, in a position of Case assignment. As a result, the S Structure representation in (134) cannot be grammatical.

Before exploring this analysis of control any further, let us take a moment to summarize the approach advocated in this section. We began with the assumption that PRO, per se, does not exist. The element that has been taken to be PRO is, from this perspective, a pair consisting of a non-overt operator and the

variable which it binds. We have adopted the standard distinction between obligatory control and non-obligatory ("arbitrary") control. In the former case, a superordinate head subcategorizes for a Comp containing a non-overt operator, in much the same way that a head may subcategorize for a [+WH] Comp. The non-overt operator in obligatory control structures is assigned a range under predication which is lexically specified by the head; in the case of adjunct clauses, the non-overt operator is assigned a range under predication with the subject of S. In both cases, the control is obligatory, although it is lexically specified only in the former instance.

Arbitrary control has been reanalyzed as a case of binding at LF; the non-overt operator may remain in situ in the syntax (under certain circumstances) and undergo QR in the mapping to LF. In this instance, the empty element is analyzed as an operator without a range; in order for the representation to be well-formed, the operator must be $\overline{A}$-bound. Only then may the non-overt operator be interpreted as having a range. One result of this analysis of arbitrary control is that we must assume that LF representations are richer than has been standardly assumed in previous works. In particular, syntactically unrealized arguments (by this I mean slots in a head's thematic grid which have not been associated with an argument position in the syntax) may be realized in LF representations. These newly realized argument positions are operator bound at LF; the nature of the

operator that binds these positions is, presumably, a lexical property of individual heads. Thus, a universal quantifier is specified by <u>easy</u> in:

(135)a. John is $[_{AdjP_i}$ easy $[_{\bar{S}}$ Op$_i$ $[_S$ [e] to please]]]

b. John is $[_{AdjP_i}$ (Ax) $[_{AdjP}$ easy $[_{\bar{S}}$ Op$_i$ $[_S$ Op$_x$ $[_S$ $\underline{t}_x$ to please $\underline{t}_i$]]]]]

while an existential operator is involved in:

(136)a. John $[_{VP_i}$ bought the book $[_{\bar{S}}$ Op$_j$ $[_S$ [e] to read $\underline{t}_j$ to the children]]]

b. John $[_{VP_i}$ (Ex) $[_{VP}$ bought the book $\underline{x}_j$ $[_{\bar{S}}$ Op$_j$ $[_S$ Op$_x$ $[_S$ $\underline{t}_x$ to read $\underline{t}_j$ to the children]]]]]

There will be a more extended discussion of arbitrary control and properties of LF representations in the next section.

The surface distribution of non-overt operators is regulated by Case theory. Non-overt operators may not appear in a Case-marked position at S Structure. Hence, a non-overt operator in a non-subject position must move to Comp in the syntax while a non-overt operator in the subject position of a non-finite clause may remain in situ until LF. Furthermore, a non-overt operator which heads an $\bar{A}$-chain bearing Case may not land in a Comp which receives Case. These assumptions accounted for the distribution of tough constructions, infinitival relatives, parasitic gap constructions, purposive clauses, environments of obligatory and optional control, and certain sentential predicates.

Finally, the analysis of control proposed in this chapter

has unified the type of binding found in all of the above constructions. In every case, a non-overt operator, which must be identified in order for the structure to receive an interpretation, is involved.

5.3 LF representations and the interpretation of arbitrary control

The LF representations we have been assuming so far in this chapter are more articulated than the LF representations assumed in many of the standard works of generative grammar (see, for example, May, 1977). It has generally been assumed that LF is the level of representation at which scope relationships are disambiguated. The term "scope relationships" is intended to cover the relationships between quantified NPs, negation, and any other operators (e.g., adverbials) which may contribute to the interpretation of a sentence. To take a concrete example, the sentence:

(137) Every student read some book.

is taken as ambiguous between the reading where every student read a book, but not necessarily the same book (wide scope on every student, i.e., every student has scope over some book) and the reading where some book is read by every student where each student reads the same book. In the latter case, some book is said to have scope over every student.

In order to disambiguate the readings, we assume that QR may adjoin operators to S (or, to state the rule in its most general form, QR may adjoin operators to any maximal projection,

including S as a special case). We may then define the scope of an operator as the set of nodes which it c-commands (see Reinhart (1983) for some discussion of the interpretation of scope).[13] Returning to the above example, wide scope on every student implies an LF representation like that in (138):

(138) $[_{S}$ every student $[_{i\ S}$ some book $[_{j\ S}$ $\underline{t}_i$ read $\underline{t}_j]]]$

where every student c-commands, and hence has scope over, some book. The LF representation where some book has scope over every student would have the following form:

(139) $[_{S}$ some book $[_{j\ S}$ every student $[_{i\ S}$ $\underline{t}_i$ read $\underline{t}_j]]]$

where some book c-commands every student. The structural differences between (138) and (139) result from differences in the application of QR. In (138), QR first adjoins some book to S and then adjoins every student to the newly created S node. In (139), the order of application has been reversed; QR first adjoins every student to S and then adjoins some book to the newly created S node.

The approach sketched out above is purely syntactically

---

13 For present purposes, I will assume the following definition of c-command:

(i) A node A c-commands a node B if and only if the first branching node dominating A also dominates B.

This definition makes c-command into a reflexive relation. In the examples in (138) and (139), the c-command relation will guarantee that the operators are in an asymmetrical c-command relation.

driven; that is, quantified NPs and other operators are present in the syntax--they must, in fact, be base generated and, therefore, are present at D Structure. S Structure representations feed the mapping to LF and only overt operators present at S Structure are assigned scope. LF itself is taken to be a level of pure representation of scope.

In this chapter, we have made different assumptions about the role that LF plays in the theory of grammar. Placed in very rough terms, we have taken LF to be a level of complete disambiguation. To say that LF is a level of disambiguation implies, of course, that scope relationships be represented in an unambiguous way. It also has a number of other consequences; in particular, we have required that certain positions that are present in the argument structure of a head, but syntactically unrealized, may come to be represented explicitly at LF.

The idea that syntactically unrealized arguments play a role in the interpretation of sentences is by no means a new one. One example of this sort of phenomenon is control by an implicit argument (see Keyser & Roeper (1984) for an extensive discussion of the following facts).[14] Adopting a standard analysis of passivization (see, for example, Chomsky (1981) or Jaeggli (1984)), we can say that passive morphology blocks assignment of an external thematic role to the subject of S; while this

[14] Control by an implicit argument was, to my knowledge, discovered by Manzini.

thematic role is syntactically suppressed, it is not absent from the argument structure of the verb. The external thematic role may act as a controller of PRO, as the following examples show:

(140)a. the ship was sunk [PRO to win the war]
b. the building was burned [PRO to collect the insurance]
c. the ice was melted [PRO to provide water]

In each of the examples in (140), the implicit agent of the verb is understood as the controller of the PRO in the adjunct clause. Thus, in (140a), the implicit agent who is responsible for the sinking of the ship is the winner of the war; in (140b), the party responsible for the burning of the building is the one who wants to collect the insurance; in (140c), the melter of the ice is also the provider of the water.

The examples in (140) should be contrasted with ergative constructions. Ergative morphology in English destroys the external thematic role completely; meaning that the external thematic role comes to be completely absent from the argument structure of the verb that bears the ergative morphology. In short, there is no implicit argument available for the control of PRO:

(141)a. *the ship sank [PRO to win the war]
b. *the building burned [PRO to collect the insurance]
c. *the ice melted [PRO to provide water]

It is fairly easy to see how the analysis of control in this chapter accounts for the above examples. In the case of the passive examples, the empty element in the subject position need

not move to Comp until the mapping to LF, since the subject position is Caseless and no selectional restrictions are placed on the Comp of the adjunct clause. During the mapping to LF, furthermore, the suppressed external thematic role may be realized syntactically, presumably as a variable in an A-position with an existential quantifier adjoined to the category that would assign the external thematic role. We may, thus, posit the following LF representation for the examples in (261):

(142)a. [$_S$ [e] [$_{Infl'}$ (Ex) was [$_{Infl'}$ [$_{VP}$ sunk the ship] $\underline{x}$] [$_{\bar{S}}$ Op$_x$ [$_S$ $\underline{t}_x$ to win the war]]]]

b. [$_S$ [e] [$_{Infl'}$ (Ex) [$_{Infl'}$ was [$_{VP}$ burned the building] $\underline{x}$] [$_{\bar{S}}$ Op$_x$ [$_S$ $\underline{t}_x$ to collect the insurance]]]]

c. [$_S$ [e] [$_{Infl'}$ (Ex) [$_{Infl'}$ was [$_{VP}$ melted the ice] $\underline{x}$] [$_{\bar{S}}$ Op$_x$ [$_S$ $\underline{t}_x$ to provide water]]]]

(Note that in each of the above cases, the S Structure subject is returned to its D Structure position, leaving a null expletive in subject position.) The non-overt operator must pick up a range from some $\bar{A}$-binder at LF. The only available element is the "implicit" argument, which we are assuming is bound by an existential quantifier; hence, the non-overt operator picks up its range from this existential quantifier and the interpretation of the non-overt operator is obligatorily related to the "implicit argument".

The LF representations in (142) should be contrasted with

those in (143):

(143)a. $*[_{S}\ [e]\ [_{VP}$ sank the ship$]\ [_{\bar{S}}\ Op_{\emptyset}\ [_{S}\ \underline{t}_{\emptyset}$ to win the war$]]]$

b. $*[_{S}\ [e]\ [_{VP}$ burned the building$]\ [_{\bar{S}}\ Op_{\emptyset}\ [_{S}\ \underline{t}_{\emptyset}$ to collect the insurance$]]]$

c. $*[_{S}\ [e]\ [_{VP}$ melted the ice$]\ [_{\bar{S}}\ Op_{\emptyset}\ [_{S}\ \underline{t}_{\emptyset}$ to provide water$]]]$

Recall that the ergative morphology completely eliminates the external thematic role of the verb it attaches to. Thus, in the mapping to LF, none of the examples in (143) have an additional thematic role that may be realized as an argument position. If this is the case, the non-overt operator in the adjunct clause has no potential $\bar{A}$-binder. Therefore, the non-overt operators in (143) have no range at LF and the variables which they bind must be interpreted as expletives. Expletive elements may not bear a thematic role, however; since each of the variables in (143) occupies a θ-position, despite being obligatorily interpreted as expletives, the examples are ruled out.

A cautionary note is in order at this point. We have so far been treating the relationship between implicit arguments and non-overt operators as cases of pure structural binding in the sense that once an operator binding an implicit argument has been spelled out, it may bind a non-overt operator just so long as the newly spelled out operator occupies a structural position which c-commands the non-overt operator. Empirically, it would appear that other factors must be taken into account in order to

adequately capture the complete array of facts. Let us return to the case of ergative constructions in English and consider a wider range of facts:

(144)a. *every naval vessel sank [PRO to win the war]
b. *some building burned [PRO to collect the insurance]
c. *some ice cube melted [PRO to provide water]

The examples in (144), unlike the parallel examples in (141), have quantified NPs in the surface subject position. Therefore, in the mapping to LF, the NPs in surface subject position should undergo QR. Assuming that the empty element in subject position undergoes QR and is analyzed at LF as a non-overt operator, we should have LF representations similar to (I will continue to assume that an empty expletive occupies the matrix subject position in these examples and the true variable is in object position):

(145)a. $[_{S}$ every naval vessel $[_{S_i}$ [e] $[_{VP}$ sank $\underline{t}_i]$ $[_{\bar{S}}$ Op$_i$ $[_{\bar{S}_i}$ $\underline{t}$ to win the war]]]]
b. $[_{S}$ some building $[_{S_i}$ [e] $[_{VP}$ burned $\underline{t}_i]$ $[_{\bar{S}}$ Op$_i$ $[_{\bar{S}_i}$ $\underline{t}$ to collect the insurance]]]]
c. $[_{\bar{S}}$ some ice cube $[_{S_i}$ [e] $[_{VP}$ melted $\underline{t}_i]$ $[_{\bar{S}}$ Op$_i$ $[_{\bar{S}_i}$ $\underline{t}$ to provide water]]]]

Since each of the quantified NPs in (145) occupies an $\bar{A}$-position which c-commands the non-overt operator, it should be possible for the quantified NP to assign the non-overt operator a range. Nevertheless, the above examples are no measurable way better than the examples in (141) which were ruled out on purely

structural grounds.

The examples in (144) could be ruled out on the grounds of selectional restrictions or because the semantic reading is simply anomalous. In all of the examples in (144), the predicates in the adjunct clauses assign an agent thematic role to their subjects and, presumably, have an animate selectional restriction on the subject position. The range assigned by the quantified NPs, however, is strictly inanimate--naval vessels, buildings and ice cubes. Hence, the sets over which the variables in the adjunct clauses will systematically violate the selectional restrictions placed upon the variable position. Similarly, one could argue that the adjuncts in (144) predicate an intension on the part of the quantified NPs in the matrix clauses. Since vessels, buildings and ice cubes are not normally taken to be capable of having intent, the resulting semantic representations are odd.

The above problem, I think, falls under a rather different generalization. Compare the examples in (144) with the following example (attributed to H. Lasnik):

(146) *the ship was sunk [PRO to become a hero]

Since the matrix verb in the example is a passive, it contains a syntactically suppressed external thematic role as part of its argument structure. This thematic role, under our assumptions, may be realized at LF, resulting in the following representation:

(147) $[_{S}$ [e] $[_{Infl'}$ (Ex) $[_{Infl'}$ was $[_{VP}$ sunk the ship] $\underline{x}]$
$[_{\bar{S}}$ Op$_{x}$ $[_{S}$ $\underline{t}_{x}$ to become a hero]]]]

In (147), the existential quantifier should be in a position to identify the range of the non-overt operator. Presumably, the existential quantifier ranges over the set of humans so there should be no problem either with the selectional restrictions placed on the variable in the adjunct clause or with the presupposition of intent on the part of the lexically induced agent of the verb <u>sink</u>. Clearly, an individual may perform an act with the intent of being glorified as a hero:

(148) The sailor$_{i}$ sank the ship $[_{\bar{S}}$ Op$_{i}$ $[_{S}$ $\underline{t}_{i}$ to become a hero]]

The example in (146) is, nevertheless, odd. Such examples indicate that the relationship between implicit arguments and non-overt operators is constrained by factors other than purely structural relations, selectional restrictions and presuppositions of intent.

Tim Stowell (personal communication) has pointed out that examples like that in (146) have been used to support the argument that control of PRO may involve control by a VP or S. In particular, (146) is ungrammatical because the entire S, <u>the ship was sunk</u> controls the PRO; <u>the ship was sunk</u>, however, is not a possible subject for the predicate <u>become a hero</u> and, therefore, the example is filtered out on the basis of selectional restrictions placed on its subject by <u>become a hero</u>.

It should be noted that VPs (or Ss) may be antecedents; consider an example like:

(149) John jumped around, which was a stupid thing to do.

In (149), it would appear that jumped around is the antecedent for which since be a stupid thing to do seems to be predicated of jump around.

Let us assume, for the sake of argument, that the above analysis of (146) is correct and that we can uncover some mechanism that determines whether an NP or a VP (or S) controls a particular PRO. Notice, now, that become does a allow an S in its subject position, depending on the complement following become. Consider the following, for example:

(150) Smoking marijuana became illegal in the 1930s.

Example (150) is, to my ear, perfectly grammatical. Now, compare example (146) with:

(151) *marijuana was smoked [to become illegal in the 1930s]

Example (151) is ungrammatical despite the fact that an S is a possible subject for the predicate become illegal. It would appear, then, that appealing to control by S or VP to explain example (146) will not solve our problem since ungrammaticality remains constant even if an S is a possible subject for the predicate in the adjunct clause. Let us look for another mechanism that will help us to explain these examples.

Whatever constrains the interpretation of arbitrary PRO at LF is over-ridden in environments of obligatory control as shown

in (148). We might suppose, therefore, that LF interpretation is subject to some additional constraint that is not imposed on binding relations in the syntax. We have already seen that this additional constraint does not make reference to structural position, although it does work in tandem with the binding theory so that a potential antecedent for a non-overt operator must obey certain structural conditions. Let us suppose then that this additional constraint on binding at LF makes reference to logical relations between the antecedent operator and the non-overt operator. To be more explicit, let us suppose that this constraint makes reference to thematic roles:

(152) Thematic Compatibility
An element A is a possible antecedent for an element B at LF if the thematic role assigned to the chain headed by A is compatible with the thematic role assigned to the chain headed by B.

We can assume that the examples in (145) and (147) are ruled out because the operator which we have taken as the antecedent for the non-overt operator is not a possible antecedent; if so, then the non-overt operators in these examples simply lack a range at LF and are, as a result, ruled out. If this assumption is correct, then it must be the case that the thematic role assigned to the $\bar{A}$-chain headed by the operator (which binds the implicit argument) is not compatible with the thematic role assigned to the $\bar{A}$-chain headed by the non-overt operator.

Let us return to our examples. The first case is the case of the ergative verbs:

(153)a. $[_{S}$ every naval vessel $[_{i}$ $[_{S}$ [e] $[_{VP}$ sank $\underline{t}_{i}$ $]$ $[_{\bar{S}}$ Op$_{i}$ $[_{S}$ $\underline{t}_{i}$ to win the war]]]]

b. $[_{S}$ some building $[_{i}$ $[_{S}$ [e] $[_{VP}$ burned $\underline{t}_{i}$ $]$ $[_{\bar{S}}$ Op$_{i}$ $[_{S}$ $\underline{t}_{i}$ to collect the insurance]]]]

c. $[_{\bar{S}}$ some ice cube $[_{i}$ $[_{S}$ [e] $[_{VP}$ melted $\underline{t}_{i}$ $]$ $[_{\bar{S}}$ Op$_{i}$ $[_{S}$ $\underline{t}_{i}$ to provide water]]]]

The variable in the object position of the matrix clause is assigned the theme role while the variable in the subject position of the adjunct is assigned the agent role. The fact that the $\bar{A}$-chain headed by the putative antecedent of the non-overt operator eliminates it as a real antecedent since the $\bar{A}$-chain headed by the non-overt operator bears the agent thematic role. Hence, because of the functioning of (152), the non-overt operator in these examples cannot find an antecedent. As a result, it lacks a range and the variable it binds must be treated as non-referential which, we have already seen, is impossible.

The next example is the case in (147) (repeated here):

(154) $[_{S}$ [e] $[_{Infl'}$ (Ex) $[_{Infl'}$ was $[_{VP}$ sunk the ship] $\underline{x}$] $[_{\bar{S}}$ Op$_{x}$ $[_{S}$ $\underline{t}_{x}$ to become a hero]]]]

The existential operator binds an argument that is the realization of the agent thematic role of the verb <u>sink</u>; hence, the existential operator heads an $\bar{A}$-chain that bears an agent thematic role. The non-overt operator, on the other hand, heads an $\bar{A}$-chain that receives whatever thematic role the predicate

become a hero assigns. It seems quite likely that the thematic role assigned to the subject position of the adjunct is distinct from the agent thematic role and it is equally likely that the two thematic roles are not sufficiently compatible for the existential operator to act as a potential antecedent for the non-overt operator. If this is correct, then the non-overt operator lacks an antecedent at LF and its variable must, once again, be treated as a non-referential element.

The constraint in (152) bears some conceptual similarity to the Constraint on Thematic Domains discussed in chapter 3. In both cases, thematic theory narrows the range of a priori possible structural relationships. Notice, however, that (152) must make reference to particular thematic roles and calls a metric (i.e., compatibility) into play. The general tendency in generative grammar has been to make reference to thematic roles in only the most general way. Thus, the θ-criterion requires that all arguments bear a thematic role, but has nothing to say about the specific nature of the thematic roles. Similarly, the constraint developed in chapter 2 makes no reference to particular thematic roles, but, rather, quantifies over thematic roles.

For the Thematic Compatibility constraint to be completely useful, we must, of course, give a coherent definition of the notion of compatibility. We could either give an exhaustive list of the possible thematic roles available to natural languages

plus all the pairs of thematic roles that we will call compatible or we could give a set of propositions from which we could derive the compatibility of any two thematic roles. The former approach would present us with some enormous difficulties. The list of possible thematic roles is of unknown length and some have argued (see, for example, Marantz (1984)) that it could not be given for the simple reason that external thematic roles are defined on the basis of the compositional semantics of the predicate that assigns them. If this is correct, then the list of thematic roles is potentially infinite. Assuming that we could give a finite listing of the set of possible thematic roles, we would then have to provide a listing of all the possible combinations of thematic roles taken two at a time. This final list of compatible thematic roles would, in all likelihood, be quite large and, in the final analysis, taxonomic in the sense that it provides no explanation as to why any two thematic roles are compatible. I will, therefore, not pursue this line of analysis.

The idea of deriving thematic compatibility from some set of principles seems the more promising approach since it circumvents the combinatorial difficulties of the previous method and it is likely to deliver a more satisfying account of the phenomenon. Let us continue with the assumption that LF is not a level of semantic interpretation but is, rather, a level of disambiguation which is then mapped to a semantic representation. In the spirit of Keenan & Faltz (1984) and Keenan & Timberlake (forthcoming), I will assume that part of the task of semantic theory is to

provide an account of predicate formation, where I assume that a predicate is a semantic unit that maps some argument onto a value. We can take a value to be a truth value, a set of individuals, or any other entity defined in the semantic system (see Keenan & Faltz (1984) for discussion).

Notice that in the previous paragraph, I referred to predicates as "semantic units" as opposed to "syntactic units". In general, the relationship between syntactic constituency and predicate structure is so close as to be virtually isomorphic; for example, in a sentence like:

(155) [[John] [tossed the book to Bill]]]
S NP VP

the VP can be treated as mapping directly onto a one-place predicate that takes John as its argument to give a proposition (which, in turn, maps onto a truth value depending on whether John is a member of the set of things that tossed the book to Bill). The argument structure of a head (i.e., the set of thematic roles associated with the head) may be viewed as a set of instructions to the syntax on how to form a predicate/argument structures.

The mapping between syntactic constituents and the semantic units that correspond to predicates is not necessarily so simple as described in the preceding paragraph. The role of adjuncts in this approach is to form new predicates from those predicates that can be mapped by the syntax directly from the argument structure of a head. To take a particularly extreme example of

this approach, consider the case of purposive clauses, as in:

(156) Bill bought the old car $[_{\bar{S}}$ Op$_i$ $[_S$ [e] to fool around with $\underline{t}_i$]]

Keenan (personal communication) treats the predicate/argument structure of the example in (156) as:

(157) Bill [bought to fool around with] the old car

where bought to fool around with is a two-place predicate that takes the old car as argument to yield the one-place predicate bought to fool around with the old car. If this is correct, then the relationship between syntactic constituency and the formation of predicates in the semantics may be relatively indirect.[15] Notice that the non-overt operators act as an interface between the syntactic structure and the more remote level of semantic representation in that they keep an account of the relationship between the arguments (in the syntactic sense) of syntactic elements and the arguments (in the semantic sense) of the predicates.

---

15 Keenan (personal communication) and Stowell (personal communication) have advocated a similar treatment of small clause structures. Namely, the predicate of the small clause may form a complex predicate with the verb in the matrix clause. Thus, an example like (i) would, at some level, be mapped onto a representation like (ii):

(i) John considers [Bill ugly]

(ii) John [considers ugly] Bill

Again, the relationship between syntactic constituency and predicate structure is fairly remote in these examples.

Underlying this approach is the attempt to give an account of the formation of possible predicates so that all and only the possible predicates of a natural language would be formed. In order to give a complete account of this phenomenon, one must be able to describe the set in which such predicates take their denotation and, related to this, the type of denotation that the arguments of these predicates must have. While I will not pretend to have such a theory of possible predicates, it is possible to make some observations. Let us return to the following set of contrasts:

(158)a. The sailor$_i$ sank the ship [$_{\bar{S}}$ Op$_i$ [$_S$ $\underline{t}_i$ to become a hero]]

b. *[$_S$ [e] [$_{Infl'}$ (Ex) [$_{Infl'}$ was [$_{VP}$ sunk the ship] $\underline{x}$] [$_{\bar{S}}$ Op$_x$ [$_S$ $\underline{t}_x$ to become a hero]]]]

c. [$_S$ [e] [$_{Infl'}$ (Ex) [$_{Infl'}$ was [$_{VP}$ sunk the ship] $\underline{x}$] [$_{\bar{S}}$ Op$_x$ [$_S$ $\underline{t}_x$ to win the war]]]]

The contrasts in (158) would seem to be based both on the referential properties of the element that the complex one-place predicate may take as an argument and on the denotations of the complex one-place predicates. On the one hand, the subject of (158a) takes its denotation in the denotation set of definite descriptions while the "subjects" (arguments that receive the external thematic role of <u>sink</u>) in (158b) and (158c) are examples of unrestricted existential quantification. We also have a contrast between (158a) and (158b), on the one hand, and (158c),

on the other. The former examples involve the predicate "become y" which involves a relation between possible worlds (or points in time) while the latter example involves the predicate "win the war". We may, therefore, hypothesize that higher order one-place predicates such as sink the ship to become a hero may not take an element under unrestricted quantification as an argument. The somewhat simpler predicate sink the ship to win the war may, however, take an element under unrestricted quantification as its argument.

The approach outlined above would maintain our theory of control subject to filtering by an interpretive component that forms semantic predicates from complex syntactic structures. This latter component involves a theory of the denotation sets for predicates and arguments (see Keenan & Faltz (1984) and Keenan & Timberlake (forthcoming) for some discussion). We will, therefore, allow the syntax to generate examples like (158b) since such structures are subject to filtering by the interpretive component.

We have so far investigated some of the ramifications of realizing syntactically suppressed arguments at LF and the interrelationship between these implicit arguments and the theory of control. Let us now turn to some further examples of elements that are present at LF but absent from the syntax and explore how such elements interact with the theory of control.

One suggestive piece of evidence for the presence of operators at LF which are not, strictly speaking, visible in the

syntax comes from the work of Heim (1982) in which she investigated the interpretation of indefinite NPs. The basic form of the puzzle is exemplified by the following contrast:

(159)a. A man won a prize.
b. If a man wins a prize, he displays it in his trophy case.

The sentence in (159a) contains two indefinite NPs, a man and a prize. Both of these NPs apparently have the force of existential quantification; that is the following seems to be a fair paraphrase of (159a):

(160) (for some x : x a man (for some y : y a prize (x won y)))

The sentence in (159b) also contain two indefinite NPs, a man and a prize; in contrast to (159a), the indefinite NPs in (159b) do not seem to have the force of existential quantification but, rather, universal quantification. Thus, (161) seems to be a close paraphrase of (159b):

(161) (for any x : x a man (for any y : y a prize (if x wins y then x displays y in x's trophy case)))

Since the indefinite NPs in (159a) are formally the same as the indefinite NPs in (159b)--in both examples the NPs have an indefinite article--how are we to account for the drastic difference in the interpretation of the indefinite NPs in the two examples?

One popular approach to this problem has been to give a number of rules for transforming existential quantification to universal quantification in certain environments (see Heim (1982)

for details). Heim argues quite persuasively that such a system of rules can be little more than taxonomic and that they obscure the correct generalization completely (see Heim (1982) for discussion). Her approach to the puzzle is quite different. Following the work of Lewis (1975) she develops a theory of unselective quantifiers.

Lewis (1975) notes sentences like the following (from Heim (1982)):

(162)a. If a man owns a donkey he always beats it.
b. In most cases, if a table has lasted for 50 years, it will last for another 50.
c. Sometimes, if a cat falls from the fifth floor, it survives.
d. If a person falls from the fifth floor, he or she will very rarely survive.

Lewis treats the adverbs in the above sentences as quantifiers ~~which have the~~ property of binding an unlimited number of variables simultaneously, unlike the more familiar "selective" quantifiers (e.g., <u>some</u>, <u>all</u>, etc.). These "unselective" quantifiers have interpretations like the following:[16]

(163)a. "always(<u>A</u>,<u>B</u>)" is true if and only if every assignment to the free variables in <u>A</u> which satisfies <u>A</u> also satisfies <u>B</u>.

b. "in most cases(<u>A</u>,<u>B</u>)" is true if and only if most assignments to the free variables in <u>A</u> which satisfy <u>A</u> also satisfy <u>B</u>.

---

[16] In the following interpretations, <u>A</u> and <u>B</u> are variables quantifying over propositions. An assignment to free variables in a proposition <u>satisfy</u> that proposition if the assignment yields the value "true" to the interpretation of that proposition.

c. "sometimes(A,B)" is true if and only if some assignments to the free variables in A which satisfy A also satisfy B.

d. "very rarely(A,B)" is true if and only if very few assignments to the free variables in A which satisfy A also satisfy B.

Heim argues that we will derive the correct interpretation for examples like those in (159) and (162) if the indefinite NPs are not taken as having quantificational force but, rather, if they contribute nothing but their status as variables to the interpretation of the sentences they occur in. Thus, she argues that an example like (162a) has a representation like (I have simplified her representations somewhat):

(164) (always x : a man, y : a donkey)(if x owns y then x beats y)

Given the interpretation for "always" given in (163a), (164) is equivalent to:

(165) (for every x : x a man (for every y : y a donkey (if x owns y then x beats y)))

The interpretation of (162a) should be contrasted with a sentence like:

(166) A man very rarely kisses a strange woman.

The interpretation of (166) should be true if the kissing relation contains very few pairs consisting of a man and a strange woman, just as the interpretation of very rarely given in (163d) would predict:

(167) (very rarely x : x a man, y : y a strange woman)(x kisses y)

Let us take it as a given, then, that certain operators may

be unselective in that they may bind any (unbound) variable in their scope (i.e., in the set of nodes which they c-command). We may now reconsider the original puzzle of why some indefinites seem to have existential force while other indefinites have universal force:

(168)a. A man won a prize.
b. If a man wins a prize, he displays it in his trophy case.

Heim's hypothesis that indefinites have no inherent quantificational force seems to be counterexemplified by the examples in (168). Unlike the examples in (162), neither of the sentences in (168) contain an overt unselective operator (like always) to bind the variables left by the indefinite NPs. The fact remains, however, that the quantificational force of the indefinite NPs in (168) is not a constant, as pointed out above.

Heim eliminates the counterexample in (168a) by developing a rule of Existential Closure; I will have little to say about this rule here, but see Heim (1982) for a formulation of the rule and extensive discussion. Existential Closure is a rule of sentence grammar that inserts an unselective existential quantifier in LF representations. Application of Existential Closure to the LF representation of (168a) would yield:

(169) (there exists $\underline{x}$: $\underline{x}$ a man, $\underline{y}$: $\underline{y}$ a prize)($\underline{x}$ won $\underline{y}$)

Since the operator inserted by Existential Closure is unselective it may bind as many free variables as are in its scope.[17] Hence,

---

17 We still require some mechanism for attaching the lexical

in the absence of any other unselective operators, indefinite NPs will be interpreted as having existential force.

Existential Closure cannot be responsible for the interpretation of the indefinite NPs in (168b) for the simple reason that these NPs are interpreted as having the force of universal quantification. The binding of an indefinite NP by means of Existential Closure is, of course, a default case since the binding transpires only when no other operator is available to bind the indefinite. By surface appearances, there is no other operator available in (168b) to bind the indefinite NPs. In order to account for examples like (168b), Heim rejects the hypothesis that the conditional is true material implication. Instead, she develops the following two hypotheses: First, conditionals express a form of conditional necessity with the <u>then</u> clause read as under the scope of a necessity operator which is restricted by the <u>if</u> clause; second, the invisible necessity operator is an unselective operator which may bind any number of variables.

The necessity operator involved in the interpretation of conditionals involves quantification over a restricted set of possible worlds. The details of this quantification over

---

content of indefinite NPs as a restriction on the unselective operator. There are many ways to do this, see, for example, Heim's treatment (Heim, 1982). For present purposes I will assume that the indefinites undergo QR and then may undergo a form of Absorption (see Higginbotham & May, 1981).

possible worlds need not concern us here (see Heim (1982) for the formal definition of this operator); it suffices to observe that this quantification will give the indefinite NPs in (168b) the force of universal quantification. Example (168b) will be associated with the following LF representation (I notate the necessity operator with "N"):

(170) (N x : x a man, y : y a prize)(if x wins y then x displays y in x's trophy case)

In (170) the necessity operator, being unselective, may bind the indefinite NPs, parallel to the example in (169). The representation in (170) will ultimately be interpreted in a way that requires for any man and any prize, if the man wins the prize in any possible world, then the man puts the prize in his trophy case in that world.

If Heim's treatment of the interpretation of indefinite NPs is correct, then LF representations contain "syntactically invisible" operators that may bind null elements. We have so far seen cases of invisible necessity operators associated with conditionals, and invisible existential operators which are introduced by Existential Closure. We have also seen that certain implicit arguments must be realized and operator bound at LF under the assumption that LF is a level of disambiguation.[18]

---

18 I have been assuming that the operator which binds the realized implicit argument follows as a lexical property of the head that contains the implicit argument in its thematic grid. The degree to which this assumption is necessary is not obvious given the existence of Existential Closure. That is, an implicit argument may be realized in the mapping to LF and bound as a

Under the framework developed in Lewis (1975) and Heim (1982), we have also seen that certain adverbs count as quantifiers at LF. Finally, we may assume, following Jackendoff (1972) and Safir (1985) to name only a few examples, that modals count as operators at LF.

Given the relative richness of the LF representations we are assuming, it should come as no surprise that the various operators listed in the previous paragraph are available for binding non-overt operators. The binding relation between these operators and non-overt operators is constrained by the fact that the binding conditions apply at LF and non-overt operators are treated as pure $\overline{A}$-anaphors by the binding conditions. It follows that a non-overt operator may only be $\overline{A}$-bound by a "syntactically invisible" operator (hereafter, the "induced operator") if that operator is in the governing category of the non-overt operator. As we have seen above, this means that the induced operator must be no farther away than the $\overline{S}$ that is superordinate to the $\overline{S}$

---

result of inserting an operator under Existential Closure. If this is the case, then the lexical head need not specify the type of operator needed to bind the implicit argument.

This approach seems to work for verbs. If one considers cases like the implicit agent of passive participles or implicit benefactives in verbs like buy (see the text, above), the quantification of the implicit argument seems to have existential force. In the case of adjectives, however, existential closure seems to give the wrong result. Thus, it is easy to trick John means it is easy for anyone to trick John (universal quantification) rather than it is easy for someone to trick John (existential quantification). While this may be a property of adjectives, I will argue below that it follows from certain properties of the copula.

containing the non-overt operator.

To make the discussion more concrete, let us turn to some examples. We can begin with an adverb of quantification (see above), as in:

(171)a. It is always easy to annoy John.
b. It is very rarely easy to please John.

If the analysis developed up to this point is correct, then the "arbitrary PROs" in the examples in (171) should differ as to quantificational force. In the mapping to LF, the adverbs of quantification in these examples will be assigned scope; I will assume that these adverbs take matrix scope. Furthermore, the empty category in the subject position of each of the complement clauses will be moved to an $\bar{A}$-position. Finally, we may assume that the implicit argument of the adjective <u>easy</u> may also be realized. The sentences will have the following representations (prior to assignment of indices):

(172)a. $[_{S}$ always $[_{S}$ it is $[_{AdjP}$ easy $[_{\bar{S}}$ Op$_{\emptyset}$ $[_{S}$ $\underline{t}_{\emptyset}$ to annoy John$]]$ $\underset{\emptyset}{\underline{x}}$ $]]]$

b. $[_{S}$ very rarely $[_{S}$ it is $[_{AdjP}$ easy $[_{\bar{S}}$ Op$_{\emptyset}$ $[_{S}$ $\underline{t}_{\emptyset}$ to please John$]]$ $\underset{\emptyset}{\underline{x}}$ $]]]$

Prior to assignment of indices, the non-overt operator, the variable which it binds, and the induced argument of <u>easy</u> have no inherent reference and no range; I have notated this by using a special index, "Ø". Notice, however, that the non-overt operator and the implicit argument of <u>easy</u> are within the scope of the

adverbs of quantification, always and very rarely. These adverbs, furthermore, occupy $\bar{A}$-positions in the governing category of the non-overt operators. Thus, the adverbs of quantification may act as antecedents for the non-overt operators (and for the implicit argument of easy). The only indexation which leads to well-formed LF representations of these examples is:

(173)a. $[_{S}$ always $[_{S_i}$ it is $[_{AdjP}$ easy $[_{\bar{S}}$ Op$_i$ $[_{S}$ $\underline{t}_i$ to annoy John$]]$ $\underline{x}_i]]]$

b. $[_{S}$ very rarely $[_{S}$ it is $[_{AdjP}$ easy $[_{\bar{S}}$ Op$_i$ $[_{S}$ $\underline{t}_i$ to please John$]]$ $\underline{x}_i]]]$

Although the derivations leading to the two LFs in (173) are very similar, the non-overt operator in each of the examples is bound by a different adverb of quantification. A review of the informal interpretations assigned to these adverbs (given in (163)) shows that these adverbs of quantification have different truth conditions. The readings of the examples in (173) should differ accordingly. The elements bound by always in example (173a) should have the force of universal quantification; roughly, every individual should find it easy to annoy John. This is so because the interpretation of always given in (163) makes it equivalent to a universal quantifier. The example in (173b), however, should not have the same quantificational force since the non-overt operator is bound by very rarely; roughly, (173b) should be true only if, at any given time, the predicate

easy to please John is satisfied by very few individuals in domain D. This reading is in accord with the interpretation give very rarely in (163d) and it accurately reflects the quantificational force of the variables in (173b).

Notice that this interpretation of (173b) is somewhat different from the expected interpretation given the reading normally associated with universal quantification; there are two possible ways to assign scope with respect to the adverb of quantification and the putative universal quantifier:

(174)a. (for all x (very rarely (it is easy for x to please John)))
b. (very rarely (for all x (it is easy for x to please John)))

Let us assume that very rarely in the examples of (174) quantifies over points of time only, perhaps with the interpretation that there are very few points of time at which the proposition over which it has scope is true. Example (174b) implies that there are very few points in time at which everyone finds it easy to please John. This interpretation seems to be fairly remote from intuitions about the meaning of the example at hand.

Example (174a) is closer to matching the intuitions; we can paraphrase the reading with "for every x there are very few points of time at which it easy for x to please John." This reading seems, however, to allow for the situation some individual always finds it easy to please John, but for any one individual the points of time at which that individual finds it

easy to please to please John are few and far between; thus, although John is always pleased, he is always being pleased by a different individual. This situation does not seem to match up very well with the meaning of the sentence. Assuming that we could provide an analysis of the interaction between the universal quantifier and the adverb of quantification to eliminate this problem, we would still have to provide an account of why the universal quantifier always seems to take scope over the adverb of quantification in examples like (174a). The account given here of the interaction between the adverb of quantification and the non-overt operator faces no such difficulties since it is precisely the adverb of quantification which provides the range of the non-overt operator. There is no potential scope ambiguity between the two operators.

Let us turn, now, to a phenomenon first discussed in Lebeaux (1984). For purposes of exposition, I will include an adverb of quantification in the example before generalizing the analysis to include examples that do not contain an adverb of quantification. The examples in question involve two non-finite clauses, as in:

(175) $[_{\bar{S}}\ [_{S}$ [e] to be famous]] is seldom $[_{\bar{S}}\ [_{S}$ [e] to be happy]]

Lebeaux points out that the two arbitrary PRCs in the above example must be interpreted as covarying. In other words, the example in (175) may not be taken as meaning that for someone to be famous is seldom for someone else to be happy. We can paraphrase the example in (175) more closely with:

(176) for any individual $\underline{x}$, for $\underline{x}$ to be famous is seldom for $\underline{x}$ to be happy.

Loosely speaking, in performing Fregean substitution on individuals that may satisfy the above example, one must always select the same individual to satisfy the famous property and the happy property.

Let us assume that <u>seldom</u> is an adverb of quantification with an interpretation that is equivalent to the interpretation given to <u>very rarely</u> in (163d):

(177) "seldom($\underline{A}$,$\underline{B}$)" is true if and only if very few assignments to the free variables in $\underline{A}$ which satisfy $\underline{A}$ also satisfy $\underline{B}$.

Given our assumptions that adverbs of quantification are assigned scope and that arbitrary PRO involves movement of non-overt operator to an $\bar{A}$-position in the mapping to LF, (175) will have the following structure prior to indexation:

(178) $[_{S}$ seldom $[_{S}$ $[_{\bar{S}}$ Op $_{\emptyset}$ $[_{S}$ $\underline{t}_{\emptyset}$ to be famous$]]$ is $[_{\bar{S}}$ Op $_{\emptyset}$ $[_{S}$ $\underline{t}_{\emptyset}$ to be happy$]]]]$

The adverb of quantification, <u>seldom</u>, will take scope over the entire proposition and both of the embedded non-finite clauses will have a non-overt operator (with no inherent range or reference) in their Comps. These non-overt operators bind variables in the subject positions of the non-finite clauses. The non-overt operators are taken as pure $\bar{A}$-anaphors and, in order for the variables that they bind to bear a thematic role, they must be assigned a range at LF. Both non-overt operators

have the superordinate $\bar{S}$ as their governing category by assumption.[19] The only well-formed indexation of the above example is:

(179) $[_{S}$ seldom $[_{\bar{S}_i}$ $[_{S_i}$ Op $[_{S_i}$ t to be famous]] is $[_{\bar{S}_i}$ Op $[_{S_i}$ t to be happy]]]]

Roughly speaking, (179) should be true only if there are few individuals who are both famous and happy. This follows from the interpretation of seldom and the fact that seldom binds the non-overt operators in the Comps of the embedded clauses.

It should be noted that adverbs of quantification may apparently supersede the operator induced to bind an implicit argument. Consider an example like:

(180) In most cases, ships were sunk to annoy hostile countries.

In (180), the adverb of quantification, in most cases (see the interpretation in (163b), may bind the suppressed external argument of sink and the non-overt operator in the Comp of the adjunct clause:

---

19 Due to the effects of the i-within-i condition, the non-overt operator in the Comp of the clause in subject position does not have Infl as an accessible SUBJECT. Notice that there is a potential binder in a non-argument position that does not yield a violation of the i-within-i condition under coindexation with the non-overt operator. If we assume that accessible SUBJECT means a potential binder for a particular element, then we can say that seldom acts as an accessible SUBJECT for the non-overt operator in the Comp of the sentential subject. Hence, this non-overt operator must find its antecedent in the superordinate $\bar{S}$ in order to be well-formed.

(181) [in most cases [[e] [were [[sunk
S i S Infl´ VP VP
ships] $\underline{x}_i$] [$\underline{\ }_S$ Op$_i$ [$\underline{t}_i$ to annoy hostile
countries]]]]]

Example (181) is true if there are many individuals of whom predicate <u>sink</u> <u>ships</u> <u>to</u> <u>annoy</u> <u>hostile</u> <u>countries</u> holds. This seems to be an intuitively correct interpretation of (180), although, unlike other cases we have seen of implicit arguments, it does not involve existential quantification.

We have so far been able to give a fairly simple account of the interaction between adverbs of quantification and non-overt operators in fixing the interpretation of structures involving "arbitrary control". We have also considered how a number of "syntactically invisible" operators contribute to LF representations. As a rule of thumb, we have proposed that syntactically suppressed elements of a head´s thematic grid be realized in the mapping to LF and that these newly realized arguments be bound by operators at LF. The nature of these "induced" operators bears some investigation; in some cases, particularly with verbs, the implicit argument seems to be bound by an existential quantifier, although this existential operator may be superseded by an adverb of quantification, as shown in (182d):

(182)a. [John [(Ex) [bought the book $\underline{x}$] [$\underline{\ }_S$ Op$_i$
S VP VP
[Op$_S$ [$\underline{t}_x$ to read $\underline{t}_i$ to the children]]]]]
S x S x

b. [[e] [(Ex) [was [sunk the ship] $\underline{x}$]
S Infl´ Infl´ VP
[$\underline{\ }_S$ Op$_x$ [$\underline{t}_x$ to win the war]]]]

c. [in most cases [[e] [were [[sunk
$_S$ $_i$ $_S$ $_{Infl'}$ $_{VP}$ $_{VP}$
ships] $\underline{x}_i$] [$_{\bar{S}}$ Op$_i$ [$_S$ $\underline{t}_i$ to annoy hostile
countries]]]]]

We have also seen that implicit arguments of adjectives are often bound by universal quantifiers, although the universal quantifier may again be superseded by an adverb of quantification:

(183)a. [(Ax) [it is [easy [$_{\bar{S}}$ Op$_x$ [$_S$ $\underline{t}_x$ to please
$_S$ $_S$ $_{AdjP}$
John]] $\underline{x}$]]]

b. [seldom [it is [fun [$_{\bar{S}}$ Op$_i$ [$_S$ $\underline{t}_i$ to annoy
$_S$ $_i$ $_S$ $_{AdjP}$
a gorilla]] $\underline{x}$]]]

It would seem that constructions involving predicative <u>be</u> often involve a sort of universal quantification which may be over-ridden by an adverb of quantification. This becomes clear if we reconsider the example in (175):

(184)a. [[e] to be famous] is [[e] to be happy]
b. [[e] to be famous] is seldom [[e] to be happy]

The examples in (184) will have LF representations similar to those in (185):

(185)a. [(Ai) [[$_{\bar{S}}$ Op$_i$ [$_S$ $\underline{t}_i$ to be famous]] is [$_{\bar{S}}$ Op$_i$
$_S$ $_S$ $_S$ $_i$ $_S$ $_i$
[$_S$ $\underline{t}_i$ to be happy]]]]

b. [seldom [[$_{\bar{S}}$ Op$_i$ [$_S$ $\underline{t}_i$ to be famous]] is [$_{\bar{S}}$ Op$_i$
$_S$ $_i$ $_S$ $_S$ $_i$ $_S$ $_i$
[$_S$ $\underline{t}_i$ to be happy]]]]

In the absence of an adverb of quantification, example (184a) seems to require that all (or almost all) individuals who are famous are also happy. Thus, the class of examples illustrated

above shares the property that, in the absence of other unselective operators, its non-overt arguments are bound by an apparently universal quantifier with the predicate adjective examples.

Heim (1982) points out the similarity in meaning between the following examples:

(186)a. If a cat has been exposed to 2,4-D it goes blind.
b. A cat that has been exposed to 2,4-D goes blind.

Example (186a) is another example of the invisible necessity operator associated with conditionals. Example (186b), while not a conditional in the sense that example (186a) is, has a generic NP as its subject and seems to be a close paraphrase of example (186a). It would appear, then, that the invisible necessity operator is present in both examples:

(187)a. (N $\underline{x}$: $\underline{x}$ a cat)(if $\underline{x}$ has been exposed to 2,4-D then $\underline{x}$ goes blind)
b. (N $\underline{x}$:$\underline{x}$ a cat that has been exposed to 2,4-D)($\underline{x}$ goes blind)

Heim points out, however, that conflating the operator in (187a) with the operator in (187b) is a very rough cut indeed. The invisible necessity operator has the force of universal quantification while the quantification found with generics is somewhat weaker. To illustrate this, consider the following:

(188) A human is bipedal.

Following Heim, we will assume that bipedalness is a property that holds of an individual by virtue of its membership in the class of humans. Example (188) is not falsified if I produce a

monopedal human, however.[20] Let us suppose that the operator in these cases ranges over a set of worlds that are ordered in terms of some ideal of normality (see Heim (1982) for some discussion) and that a proposition "An X is Y" is true if some significantly large subset of the set of individuals in the set X have the property Y. Thus, this operator, call it "G", will bear some similarity to the invisible necessity operator, its quantificational force will be somewhat weaker.

The question now becomes one of explaining how it is that this operator is induced from S Structure representations in the mapping from S Structure to LF. We have grounds for believing that this operator is generally present in constructions involving predicative <u>be</u> (see Safir (1985) for some discussion of predicative <u>be</u>) insofar as generics (like example (188)), constructions involving the predication of an adjective (or some other predicate, see examples like those in (183)), and constructions involving "linked" non-overt operators (as in (185)) all seem to involve this form of quantification. We have also seen that adverbs of quantification can modify the force of this quantification (for example the contrast exemplified by (185a) and (185b)). We can assume, then, that this operator is induced by making reference to properties of Infl--the tense of

---

20 I assume that (188) is false in a world where it can be demonstrated that the bipedal property is distributed more or less randomly over the class of humans; for example, in a world where it is equally likely that a human is bipedal or monopedal.

the Infl and possibly the presence of predicative be--along with any adverbial modification of Infl--adverbs of quantification like seldom, often, or never.

Having identified a plausible source for the induced operator found in many predicative sentences, we can go on to consider further evidence that control involves a non-overt operator that must be $\bar{A}$-bound in its governing category. Consider the following contrast, first noted in Lebeaux (1984):

(189)a. [[e] having relatives in the Old World] makes [[[e] winning the West] difficult]
b. [[e] having relatives in the Old World] makes [[stories about [[e] winning the West] implausible]

Let us assume that the Infl in the examples of (189) is such that we can induce the operator "G". The sentence in (189) has a sentential subject with an empty category in its subject position. In the mapping to LF, this empty category may be moved to an $\bar{A}$-position and identified as a non-overt operator. Given that we have induced G in the matrix clause, this non-overt operator will have the matrix clause as its governing category. Furthermore, a small clause occupies the object position of make in (189a). This small clause has a sentential subject which, like the matrix subject, has an empty category in its subject position; this empty category will move to an $\bar{A}$-position in the mapping to LF and also be identified as a non-overt operator. Since the small clause does not contain a SUBJECT accessible to this non-overt operator, its governing category will be the matrix clause. The following approximates the LF representation

of example (189a):

(190) $[_{S_i}\ G\ [_S\ [Op_i\ [\underline{t}_i$ having relatives in the Old World]]
makes $[[Op_i\ [\underline{t}_i$ winning the West]] difficult]]]

The non-overt operators in (190) are identified in their governing category by the operator, G, as required by condition A of the $\bar{A}$-system of binding and by the requirement that operators assign a range to they variables which they bind. Notice that we have derived the linked interpretation of the non-overt operators as a consequence of these requirements.

While (189a) is yet another case of the linked interpretation of non-overt operators, the non-overt operators in example (189b) do not have the linked interpretation. The people who have relatives in the Old World are not necessarily the people responsible for winning the West. Superficially, this interpretation of (189b) may seem surprising; example (189b) differs in structure from (189a) in that one of the embedded clauses is a nominal complement, but this nominal does not, on the surface, contain a SUBJECT accessible to the non-overt operator. If this is indeed the case then the two non-overt operators in (189b) should have the same governing category and, hence, should have the same $\bar{A}$-antecedent (namely, the induced operator, G). How are we to account for the difference in interpretation between (189a) and (189b)?

Recall that we are allowing syntactically unrealized arguments of heads to be overtly represented at LF. These

arguments may be bound by an already existent unselective operator or by an existential operator (perhaps as a result of Heim's rule of Existential Closure, see above). If we extend this analysis to nouns, then the solution to the puzzle posed by (189b) becomes transparent. In particular, let us suppose, as is reasonable, that some nouns have an articulated argument structure and that these arguments may be realized at LF as bound, in the unmarked case, by an existential operator. The noun, story, for instance, may have a syntactically unrealized possessor so that (189b) is analogous to (191):

(191) [[e] having relatives in the Old World] makes [[someone's stories about [[e] winning the West] implausible]

An operator binding this implicit possessor may be realized in the mapping to LF. If this is so, then the non-overt operator inside the sentential complement to the noun, stories, will have the NP as its governing category rather than the matrix S:

(192) $[_{S_i}\ G\ [_{S_i}\ [Op_i\ [t_i$ having relatives in the Old World]] makes $[[_{NP}\ (Ex)\ [_{NP}\ x$ stories about $[Op_{\emptyset}\ [t_{\emptyset}$ winning the West]]]] implausible]]]

The induced existential operator in the subject position of the small clause in (192) acts as an accessible SUBJECT for the non-overt operator in the clausal complement to the noun, stories; as a result, the non-overt operator must find its antecedent inside the NP. Let us suppose that the non-overt operator in the clausal complement takes on the index of the induced existential operator at LF as shown in (193):

(193) $[_{S_i}\ G\ [\ [Op_i\ [\underline{t}_i$ having relatives in the Old World]]
makes $[[_{NP}\ (Ex)\ [_{NP}\ \underline{x}_x$ stories about $[Op_x\ [\underline{t}_x$ winning
the West]]]] implausible]]]

In (193) the induced subject of the NP headed by <u>stories</u> locally A-binds the non-overt operator. This state of affairs violates the conditions we have placed on the distribution of non-overt operators at LF. Since the induced subject of the NP is not present at S Structure, the clausal complement may not be predicated of this subject in the syntax. The only viable way to identify the non-overt operator is by $\bar{A}$-binding at LF.

While we have been able to account for the fact that the interpretation of the non-overt operator inside the NP in example (192) is not linked with the interpretation of the non-overt operator in the sentential subject of the matrix S, we have, as yet not accounted for what assigns the non-overt operator a range at the level of LF representation. There remains, then, the task of outlining a possible solution to this apparent contradiction.

A possible solution may be found by developing Heim's (1982) rule of Existential Closure, a rule which I have briefly alluded to, above. This rule inserts an unselective existential quantifier in LF representations by adjoining this operator to S. Heim invokes this rule in order to account for the existential force associated with indefinite NPs in sentences like:

(194)a. A man walked into the room.
b. A dog bit a cat.
c. A bureaucrat gave a form to an applicant.

After applying QR to each of the indefinite NPs in (194) and then applying Existential Closure to the result, the examples in (194) will have LF representations like:

(195)a. (Ex $\underline{x}$: $\underline{x}$ a man)($\underline{x}$ walked into the room)
b. (Ex,$\underline{y}$ $\underline{x}$: $\underline{x}$ a dog, $\underline{y}$: $\underline{y}$ a cat)($\underline{x}$ bit $\underline{y}$)
c. (Ex,$\underline{y}$,$\underline{z}$ $\underline{x}$: $\underline{x}$ a bureaucrat, $\underline{y}$: $\underline{y}$ a form, $\underline{z}$:$\underline{z}$ an applicant)($\underline{x}$ give $\underline{y}$ to $\underline{z}$)

The immediate problem posed by Existential Closure is one of constraining its operation. In general, the existential operator will be unable to license unbound gaps in the syntax (i.e., levels of representation prior to and including S Structure) but it may license variables at LF. Thus, the modular character of the theory of grammar that we are assuming here restricts the operation of Existential Closure to the interpretation of indefinites, implicit arguments that have been induced at LF, and the interpretation of "arbitrary PRO". Let us assume that Existential Closure adjoins an operator in one of two environments: a tensed S or a "propositional NP" such as an NP headed by _fact_, _rumor_, _story_, etc. We can assume, in fact, that Existential Closure is responsible for providing a range to the induced subject of _story_ in (310). Since the operator inserted by Existential Closure is unselective, it may bind any number of distinct positions; hence, the non-overt operator may bind the non-overt operator in (310) just so long as the index assigned to the non-overt operator under this binding is disjoint from the index associated with the induced subject of the NP headed by _story_:

(196) $[_{S_i}\ G\ [\ [Op_i\ [t_i$ having relatives in the Old World]]
makes $[[_{NP}\ (Ex,y)\ [_{NP}\ \underline{x}$ stories about $[Op_y\ [t_y$
winning the West]]]] implausible]]]

Thus, the interpretation associated with (196) follows from the interaction of the binding theory, the assumption that implicit arguments are realized in LF and the functioning of the independently motivated rule of Existential Closure. The empty element in the subject position of the non-finite clause is moved to an $\bar{A}$-position by QR and, as a result, is treated as a non-overt operator. It is an $\bar{A}$-anaphor that must be locally $\bar{A}$-bound in order to receive an interpretation. At the same time, the implicit subject of the noun <u>story</u> is realized in the mapping to LF. This realization of the subject of the NP results in the fact that the NP is the governing category for the non-overt operator (thus accounting for the "non-linked" reading, noted above). The realization of the subject of <u>story</u> feeds the rule of Existential Closure which, in turn, provides an $\bar{A}$-antecedent for the non-overt operator. Finally, since the subject of the NP c-commands the non-overt operator, the binding conditions require that the non-overt operator have a distinct index from the subject of the NP; the indexing shown in (196) follows as a natural result. The interpretation of examples like that in (196) provides us with additional evidence that the empty category found in various control structures (in particular, arbitrary control structures) has properties that are identical to properties associated with non-overt operators insofar as the

element in question is an $\bar{A}$-anaphor that must be bound in its governing category.

Before concluding this section, let us consider some examples that are, at least on the surface, more problematic for the analysis of control developed in this chapter. We have been assuming that arbitrary control involves syntactically invisible operators that often result from the realization, in the mapping to LF, of so-called implicit arguments. This analysis gave us a fairly simple account of the "double control" found in many purposive clauses. We have so far ignored cases where all the possible arguments of a verb are realized, as in (197) and (198):

(197)a. John bought a book for Mary.
b. Bill cooked an artichoke for Mary.
c. Bill read a book to the children.

(198)a. John bought Mary a book.
b. Bill cooked Mary an artichoke.
c. Bill read the children a book.

Prima facie, we would not expect the "double control" phenomenon since no implicit argument will be induced in such examples. Nevertheless, in some examples, the "double control" reading is still possible:

(199)a. John bought a book for Mary to read.
b. John bought Mary a book to read.

We can eliminate (199a) as a counterexample since the PP headed by <u>for</u> is not necessarily a complement of the verb <u>buy</u>; we could assign the following structure to (199a):

(200) [John [bought a book] [Op for [Mary to
S VP i $\bar{S}$ i S
read $\underline{t}_i$]]]

where for is a complementizer and Mary is in the subject position of the adjunct clause. If (200) is the S Structure for (199a), then we can assume that the implicit benefactive of buy is induced in the mapping to LF and Existential Closure then applies:

(201) [John [(Ex) [bought a book] x_i] [$_{\bar{S}}$ Op$_i$ for [Mary to read t_i]]]
S VP VP S

At some level of interpretation after LF, Mary could be identified with the argument bound by the existential quantifier, thus accounting for the tendency to identify Mary as the benefactive argument of buy as well as the agent of read. There is some evidence for identifying (200) as the S Structure of (199a). Notating a major intonational break with "||", it appears that a major intonational break may be placed between a book and for Mary but such a break is unnatural between for Mary and to read:

(202)a. John bought a book || for Mary to read
b. ?John bought a book for Mary || to read

If intonation breaks correspond to syntactic constituency, then it is natural that an intonational break may occur between a VP and an $\bar{S}$ (which corresponds to the S Structure in (200)) rather than occurring in the middle of VP (which corresponds to an analysis that treats for Mary as a complement of the verb buy). If this analysis is correct, then examples like (199a) do not contain an "arbitrary PRO" in the subject position of the adjunct

clause and, so, do not constitute a counterexample to the analysis of control structures presented in this chapter.

The example in (199b) is more problematic; in this case the S Structure must be as shown in (203) (irrelevant details omitted):

(203) $[_{S}$ John $[_{VP}$ bought Mary a book $]_{i}$ $[_{\bar{S}}$ Op$_{i}$ $[_{S}$ [e] to read $\underline{t}_{i}$]]]

In (203), the benefactive NP <u>Mary</u> undoubtedly has VP as an ancestor, so the counter-analysis we exploited for (199a) is unavailable to us here. On the face of it, we cannot appeal to the realization, at LF, of an implicit argument that will be bound under Existential Closure; but then the empty category in the subject position of the adjunct clause cannot be bound by a non-overt operator since the non-overt operator will have no $\bar{A}$-antecedent to assign it a range. We would therefore predict that the empty category in the subject position of the adjunct should be treated as an expletive and the entire representation should be ruled out given that an expletive may not occupy a θ-position.

Let us investigate the status of the putative benefactive NP in the representation of (199b). Recall that this NP obeys very strong subject condition effects; that is, the NP in question is an island to extraction. Kayne (1984) has provided persuasive evidence that the benefactive NP in examples like (199b) is a

subconstituent of the subject of a small clause:[21]

(204) [John [bought [Mary a book]] [$_{\bar{S}}$ Op [[e]
S VP sc i $\bar{S}$ i S
to read $\underset{i}{\underline{t}}$]]]

If this analysis of dative shift constructions is correct, then the benefactive NP is not, strictly speaking, a complement of the verb; we might plausibly assume that the benefactive receives its thematic role internally to the small clause. There is some support for this idea insofar as such constructions may easily take a second benefactive; compare the following:

(205)a. John bought Mary a book for Bill.
b. ?John bought a book for Mary for Bill.

This suggests that the verb, buy in this case, still has an implicit benefactive argument which may be realized, and bound under Existential Closure, during the mapping to LF. The LF representation of (199b) may very well be like the following:

---

21 Kayne, in fact, argues that the benefactive NP in such examples is governed by a phonologically null preposition, as in (i):

(i) [bought [[[e] [Mary]] [a book]]]
VP PP P NP NP

This accounts, in his framework, for the impossibility of extracting this benefactive NP:

(ii) *who$_i$ did John buy $\underline{t}_i$ a book

I will have nothing to say here about the properties of this null preposition. What is crucial for our purposes is that the benefactive NP forms a small clause with the direct object and, so, is not in construction with the verb.

(206) [John [(Ex) [bought [Many a book] x]
S VP VP SC i
[Op [Op [t to read t]]]]
S i S x S x i

The "true" benefactive argument of the verb in (206) has been realized and bound in the mapping to LF. This, in turn, makes an antecedent available for the non-overt operator associated with the subject position of the adjunct clause.

It should be noted that the entire problem probably disappears if we adopt the analysis of dative shift found in Stowell (1981). He argues that the "dative shifted" NP is in fact cliticized onto the verb; it is then plausible to treat this cliticized NP as an $\bar{A}$-binder which could plausibly bind the non-overt operator in question at LF. Thus, the analysis of control presented in this chapter seems to be independent of the analysis of dative shift constructions.

The analysis accorded to the counterexamples in (199) is, of course, largely speculative in nature and there are, no doubt, many other potential counterexamples to this analysis. The direction of research is, however, clear; properties of control follow from the interaction of the binding theory with argument structure. In explicating the argument structure of particular lexical items, many, if not all, of the properties of control will follow naturally. From this perspective, there is no independent module of grammar that is concerned exclusively with control, but rather an interaction of binding and argument structure.

5.4 Summary

We began our analysis of control with the puzzle of the "PRO gate." Examination of cases of weak cross-over led to the conclusion that pronominals may not be directly $\bar{A}$-bound. Pronominals may be treated as bound variables just in case they are, in turn, A-bound by a variable. If PRO has a pronominal feature, however, then we must weaken this restriction since PRO may be directly $\bar{A}$-bound at LF:

(207)a. $[\text{PRO}_i \text{ kissing his}_i \text{ mother}]$ annoys every boy$_i$

b. $[_S \text{ every boy}_i \; [_S \; [\text{PRO}_i \text{ kissing his}_i \text{ mother}]$ annoys $\underline{t}_i]]$

This direct $\bar{A}$-binding of PRO by operators at LF also implied that empty categories must have inherent features; if the properties of empty categories are derived solely with reference to the structure in which they occur in the manner discussed in Chomsky (1982), then the PRO in (207b) will be identified as a variable since it occurs in an A-position and is locally $\bar{A}$-bound by <u>every boy</u>. But a variable in the position of the PRO in (207b) is not head governed and, so, violates the ECP. Thus, we must simply stipulate that this empty category is a PRO and put aside the idea that empty categories may be contextually defined.

In order to circumvent these problems, we have replaced PRO with a non-overt operator/variable pair. Notice that the conclusion that a variable occupies the subject position of the embedded S in (207b) is exactly what is predicted by the

contextual definition of empty categories. The difference is that the relationship between every boy and the variable is mediated by a non-overt operator; thus, the ECP is not violated since the variable is governed by the non-overt operator. In this framework, the PRO gate comes as no surprise since non-overt operators are anaphors which must pick up a range from some $\bar{A}$-binder.

The hypothesis that control involves a non-overt operator led us to the conclusion that control is simply another instance of the type of binding found in other constructions which involve a non-overt operator, like tough movement, parasitic gaps, infinitival relatives and so on. The analysis of these latter constructions and obligatory control was as follows: Syntactic movement of an empty element to an $\bar{A}$-position was forced either by subcategorization (obligatory control) or by the fact that the empty element occupies a Case-marked position. In both cases, the distribution is regulated by Case theory. Once having moved to an appropriate $\bar{A}$-position, the empty element is identified as a non-overt operator and must be assigned a range so that the variable which it binds will not be taken as an expletive. In the syntax there are two ways to identify a non-overt operator--predication or $\bar{A}$-binding. Predication identifies the non-overt operator in obligatory control structures while $\bar{A}$-binding identifies the non-overt operator in parasitic gap constructions.

In cases of arbitrary or non-obligatory control, syntactic movement is not forced by either subcategorization or Case

theory. If the empty category remains in situ, however, it will be taken as an expletive element and may not bear a thematic role. If the empty element moves to an $\bar{A}$-position, it will be taken as a non-overt operator. There is only one way to identify a non-overt operator at LF, however--$\bar{A}$-binding. This last assumption led us into a consideration of properties of LF. We concluded that LF is a very rich level of representation where operators that are induced from tense and mood, or operators associated with adverbs of quantification are realized and assigned scope.

Furthermore, we argued that syntactically unrealized arguments may be realized at LF. We thus view LF as a completely unambiguous level where argument structure is realized in its purest form. We can therefore contrast LF with D Structure. In this framework, D Structure is a level where syntactically obligatory thematic positions like "subject" or "direct object" are represented; from our perspective, D Structure is thematically pure with respect to the syntax. LF differs from D Structure in that suppressed arguments are realized and scope is assigned unambiguously to various operators. Thus, LF is the thematically pure level of representation which is second to none. Our assumptions about LF, in turn, provided us with an account of how the realization and binding of implicit arguments interacts with "control". Along the way we found an interaction between "control" and, for example, adverbs of quantification.

In general, the assumption that "control" involves non-overt operators not only allowed us to eliminate control theory as an independent module of grammar, but also led us to a unification of apparently divers types of binding and uncovered a variety of interesting interactions, as was the case with adverbs of quantification. All this provides us with evidence of the robustness of the hypothesis that control falls under the study of non-overt operators. The entire approach, finally, crucially assumes a grammar made up of interactive modules (cf., the relationship between Case theory and binding assumed in this chapter) which apply at different levels of representation (cf., the distinction between obligatory control--which is captured at S Structure--and non-obligatory control--which is captured at LF).

The relatively optimistic picture painted in the previous paragraph should, however, also be a reminder of the amount of work that remains. The framework developed in this chapter, and, indeed, throughout this dissertation, requires a theory of argument structure and predicate formation; this theory, however, is still in its infancy. The conclusion is that argument structure is a fundamental ingredient in explicating the mental representation of syntactic structure. Finally, there is a last puzzle; we have focused our attention in this chapter on the reduction of PRO in the subject position of S to a non-overt operator. We have not looked at the problem of PRO inside an NP, although we argued in chapter 2 that PRO could, in fact, occur in

the Spec of NP. While I will not discuss the reduction of PRO inside an NP to a non-overt operator, some preliminary observations are in order. By our logic, the non-overt operator in the structure shown in (208a) will move to the Spec of NP in order to avoid Case assignment by the preposition, as shown in (208b):

(208)a. ...need [$_{NP}$ a good [$_{N'}$ talking [$_{PP}$ to Op]]]...

b. ...need [$_{NP}$ Op$_i$ a good [$_{N'}$ talking [$_{PP}$ to t_i]]]...

Since the non-overt operator in the above structure does not appear to license a parasitic gap, we assumed that the position in the Spec of NP to which the element moves does not count as an $\bar{A}$-position. Nevertheless, it appears that the non-overt operator may remain in this position and, furthermore, it may be identified by the subject NP. This is quite reminiscent of obligatory control verbs where an obligatory coreference relation is established between an argument of the matrix predicate and the subject of an embedded clause and where the variable left by movement of the non-overt operator inside the embedded clause does not license a parasitic gap. This last fact follows from the condition that the "licensing" variable may not c-command the non-overt operator associated with the parasitic gap. We thus have a clear analogy between obligatory control structures and retroactive nominals.

Drawing on this parallelism between retroactive nominals and obligatory control structures, we may say that retroactive

nominals involve identification of a non-overt operator under lexical specification of the governing verbal and that the non-overt operator does not license a parasitic gap because the "licensing" variable (perhaps the non-overt operator itself) in the Spec of NP will c-command, and thus A-bind, the operator associated with the parasitic gap. The only way in which the element in the Spec of NP will not c-command the non-overt operator associated with the parasitic gap is if the constituent containing this latter non-overt operator is outside of the retroactive nominal. But if this last condition holds, then the non-overt operator inside the retroactive nominal can neither A- nor $\bar{A}$-bind the operator associated with the parasitic gap. Hence, the operator associated with the parasitic gap will not be identified at S Structure, contrary to the licensing conditions discussed in chapter 4. If this argument is correct, then the fact that parasitic gaps cannot be licensed in retroactive nominals should come as no surprise.

The comments in the preceding paragraph make clear what direction research on the distribution of "PRO" inside nominals should take. Many details of the exact nature of the internal structure of nominals remain to be investigated. The preceding discussion implies, for example, that nominals can contain an internal $\bar{A}$-position that may be present to receive a non-overt operator where needed (contrary to our assumptions in chapter 2). I will leave this investigation for a later date.

References

Aoun, J. (1981). The Formal Nature of Anaphoric Relations. MIT PhD Dissertation.

Aoun, J. (1985). A Grammar of Anaphora. The MIT Press, Cambridge, MA.

Aoun, J. (forthcoming). Generalized Binding. Foris Publications, Dordrecht, Holland.

Aoun, J. and R. Clark (1985). "On Non-Overt Operators". Southern California Occasional Papers in Linguistics. University of Southern California, pp. 17-36.

Aoun, J. and D. Sportiche (1983). "On the Formal Theory of Government". The Linguistic Review, 2, pp. 211-36.

Bach, E. (1982). "Purpose Clauses and Control". in Jacobsen, P and G. Pullum (eds.) The Nature of Syntactic Representation. D. Reidel Publishing Company, Dordrecht, Holland. pp. 35-57.

Baker, C. L. (1981). "Learnability and the English Auxiliary System". in Baker, C. L., and J. J. McCarthy (eds.) The Logical Problem of Language Acquisition. The MIT Press, Cambridge, MA. pp. 296-323.

Belletti, A. & L. Rizzi (1981). "The Syntax of ´ne´: Some Theoretical Implications". The Linguistic Review, 1, pp. 117-54.

Bennis, H. (1981). "A Note on Government and Binding". in May, R. and J. Koster (eds.) Levels of Syntactic Representation. Foris Publications, Dordrecht, Holland. pp. 1-8.

Bouchard, D. (1984). On the Content of Empty Categories. Foris Publications, Dordrecht, Holland.

Borer, H. (1984). Parametric Syntax. Foris Publications, Dordrecht, Holland.

Brody, M. (1984). "On Contextual Definitions and the Role of Chains". Linguistic Inquiry, 15.3, pp. 355-80.

Burzio, L. (1980). Intransitive Verbs and Italian Auxiliaries. MIT PhD Dissertation.

Chierchia, G. (1984). Topics in the Syntax and Semantics of Infinitives and Gerunds. PhD Dissertation, University of Massachusetts, Amherst.

Chierchia, G. (1985). "Formal Semantics and the Grammar of Predication". Linguistic Inquiry, 16, pp. 417-443.

Chomsky, N. (1965). Aspects of the Theory of Syntax. The MIT Press, Cambridge, MA.

Chomsky, N. (1970). "Remarks on Nominalization" in Jacobs, R. & R. Rosenbaum (eds.) Readings in English Transformational Grammar. Ginn and Company. Waltham, MA.

Chomsky, N. (1977). "On WH-Movement" in P. Culicover, T. Wasow & A. Akmajian (eds.) Formal Syntax. Academic Press, New York.

Chomsky, N. (1977). "On WH-Movement" in P. Culicover, T. Wasow & A. Akmajian (eds.) Formal Syntax. Academic Press, New York

Chomsky, N. (1980). "On Binding". Linguistic Inquiry, 11, pp. 1-46.

Chomsky, N (1981). Lectures on Government and Binding. Foris Publications, Dordrecht, Holland.

Chomsky, N. (1982). Some Concepts and Consequences of the Theory of Government and Binding. The MIT Press, Cambridge, MA.

Chomsky, N. (1985). "Barriers". ms., MIT.

Chomsky, N. (forthcoming). Knowledge of Language: Its Nature, Origin and Use.

Chomsky, N. & H. Lasnik (1977). "Filters and Control". Linguistic Inquiry, 8, pp. 425-504.

Cinque, G. (1984). "$\bar{A}$-bound pro vs. Variable". ms., Università di Venezia.

Clark, R. (1984). "Control into NP". in Proceedings of the West Coast Conference on Formal Linguistics, 3, pp. 40-47.

Clark, R. (1985). "The Syntactic Nature of Logical Form: Evidence from Toba Batak". Linguistic Inquiry. 16.4, pp. 663-669

Contreras, H. (1984). "A Note on Parasitic Gaps". ms. University of Washington.

Epstein, S. D. (1984). "Quantifier-PRO and the LF Representation of PRO arb". Linguistic Inquiry, 15.3, pp. 499-505.

Evans, G. (1980). "Pronouns". Linguistic Inquiry, 11, pp. 337-62.

Evans, G. (1982). The Varieties of Reference. Clarendon Press, Oxford.

Fiengo, R. (1980). Surface Structure. Harvard University Press, Cambridge, MA.

Finer, D. (1984). The Formal Grammar of Switch-Reference. PhD Dissertation, University of Massachusetts, Amherst.

Geach, P. T. (1962). Reference and Generality. Cornell University Press, Ithaca, NY.

Gueron, J. and R. May (1984). "Extraposition and Logical Form". Linguistic Inquiry, 15.1, pp. 1-31.

Haik, I. (1984). "Indirect Binding". Linguistic Inquiry, 15.2, pp. 185-223.

Hantson, A. (1984). "Towards an Analysis of Retroactive Gerunds" in W. de Geest & Y. Putseys (eds.) Sentential Complementation. Foris Publications, Dordrecht, Holland

Heim, I. (1982). The Semantics of Definite and Indefinite Noun Phrases. PhD Dissertation, University of Massachusetts, Amherst.

Higginbotham, J. (1980). "Pronouns and Bound Variables". Linguistic Inquiry, 11, pp. 679-709.

Higginbotham, J. (1983). "Logical Form, Binding, and Nominals". Linguistic Inquiry, 14.3, pp. 395-420.

Higginbotham, J. and R. May (1981). "Questions, Quantifiers and Crossing". The Linguistic Review, 1.1, pp. 41-80.

Hornstein, N. (1984). Logic as Grammar. The MIT Press, Cambridge, MA.

Hornstein, N. and D. Lightfoot, eds. (1981). Explanation in Linguistics. Longman, New York.

Hornstein, N. and A. Weinberg (1981). "Case Theory and Preposition Stranding". Linguistic Inquiry, 12.1, pp. 55-91.

Horvath, J. (1981). Aspects of Hungarian Syntax and the Theory of Grammar. UCLA PhD Dissertation.

Huang, C.-T.J. (1982). Logical Relations in Chinese and the Theory of Grammar. MIT PhD Dissertation.

Huang. C.-T. J. (1984). "On the Distribution and Reference of Empty Pronouns". Linguistic Inquiry, 15.4, pp. 531-74.

Jackendoff, R. (1972). Semantic Interpretation in Generative Grammar. The MIT Press, Cambridge, MA.

Jackendoff, R. (1977). $\overline{X}$-Syntax: A Study of Phrase Structure. The MIT Press, Cambridge, MA.

Jaeggli, O. (1982). Topics in Romance Syntax. Foris Publications, Dordrecht, Holland.

Jaeggli, O. (1984a). "Passive". ms., MIT & USC.

Jaeggli, O. (1984b). "Subject Extraction and the Null Subject Parameter". Proceedings of NELS 14. pp. 132-53

Jespersen, O. (1940). A Modern English Grammar. E. Munksgaard, Copenhagen

Kayne, R. (1975). French Syntax: the Transformational Cycle. The MIT Press, Cambridge, MA.

Kayne, R. (1984). Connectedness and Binary Branching. Foris Publications, Dordrecht, Holland.

Keenan, E. & L. Faltz (1985). Boolean Semantics for Natural Language. D. Reidel Publishing Company, Dordrecht

Keenan, E. & A. Timberlake (1985). "Predicate Formation Rules in Universal Grammar" to appear in Proceedings of the West Coast Conference on Formal Linguistics.

Keyser, S. J. and T. Roeper (1984). "On the Middle and Ergative Constructions in English". Linguistic Inquiry, 15.3, pp. 381-416.

Koopman, H., and D. Sportiche (1983). "Variables and the Bijection Principle". The Linguistic Review, 2.2.

Lasnik, H. and M. Saito (1984). "On the Nature of Proper Government". Linguistic Inquiry, 15.2, pp. 235-89.

Lebeaux, D. (1984). "Anaphoric Binding and the Definition of PRO". Proceedings of NELS 14. pp. 253-74

Lees, R. B. (1970). The Grammar of English Nominalizations. Mouton And Company. The Hague, Netherlands.

Lewis, D. (1975). "Adverbs of Quantification". in E. Keenan (ed.), Formal Semantics of Natural Language. Cambridge University Press, Cambridge, pp. 3-15.

Longobardi, G. (1985). "Connectedness, Scope and C-Command". Linguistic Inquiry, 16.2, pp. 163-92.

Manzini, M.R. (1983). "On Control and Control Theory". Linguistic Inquiry, 14.3, pp. 421-46.

Marantz, A. (1984). On the Nature of Grammatical Relations. The MIT Press, Cambridge, MA.

May, R. (1977). The Grammar of Quantification. MIT PhD Dissertation.

May, R. (1984). Logical Form: Its Structure and Derivation. to be published by MIT Press.

Montalbetti, M. (1984). After Binding. MIT PhD Dissertation.

Pesetsky, D. (1982a). "Complementizer Trace Phenomena and the Nominative Island Condition". The Linguistic Review, 1.3.

Pesetsky, D. (1982b). Paths and Categories. PhD Dissertation, MIT.

Postal, P. (1974). On Raising. The MIT Press. Cambridge, MA.

Reinhart, T. (1983). Anaphora and Semantic Interpretation. The University of Chicago Press, Chicago.

Riemsdijk, H. van (1978). A Case Study in Syntactic Markedness. Foris Publications, Dordrecht, Holland.

Rizzi, L. (1982). Issues in Italian Syntax. Foris Publications, Dordrecht, Holland.

Ross, J.R. (1967). Constraints on Variables in Syntax. PhD Dissertation, MIT.

Safir, K. (1982). Syntactic Chains and the Definiteness Effect. MIT PhD Dissertation.

Safir, K. (1984). "Worth". ms., Rutgers University.

Safir, K. (1984b). "Multiple Variable Binding". Linguistic Inquiry, 15.4, pp. 603-38.

Safir, K. (1985). Syntactic Chains. Cambridge University Press, London.

Sportiche, D. (1981). "Bounding Nodes in French". The Linguistic Review, 1.2, pp. 219-46.

Sportiche, D. (1983). Structural Invariance and Symmetry in Syntax. MIT PhD Dissertation.

Stowell, T. (1981). Origins of Phrase Structure. MIT PhD Dissertation

Stowell, T. (1983). "Subjects Across Categories". The Linguistic Review, 2, pp. 285-312.

Stowell, T. (1985). "Null Operators and Antecedent Government". ms. UCLA

Wahl, A. (1985). "Two Types of Locality". ms. University of Maryland & USC.

Williams, E. (1980). "Predication". Linguistic Inquiry, 11.1, pp. 203-38.

Williams, E. (1981). "On the Notions ´Lexically Related´ and ´Head of a Word´". Linguistic Inquiry, 12, pp. 245-74.

Williams, E. (1982). "The NP Cycle". Linguistic Inquiry, 13.2, pp. 277-95.

Williams, E. (1984). "Grammatical Relations". Linguistic Inquiry, 15.4, pp. 639-73.