

INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9108138

**Lexically based parsing with application to infinitival control
constructions in English**

Leacock, Claudia, Ph.D.

City University of New York, 1990

Copyright ©1990 by Leacock, Claudia. All rights reserved.

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A

LEXICALLY BASED PARSING
WITH APPLICATION TO INFINITIVAL CONTROL CONSTRUCTIONS IN ENGLISH

by

CLAUDIA LEACOCK

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1990

© 1990
CLAUDIA LEACOCK
All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

September 27, 1990
Date

John Roysne
Chair of Examining Committee

September 27, 1990
Date

Richard K. ...
Executive Officer

Virginia Telle

Steuve Lander
Supervisory Committee

Abstract

LEXICALLY BASED PARSING
WITH APPLICATION TO INFINITIVAL CONTROL CONSTRUCTIONS IN ENGLISH

by

Claudia Leacock

Adviser: Professor John Moyne

The thesis examines a current model of grammar within Chomsky's Move- α framework in relation to its implications for parsing natural language. It is demonstrated that Chomsky's arguments for the movement of elements are not conclusive. However, from the computational point of view, the model can be greatly improved by eliminating movement operations entirely.

I explore how a lexicon must be structured for a lexically based parser to operate, illustrating the proposed parsing mechanisms with an analysis of infinitive constructions in English. It is shown that the grammatical properties of infinitives can be accounted for by assuming that control and complement information are encoded lexically, and that these lexical structures are available to the parser.

Acknowledgements

I thank the members of my committee at the CUNY Graduate Center: John Moyne, chair, Richard Kayne and Virginia Teller. I especially thank Terry Langendoen for his advice and encouragement, and for remaining on my committee as an outside advisor when he moved from CUNY to the University of Arizona. I also thank my colleagues at the IBM T.J. Watson Research Center where I have been employed on a Student Thesis Agreement: Warren Plath, David Johnson, Paul Postal, Fred Damerau, and Edwin Battistella, who was a visiting scientist last year. I also thank Alice Greenwood, Dana McDaniel, and Yael Ravin. Of course, I owe many many thanks to my husband, Ray Dougherty. I also thank our children, Sean, Tess and Tracy for occasionally allowing me to work.

Finally, I dedicate this dissertation to the memory of my mother:

Eleanor Burke Leacock

TABLE OF CONTENTS

1.0	Introduction	1
1.1	The Parser as Licenser	12
2.0	Models of Language	20
2.1	The Move- α Model	20
2.2	The Lexical Model in "Knowledge of Language"	28
2.3	Licensing Structures	35
2.3.1	Replacing Base Rules with Lexical Structures and X-bar Theory	36
2.3.2	The Function of Licensing Principles: Construct Phrase Markers or Validate Phrase Markers	41
2.3.3	Chomsky's Licensing Principles	43
2.3.4	Abney's Licensing Principles	45
2.4	Chain Formation in the Two Models	48
2.4.1	Forming Chains with Move- α	48
2.4.2	Determining Features of Empty Categories	51
2.4.3	Forming Chains with a Lexical Parser	54
2.4.4	Types of Chains	55
2.5	Chain Structures and Binding Structures	56
2.5.1	The Move- α Model -- Binding as a Filter	57
2.5.2	Redundancy between Inherent Chains and Binding Theory	60
3.0	Motivating D-structure	67
3.1	Chomsky's LGB Argument	68
3.2	Chomsky's Logical Non-Equivalency Argument	77
3.3	'Stubborn' versus 'Illegal'	78
4.0	A Lexically Driven Analysis of 'Stubborn' and 'Illegal'	89
4.1	The Empty Operator and Strong Binding	90
4.1.1	The Empty Operator in Infinitive Clauses	93
4.1.2	The Empty Operator in Parasitic Gaps	95
4.2	The Lexical Structures of 'stubborn' and 'illegal'	95
4.2.1	External θ -role Assignment	96
4.2.2	Complement Structure	97
4.2.3	Comparative Formation	98
4.2.4	Lexical Structures for 'stubborn' and 'illegal'	100
4.3	Control Violations in the Crucial Examples	101
4.4	The Constituent Structure of 'John is too stubborn to talk to'	103
4.5	The Constituent Structure of 'John is afraid to talk to Mary'	111
4.6	Lexical Structures in 'John is illegal to talk to'	112
4.7	The Constituent Structure of 'John is easy to talk to'	117
4.8	The Constituent Structure of 'Mary is pretty to look at'	125
4.9	The Two Constituent Structures of 'John is illegal to talk to'	133
4.10	Lexical Properties of Adjectives	136
5.0	Control Facts and Control Theories	141
5.1	The Control Properties of "Ask"	141
5.1.1	The Lexical Structure of "Ask"	143
5.2	Chomsky's Control Theory in "On Binding"	144
5.2.1	Chomsky (1980a) Control Theory and "Ask"	150

5.3	Williams on Predication and Control	153
5.3.1	"Ask" and the Theory of Predication	159
5.4	Manzini's Control Theory	168
5.4.1	"Ask" in Manzini's Control Theory	173
5.5	Koster on Governed PRO and Control	174
5.5.1	"Ask" and the Theory of Governed PRO	178
5.6	Hornstein and Lightfoot's Theory of PRO	182
5.6.1	"Ask" in the Hornstein and Lightfoot Theory of PRO	190
5.7	Chomsky's Description of Control Properties in "Knowledge of Language"	193
5.7.1	Control in Declarative Sentences	194
5.7.2	Control in Interrogative Complements	196
5.7.3	Control in Adjuncts	197
5.7.4	Control in Subject Sentences	198
5.8	C-command	199
5.9	Control in Passive Constructions	201
5.10	Summary	202
6.0	The Distribution and Interpretation of PRO	204
6.1	The Distribution of PRO	204
6.2	The Pronominal Anaphor 'One'	209
6.2.1	The Person, Number, and Gender of 'One' and PROone	210
6.2.2	The Reference Properties of Overt 'One' and Empty 'One'	217
6.2.3	The Feature [HUMAN]	219
6.2.4	The History of the Overt Pronoun 'One' and Empty 'One'	223
6.3	The Interpretation of PRO	225
6.3.1	The Interpretation of the Features: Generic and Definite	226
6.3.2	The Interpretation of the Feature Bundles	228
7.0	Licensing Mechanisms for Infinitive Constructions	233
7.1	Lexical Specification of Features and Antecedent	236
7.1.1	Lexically Specified Anaphora	237
7.1.2	Lexically Specified Anaphora on Verbs	243
7.1.3	Lexically Specified Anaphora on Adjectives	244
7.2	Long-Distance Control and Binding	245
7.2.1	The reference of the Indefinite Empty Pronoun	246
7.2.2	Long-Distance Control	249
7.3	Feature Assignment by Grammatical Formatives	254
7.3.1	The Assignment of Structure and Control to Passives	254
7.4	The Constituent Structure of Adjuncts	261
7.4.1	Types of Adjuncts	262
7.4.2	The Tree Building Mechanisms for Infinitival Adjunct Clauses	267
7.5	The Control Mechanisms for Adjuncts	278
7.6	VP Adjuncts: 'in order to' clauses	283
7.6.1	The Domain of the VP-Adjunct Mechanism	286
7.6.2	Adjuncts and Complements in the Verb Phrase	287
7.6.3	The Theory of Implicit Controllers	289
7.7	Adverbial Adjuncts: Degree Adverbial Phrases	296
7.7.1	The Mechanism of Indirect Control: Case I	298
7.7.2	The Indirect Referent Mechanism: Case II	301
7.7.3	Indirect Reference and Long-Distance Control: Case III	302

7.7.4	The internal structure of the ADV adjunct infinitive clause.	302
7.8	Noun Infinitival Adjuncts: Infinitival Relatives	306
7.9	Adjectival Adjuncts: Copula Clauses	310
7.10	A Classification of Control Structures According to Features	312
8.0	Lexical Structures of Verbs and Adjectives	315
8.1	Verbs with External Theta Role	316
8.2	Verbs with No External Theta Roles	317
8.2.1	Literature on verbs that Assign no External Theta Role	323
8.3	Distribution and Interpretation of the Empty Operator	325
8.3.1	The motivation for an empty operator.	327
8.4	Lexical Structure of Adjectives	333
8.4.1.1	Control in "Pretty" Adjuncts	334
8.4.1.2	Control in "Tough" Predicates	336
9.0	The Domain of Licensing PRO	338
9.1	The C-command Theory of Control	339
9.2	Thematic Role Theories	341
9.3	The Domain of Licensing of PRO	347
9.4	The Lexical V_1 - V_2 Mechanism of Control	352
9.5	Double-Passive Constructions	368
9.6	The V_1 - V_2 Control Theory Versus the Thematic Role Theory.	376
9.7	The V_1 - V_2 Control Theory Versus the C-command Theory.	379
10.0	Conclusion	381
Appendix A.	Types of Adjectives	388
Appendix B.	Nominalizations	391
Bibliography	393

1.0 INTRODUCTION

Within the current Principles-and-Parameters framework, the grammar, or I-Language (internalized language) can be characterized as a system of rules and principles. Chomsky writes:

As a tentative empirical hypothesis, we might take the I-language to be a rule system of some sort, a specific realization of the options permitted by UG, fixed by presented experience. The rule system assigns to each expression a structure, which we may take to be a set of representations, one on each linguistic level, where a linguistic level is a particular system of mental representation. This structure must provide whatever information about an expression is available to the person who knows the language, insofar as this information derives from the language faculty; its representations must specify just what the language faculty contributes to determining how the expression is produced, used, and understood.

A linguistic level is a system consisting of a set of minimal elements (primes), an operation of concatenation that forms strings of primes, as much mathematical apparatus as is necessary to construct appropriate formal objects from these elements, the relevant relations that hold of these elements, and a class of designated formal objects (markers) that are assigned to expressions as their representations on this level. The rule system expresses the relations among the various levels in the language in question and determines the elements and properties of each level. (Chomsky 1986a:46)

The rule system is generally described as being a collection of principles and parameters. These principles and parameters can be characterized as a finite set of well-formedness conditions. Along with these generalizations, the only other information required is that contained in language specific lexical structures. The combination of the general principles and language-specific lexical structures are the sole requirements for the knowledge and use of language.

If there are n -levels in the I-language model, then any sentence can be considered an n -tuple (X_1, \dots, X_n) with a structure X_α representing the sentence at level $_\alpha$. The diagram in Figure 1 on page 2 as-

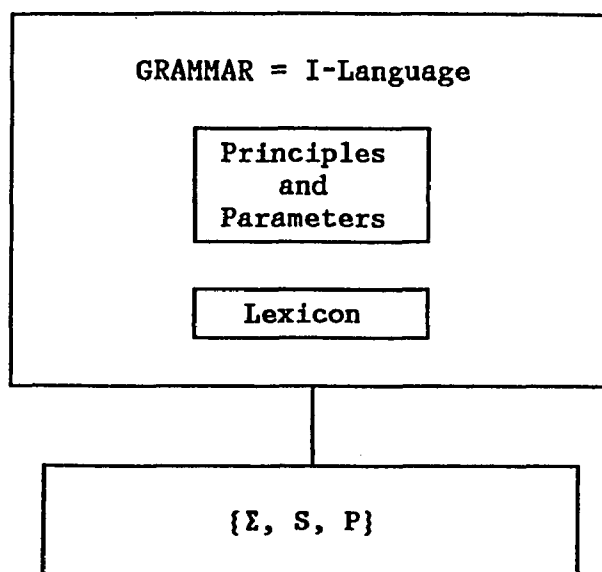


Figure 1. An I-Language Model

sumes that a sentence is a triple with a semantic representation S , a phonetic representation P , and a syntactic representation, Σ . The model, as depicted, is neutral as to whether it contains a transformational component. If there is a transformational component, then Σ consists of a set of phrase markers ($PM_1 \dots PM_n$) where there is a mapping between S and PM_n and there is a mapping between P and PM_n . In a non-movement model, Σ contains a single S-structure phrase marker with a mapping relation between both S and P . If σ consists of more than one phrase marker, then the transformational component is required to act as a link among the various levels. If there is no transformational component in the grammar, then Σ represents a single level with the S-structure phrase marker representing all structures at the syntactic level.

I consider two possibilities for the I-language syntactic representation. The Move- α model assumes that syntax defines at least two levels, a D-structure level and an S-structure level. Therefore, each sentence is a 4-tuple with S, P, D-structure, and S-structure, as in (1):

- (1) A sentence in the Move- α model:
 {S, D-structure, P, S-structure}

The Parser as Licenser model, advanced in this thesis, assumes that syntax contains only one level, S-structure. Therefore, each sentence is a triple with S, P, and S-structure, as in (2):

- (2) A sentence in a Parser as Licenser model:
 {S, P, S-structure}

With the introduction of X'-bar theory, phrase structure rules have been rendered redundant. Now phrase markers are built as projections from heads:

The device of phrase structure rules is particularly suspect, because these rules so closely reflect lexical properties. Statement of the lexical properties is ineliminable from the grammar: For example, the grammar cannot avoid stating that *claim* takes a propositional complement as part of its lexical entry. Therefore, it is to be expected that the phrase structure rules should be eliminable insofar as they merely restate, in another form, the essential content of lexical entries. In fact, it seems that such rules are eliminable more generally, that there are no rules of this type in language.
 (Chomsky 1986b:83)

Given the conditions of X-bar theory in one or another of its variants, lexical entries can be limited to a minimal form with indication of no more than inherent and selectional features, and phrase structure rules can apparently be dispensed with entirely, a highly desirable move for familiar reasons.
 (Chomsky 1986c:3)

In such a system, some mechanisms must mediate between the phonetic form, lexical structures, X-bar Theory, and the projection principle to construct phrase markers.

There are two possible mechanisms for associating a phrase marker with a string containing lexical items. A *parser* is a mechanism that projects a phrase marker from an unstructured string of elements using principles, parameters, and lexical information about the formatives in the string. The alternative to a parser is a set of *phrase structure rules* that derives a phrase marker by starting with an initial symbol (usually CP) and rewrites non-terminal symbols. Lexical items are then inserted into the phrase marker produced by the phrase structure rules. Chomsky (1957) calls the process by which rewriting rules, starting with an initial symbol S, produce a phrase marker containing a terminal string of lexical items a *derivation*.

If phrase structure rules are rejected as the mechanism for producing phrase markers and pairing a phrase marker with a string, then the I-language grammar must contain a parser.¹ In the Move- α model, the parser projects a D-structure phrase marker from a D-structure string. In the Parser as Licenser model, the parser projects an S-structure from an S-structure string. A generative grammar must provide at least two things: (a) a notation in which to represent strings at each level, and (b) an automatic mechanism which assigns the notations to strings.

¹ Terry Langendoen has pointed out to me that Chomsky's attempted elimination of the phrase structure rules could be thought of as the elimination of a top-down generative device in favor of a device, a parser, that assigns structure bottom-up.

Let us consider how a Move- α model would assign a D-structure and S-structure to sentence (3).

(3) Who was John persuaded to visit?

Some string generation mechanism must produce the D-structure string in (4).

(4) PAST be persuade John to visit who

The parser, which projects the D-structure phrase marker from the contents of the lexical entries, would produce a structure like (5).

(5) [e PAST be persuade John [e to visit who]]

The various mechanisms of the Move- α model would then operate on (5) to yield the S-structure (6).

(6) [who_i PAST be John $_j$ persuade e_j [e_i [PRO $_j$ to visit e_i]]]

As we shall see below, a Parser as Licenser model would assign an S-structure directly to the string (3).

The parser in the Move- α model, which projects the D-structure phrase marker from the D-structure string, may differ in details from the parser in the Parser as Licenser model which projects the S-structure phrase marker from the S-structure string. However, both models have one basic fact in common: Since they are both parsers which project a phrase marker from an unordered string of elements containing lexical material, both models *must* have some source for the unstructured strings to which the parser can apply.

We can assume that the strings fed to either parser are produced by a random string generator which strings together lexical items and grammatical formatives in totally random fashion. In the Move- α model, some of these random strings will be parsed and assigned a D-structure

phrase marker. In the Parser as Licenser model, some of these random strings will be assigned an S-structure phrase marker.

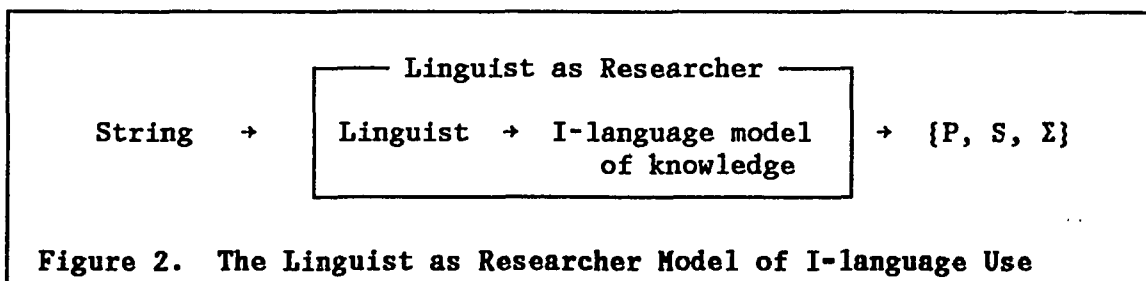
The fact that any theory of I-language requires a source of the strings from which it can project structures has not been discussed much in the literature. The question to be examined follows: What is the source of strings to input to the parser?

One must distinguish between a theory of knowledge (I-language) and a theory of use (how I-language is used). If we assume that both the Move- α parser, which constructs D-structure phrase markers, and the Parser as Licenser parser, which constructs S-structure phrase markers, operate on strings produced by a random string generator, then it is clear that they are both theories of knowledge. A theory of knowledge does not assert there is any temporal priority among the members of the n-tuple. The goal of the Move- α model is to generate well-formed quadruples {P, S, D-structure, S-structure}, and although the parser may project the D-structure from a randomly given string and then find the P, S, and S-structure, the model does not assert that this is a temporal priority with any relation to how a human "knows" that some quadruple is a sentence. The goal of the Parser as Licenser model is to generate well-formed triples {P, S, S-structure} and although some structures may precede others in calculating the three levels, there is no assertion that this is a temporal priority relating to the internalized structure of human knowledge.

A theory of language use does assert that there is a temporal priority among the members of the n-tuple. Let us consider two models

of how a knowledge of language (I-language) model might be incorporated into theories of use.

Consider first the Linguist-as-Researcher Model of I-language use, illustrated in Figure 2. We assume that a linguist has a theory of I-language which provides a formal notation in which to represent sentences. In this model of language use, the Linguist-as-Researcher is presented with a string of elements and, using the formal I-language model, produces the n-tuples for each sentence: $\{P, S, \Sigma\}$. This is a model of language use with a temporal priority to the process. The input String precedes everything else and is the stimulus for the linguist to employ the I-language to construct the n-tuple. The process is successful if the P constructed by the I-language model is the same as the input string.



Suppose the I-language model is the Move- α model. The linguist plays a crucial creative role in the Linguist-as-Researcher model in that the D-structure string from which the D-structure phrase marker will be projected is not the input string. The D-structure string and the input string differ in the order of constituents. The linguist must use imagination and creativity to produce an appropriate D-structure

string so that the Move- α mechanisms will yield the input string as the phonetic representation, P.

Consider a second model of language use: A perception/recognition model. Here, too, is a temporal order in the analysis. The perception/recognition model is presented with an input string and it must assign that string an n-tuple {P, S, Σ }. Suppose the I-language model is the Move- α model. The model, see Figure 3 on page 9, inputs a string and must produce a quadruple {P, S, D-structure, S-structure} such that P is the input string. Since the D-structure string from which the structure is assigned to the sentence does not have the same order of constituents as the input String, the perception model requires a Synthesizer Component to guess an appropriate D-structure from which the correct quadruple may be calculated. The Move- α perception/recognition model is an analysis-by-synthesis model since the device must synthesize a D-structure string, or a set of D-structure strings, and construct n-tuples until one of them contains P equal to the input string.

If we compare the Linguist-as-Researcher and the Perception models which are based on the Move-alpha theory of knowledge of language, we can see that they are very similar. Essentially, the perception model attempts to develop a Synthesizer module to replace the linguist in the Linguist-as-Researcher model. It is this which makes the move- α model difficult, if not impossible, to implement as a theory of perception: One must have a formal computational model of the creative imagination the linguist uses in guessing an appropriate D-structure string to serve as the seed from which to produce an n-tuple.

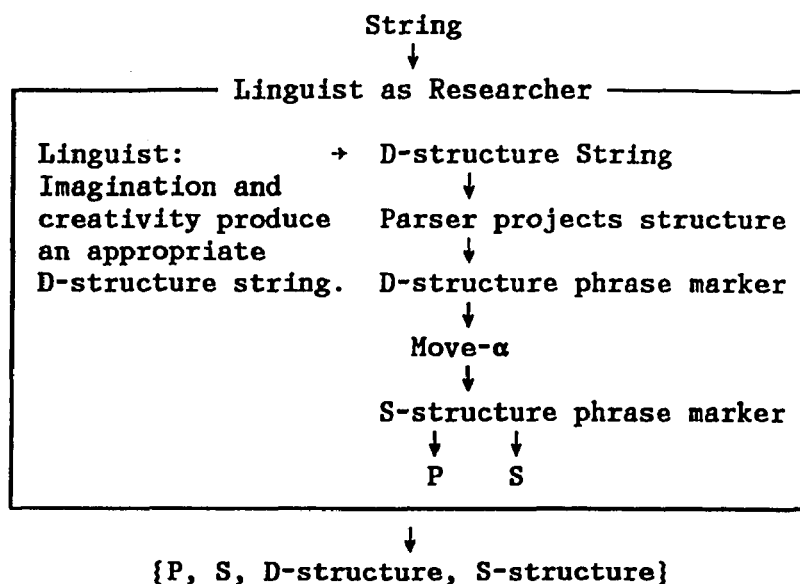


Figure 3. The Move- α Model of Linguist as Researcher

If we adopt an I-language model which develops the structure of a sentence by using a parser to project an S-structure from the string of elements in P, then the development of theories of language use is simplified. A theory of perception/recognition does not require any synthesizer component to guess an appropriate string to match the input string. The theories of language use project all structure from the given string.

Although the I-language model which projects the sentence syntactic structure from the input string lends itself to theories of language use, it is itself not necessarily a theory of language use.² In this

² The idea that a competence and performance model must be kept distinct has been discussed many times in the literature. For example:

... it is important to distinguish clearly between the function and properties of the perceptual model *PH* and the competence model *G* that it incorporates. Both *G* and *PH* relate sound and meaning; but *PH* makes use of much information beyond the intrinsic sound-meaning association de-

dissertation, the Parser as Licenser theory discussed should be evaluated as a proposed contribution to I-language theory of knowledge in terms of its characterization of knowledge of language. I claim that all questions relating to knowledge of language can be answered by an I-language theory which has only {P, S, S-structure}. The fact that the I-language model of knowledge may provide the basis for theories of language use, insofar as this is true, can only be considered a virtue of the model.

Most of the research in principles and parameters models has assumed that each sentence is a quadruple (P, S, D-structure, S-structure). Few arguments have ever been adduced to show that the D-structure level is required to describe or explain the phenomena. It is the purpose of this dissertation to investigate in some detail one of the most recent arguments purported to show that the description and explanation of language requires a D-structure. We will show that the argument is not valid.

terminated by the grammar G , and it operates under constraints of memory, time, and organization of perceptual strategies that are not matters of grammar. Correspondingly, although we may describe the grammar G as a system of processes and rules that apply in a certain order to relate sound and meaning, we are not entitled to take this as a description of the successive acts of a performance model such as PN --in fact, it would be quite absurd to do so. What we have said regarding perceptual models is equally applicable to production models. The grammatical rules that generate phonetic representations of signals with their semantic interpretations do not constitute a model for the production of sentences, although any such model must incorporate the system of grammatical rules. If these simple distinctions are overlooked, great confusion must result. (Chomsky 1968:177)

Parsing consists of assigning a structural description, or a phrase marker, to a string. A parser is an instantiation of the principles which make up the I-language. This aspect of language has been, and continues to be, widely studied. Chomsky sums up the current state of affairs as follows:

The question of how we use the knowledge we have attained has two aspects: the perception problem and the production problem. The first has to do with the interpretation of physical stimuli as linguistic expressions with form and meaning. The second has to do with the use of language for expression of thought, for play, in its esthetic functions, for such specific ends as communication, and so on. The perception problem is amenable to inquiry, and much has been learned about the topic; the production problem remains as much a mystery as it was for the Cartesians, ... The topic now falls within what we call "artificial intelligence" or "cognitive science"; it is in fact, a resurrection of Cartesian ideas in a new framework. We can observe these features of language use, but can say nothing very illuminating about them. (Chomsky 1987:17-18)

The question which I will explore in the next two chapters is: At what level does the lexical parser operate? That is, should it project a D-structure or an S-structure?

With the introduction of trace theory into the grammar, much of the motivation for multiple linguistic levels was eliminated. This has been pointed out by Koster:

... trace theory is not a complement to the standard approach, but an alternative to it: with traces represented at S-structure, it is not necessary to have a separate level of D-structure. In a sense, deep structure does not disappear, because its relevant aspects are now coded into S-structure. (Koster 1987:34)

S-structures now contain all information relevant to their derivation in the form of traces and chain structures.

In this thesis, I explore the possibility of eliminating Move- α between D-structure and S-structure in relation to binding and control

structures. I do not explore the other application of Move- α that has been proposed -- that of quantifier raising which occurs between S-structure and Logical Form (LF). However, there is no obvious reason why this proposed movement can not be worked out along the lines proposed here.

The parsing model explored here is one suggested by Chomsky:

Parsing programs are typically rule based: the parser in effect mirrors a rule system and asks how these rules can assign a structure to a string that is analyzed word by word. The examples just discussed, and many others, suggest that a different approach might be pursued. Given a lexicon, structures can be projected from heads by virtue of the Projection Principle, X-bar theory, and other subsystems of UG that are involved in licensing elements, which are furthermore associated by these principles in the manner already illustrated. Perhaps parsers should not be based on rules at all, but should rather be based on lexical properties and principles of UG that determine structures from them.

(Chomsky 1986b:27)

1.1 THE PARSER AS LICENSER

A fundamental component within the Principles-and-Parameters framework is the licensing principles of *Full Interpretation*. Every element (constituent) in a phrase marker must have a reason for being there. That is, every constituent must be licensed.

Chomsky's Move- α model (Chomsky 1981, 1986a) and my proposal, the Parser as Licenser model, both employ many of the same licensing mechanisms: (a) lexical argument structures, (b) X-bar theory, (c) the projection principle, and (d) binding. The two hypotheses diverge, however, on one basic point; stated in (7).

- (7) How is the location, distribution, type, and reference of any empty category specified?

The Move- α model claims that to answer (7), the syntactic structure of a sentence must consist of at least two phrase markers, a D-structure and an S-structure. By the projection principle, which requires that complements of heads be represented in the phrase marker at each syntactic level, the D-structure and S-structure are required to match the lexical items of each head. Thus the only way the two phrase markers can differ is in the order of the lexical projections. The order of constituents in the S-structure is the same as that of the PF (phonetic form). If Move- α has applied, the order of constituents in the D-structure will differ from that of the PF. Chains, data structures created by successive applications of the transformation Move- α , express the difference in the linear order of constituents between the D-structure and the S-structure. Chains serve only to record, in the S-structure, the history of movement for any lexical projection whose linear position in the S-structure and D-structure differ. A vacuous case of a chain, which is of little interest here, is one assigned to an unmoved element which has the same position in the D- and S-structures.

The Parser as Licensor model claims that to answer (7), the syntactic structure of a sentence can be a single phrase marker, an S-structure. This S-structure has the same order of constituents as the PF. In solving (7), there is never a data structure in the parsing mechanism in which the order of constituents differs from that of the input string. There may, however, be additional nodes (empty categories) inserted between the overt elements of the surface string.

The question that is always asked about an ill-formed syntactic structure within the Move- α Model is: "How is this structure blocked

by the grammar?" The question to be asked within the Parser as Licenser model is: "Why doesn't the parser build this structure?" To answer the first question, all possible bad structures which can be the outcome of transformations and of free indexing have to be filtered out. To answer the second question is not a matter of filtering out bad structures, it is a matter of determining which structures are impossible for the parser to build in the first place.

The two models have quite different analyses for (8). The Move- α model analyzes (8) as having a D-structure (9i) and an S-structure (9ii). A Parser as Licenser model requires only the S-structure (9ii).

(8) Who was John persuaded to visit

- (9) i. e PAST+be+en persuade John e to visit who
 ii. who_j PAST+be+en John $_j$ persuade e_j [e_j [PRO $_j$ to visit e_j]]

Chomsky (1986a:102-104) analyzes (8) using a lexical parsing model, many aspects of which have been incorporated into the Parser as Licenser model. He attempts to discover the properties and bindings of the empty categories based solely on their S-structure configuration in conjunction with lexical structures and licensing properties. In this analysis, Chomsky makes no use of D-structure, Move- α , or chains. However, as I will show in "Models of Language" on page 20, the version of a lexically driven parsing model briefly outlined in Chomsky (1986a) is an inadequate model.

My parser's goal is to project an S-structure from a string. To do this, it has to predict the properties of empty categories. It has to posit their locations and types, and bind them to their antecedents. Once a parser does this, D-structure and the transformational component

which derives S-structure from D-structure no longer have any role in the grammar.

The parser is lexically driven, generating an S-structure directly as a projection of the lexical entries of the items in the string under the constraints imposed by the modules and full interpretation. In this model, X-bar theory operates directly to define and construct S-structures. The projection principle operates between lexical structures and S-structure. In addition, the principle of Full Interpretation has to apply to S-structure. It is FI's licensing principles that will have to bear much of the burden of replacing the phrase structure rule component. Indexing structures are determined solely on the basis of the S-structures output by the parser.

It is not necessary for a parser to build ill-formed indexing structures which must then be filtered out. For example, knowing, due to the parser's access to lexical structures, that a given indexing structure results in a Theta Criterion violation, the parser would not build the structure for another component of the grammar to subsequently rule out.

People try to describe patterns of grammaticality by mechanisms specified in terms of structural configurations defined on a set of existing trees which have been assigned to a string. I argue, rather, that the best way to describe patterns of grammaticality is to formulate all constraints in terms of the principles and processes which assign the structures to strings. The basic issue is how the parser can assign a structure to an input string. More precisely, how is a tree projected

from the items in the string and lexical information about those items? It is not the parser's job to explain why some structures are ill-formed.

In "Models of Language" on page 20, I sharpen the basic assumptions underlying the discussion in *Knowledge of Language* and present two alternative theories: one with movement and one without movement. I indicate which components of the grammar are needed to drive one version of a non-movement model.

In "Motivating D-structure" on page 67, I review Chomsky's arguments concerning the relation between a model with movement versus one without movement. In Chomsky (1981), he indicates that the two are notational variants. If this is true, then the model will be chosen based on parsibility from a computational point of view.

A more serious argument against a non-movement model involves the claim that there are sentences which can be derived with movement that cannot be derived without movement. The strongest and most recent such argument for Move- α over a non-movement system is presented in *Knowledge of Language* (Chomsky 1986a:114-114), and involves the two sentences in (10).

- (10) i. John is too stubborn to talk to.
 ii. * John is illegal to talk to.

Chomsky states that his analysis of these sentences "crucially requires that we regard Move- α as an operation converting D-structure to S-structure" (Chomsky 1986a:114). However, in analyzing the sentences in (10), Chomsky presents an hypothesis which is a mixture of the movement and non-movement views. I argue that his analysis of the sentences in (10) does not hold for either model. However, the issues that these

sentences raise are crucial, and in the rest of the thesis I develop a parsing model which builds and interprets sentences containing infinitive clauses.

In "A Lexically Driven Analysis of 'Stubborn' and 'Illegal'" on page 89, I tease apart the various phenomena that combine to produce the crucial examples. Fundamental to an analysis of the sentences in (10), is a description of their control properties. In neither sentence is the subject NP of the matrix clause, *John*, the antecedent for PRO in the embedded infinitive clause. Both are instances of what has traditionally been called PRO_{arb}. I propose a non-movement analysis for comparative adverbials, *tough* predicates, *pretty* predicates, and other predicate adjective constructions, and their control properties.

Control theory is often considered to be neutral as to whether it operates on a tree which is the output of Move- α or on a tree which has been generated without movement. However, in our analysis of adjectives, *tough* predicates play a crucial role, and there is currently no way of generating *tough* predicates in the Move- α model of Chomsky (1986a). A non-movement analysis is required in order to generate the trees on which the control module operates.

In "Control Facts and Control Theories" on page 141, I present the control theories of Chomsky (1980), Williams (1980), Manzini (1983), Koster (1984), and Hornstein and Lightfoot (1987). These theories try to predict control properties based purely on the structural configurations of government and c-command. The control properties of *ask* in particular are tied to its lexical analysis, and I apply the mechanisms for the various control theories covered to see whether they can

predict the control properties of *ask*. I close the chapter by bringing together the range of control facts and organize them according to the syntactic structures in which they occur.

In "The Distribution and Interpretation of PRO" on page 204, I argue that the traditional division between instances of PRO_{arb} and PRO_{index} is impoverished--and that a much richer range of interpretations can be predicted. I introduce the two extrinsic features which are assigned by the parser: [\pm generic] and [\pm definite]. One of the four resulting feature bundles is the empty pronoun *one*--which shares the interpretation of its morphological counterpart.

In "Licensing Mechanisms for Infinitive Constructions" on page 233, I present an informal description of a lexically-based parser which uses the surface string, lexical structure, principles, and parameterized modules to generate phrase markers for strings containing English infinitive clauses.

I divide the resulting phrase markers into three basic control structures: (1) lexically specified control, (2) long-distance control, and (3) control of adjuncts. I then provide an algorithm for how the control module assigns extrinsic features and antecedent(s) to each configuration.

In "Lexical Structures of Verbs and Adjectives" on page 315, I review the range of lexical entries for verbs and adjectives in terms of argument structure, extrinsic feature, control, and θ -role assignment. I show how the syntactic properties of verbs and adjectives follow from properties of lexical structures. In particular, I show the consequences of external and internal θ -role assignment.

In "The Domain of Licensing PRO" on page 338, I offer the control mechanism for double-passive constructions. I analyze the control phenomena in terms of lexical structures and the presence of passive morphology in the embedded clause. In order to predict control in double-passive constructions, the domain in which the control module operates must be two sentences.

In conclusion, I have tried to motivate why a model without movement is an important field of investigation, and I give an informal presentation of how such a model would work. In so doing, I have developed a model of control which assigns a full range of interpretations to instances of PRO in infinitive constructions in English.

2.0 MODELS OF LANGUAGE

In *Knowledge of Language* (Chomsky 1986a), Chomsky presents two conceptually distinct models of language internal to the Extended Standard Theory (EST) framework. However, he does not formally distinguish between the two models. Instead, he merges them together. In this chapter, I will define each as a set of coherent assumptions. I will call one the Move- α model and the other the Lexical model.

The Move- α model imposes order on the unordered arguments represented in the lexicon at the level of D-structure. S-structure is derived from D-structure through the operation of the transformational rule, Move- α . This S-structure contains chains which reflect the successive positions of moved elements. Interpretation in the logical form (LF) is facilitated by these chains.

Chomsky formulates the Lexical model as a parsing model. With the aid of lexical structures, this model maps an S-structure directly from the input string. The Lexical model makes no use of a level of D-structure, and the content of empty categories is contextually determined. Therefore the relation between D-structure and S-structure, Move- α , is not required either. Since Move- α is not used in the lexical model, the chain structures which are required for interpretation have to be constructed by the parser.

2.1 THE MOVE- α MODEL

Most current research in the Principles and Parameters framework is within the Move- α model. The organization of components for the

Move- α model is essentially that of the "Extended Standard Theory" (EST) as illustrated in Figure 4 on page 22 from Chomsky (1986a:68):

In the EST framework, phrase structure rules (I) generate D-structures which "express semantically relevant grammatical functions and relations" (Chomsky 1986a:67). (I) is the only component of the EST model which has been eliminated from the Move- α model:

We have kept to the general framework of the extended standard theory (EST), while virtually eliminating the rule systems it presupposed. D-structure may now be defined as a "pure" representation of θ -structure: Each argument is in a θ -position and each θ -position is occupied by an argument, where an argument is a phrase that must be assigned a θ -role to satisfy the licensing conditions. (Chomsky 1986a:155)

The Move- α model, a version of EST, includes the following assumptions:

1. The syntactic structure of a sentence consists of four levels of representation: D-structure, S-structure, PF (phonetic form or surface structure), and LF (logical form).
2. Certain mapping operations hold between each level. The operation holding between D-structure and S-structure is the transformational component (II in Figure 4 on page 22) which, in the current principles-and-parameters framework, consists of the single transformational rule Move- α . Phonological and other rules (III in Figure 4 on page 22) hold between the S-structure and surface structure (PF). Rules of logical form (IV in Figure 4 on page 22) hold between S-structure and the structures in LF where scope is represented. An indexing mechanism which freely assigns indices to NPs also operates between S-structure and LF. Chomsky does not define the operation relating the lexicon and D-structure (I in Figure 4 on page 22) in his updated framework. Presumably it is

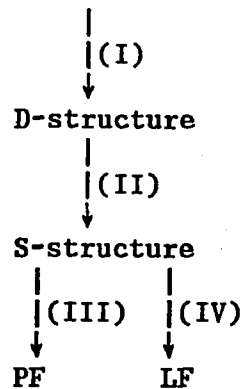


Figure 4. Components of the Extended Standard Theory (EST)

some sort of mapping between lexical structures and D-structure, since D-structure is defined as "a direct projection of lexical structure subject to other licensing conditions (e.g., those that guarantee the presence of subjects of predication)" (Chomsky 1986a:161). This mapping operation has to impose order on the arguments, since arguments in the lexicon are unordered.

3. Certain well-formedness conditions hold between two or more levels. Examples of these are the Projection Principle and Principle of Full Interpretation.
4. The S-structure reflects the history of movement in that it contains traces of moved elements in the positions they occupied at D-structure and intermediate stages of the derivation. As Move- α applies, chains are built, connecting traces to their antecedents via coindexation. By this mechanism, all the semantic and syntactic information about complement structures that is represented in the D-structure is available at the level of S-structure.

Language users are genetically endowed with a Universal Grammar (UG), a modular system of components or principles (with variable parametric settings) which apply at one or another of these levels. UG is the initial state S_0 or default setting of the parameters. These parameters may be assigned values different from the default as the result of input to the language acquisition device (LAD). Structural differences among languages are the result of this parametric variation of the modules. Chomsky's (1986a) description of UG follows:

UG consists of various subsystems--X-bar theory, binding theory, Case theory, theta theory, bounding theory ... and so forth--each containing certain principles with a limited degree of parametric variation. In addition, there are certain overriding principles such as the projection principle, FI (full interpretation), and the principles of licensing. Certain concepts, such as the concept of domain discussed earlier and the related technical notions c-command and government, to which we return, play a central role throughout these subsystems. The interaction of the principles of these various "modules" determines the structure of each possible string--its representations on each level... phrase structure rules can be largely eliminated, perhaps completely so. (Chomsky 1986a:102)

Within the Move- α model, it is not specified whether the parametric variation serves to restrict the application of Move- α , thereby constraining movement, or whether it imposes well-formedness conditions on the data structures resulting from unrestricted movement. The model allows for either of these approaches. I will refer to the principles and parameters as applying to data structures -- but they may equally apply to the processes from which the data structures are derived.

At the level of D-structure, X-bar theory and theta theory apply. Case theory applies at both D-and S-structure. At the level of S-structure, bounding theory applies. At the LF level, binding theory applies. Control theory also applies at LF, and there have been attempts

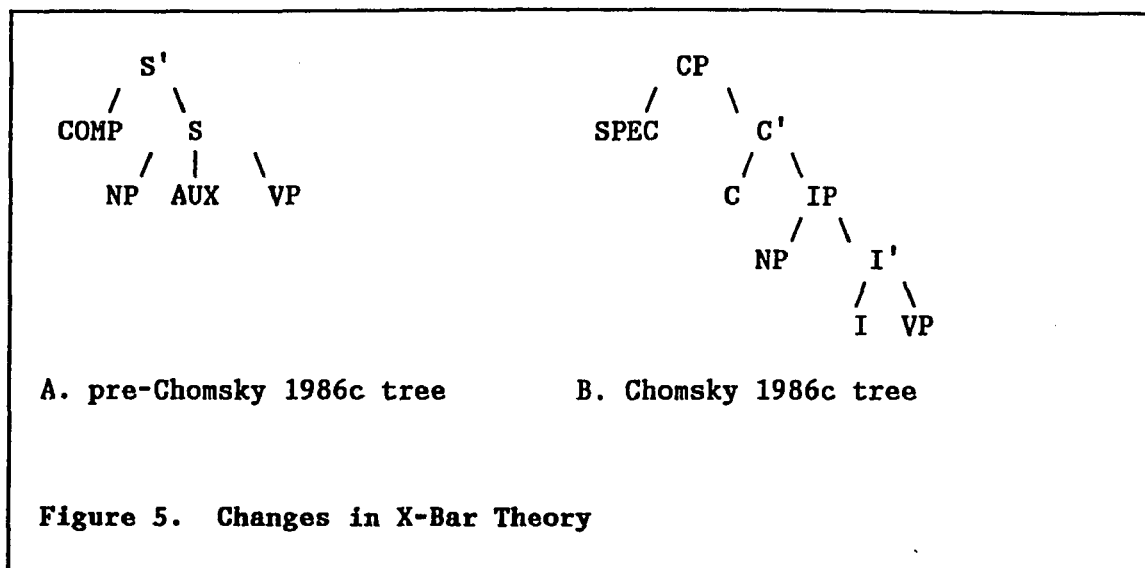
to reduce it to binding theory (Manzini 1983). These modules interact with principles that hold between levels: the projection principle and full interpretation (FI) with its licensing conditions.

The subsystems are typically defined in terms of government, c-command, and the concept of domain. The principles and parameters apply within a structurally defined domain. C-command and Government are the structural relations that define a domain relevant for grammatical processes. Ideally, domain, C-command and Government should provide all the structural conditions necessary for stating constraints on movement, binding, and control. They are the primitive vocabulary of UG.

X-bar theory captures the generalization that all of the major grammatical categories of language (N, V, A, P, S, C) have the same internal skeletal structure. It ensures that, in the D-structure, each category X'' in the grammar has the same internal structure which consists of a specifier, the head X, and the head's complements.

We assume the version of X-bar theory (Chomsky 1986c) where both CP and IP are maximal projections. X-bar theories previous to Chomsky (1986c) assume that CP is a defective node in that it does not have a specifier position. Instead CP was expanded simply to COMP S. Thus what we refer to as CP is frequently referred to as S' in the literature (especially previous to 1986) and what we refer to as IP is frequently referred to as S. These differences are illustrated in Figure 5 on page 25.

The order of constituents, the head in relation to its specifier and its complements, is determined by the setting of the head-first



parameter of X-bar theory. An alternative possibility is that the order of constituents is determined by the conjunction of the direction of Case assignment, θ -marking, and the Case adjacency principle.

Theta theory links the verb to its arguments. Arguments receive their θ -roles, are θ -marked, by the verb at the level of D-structure. In fact, D-structure is "a pure representation of theta structure" (Chomsky 1986a:157). The Theta Criterion, which is the "fundamental" principle of theta theory "expresses the intuitive idea that each argument is assigned its θ -role in exactly one θ -position (namely, at D-structure), and that each assignable θ -role must be assigned to an argument" (Chomsky 1986a:184).

Case theory accounts for the order of the arguments of a head, the distribution of the preposition *of*, and the distribution of the empty category PRO. It also distinguishes Case assigners (V and PREP) from non-Case assigners (N and ADJ). This module differs from the other

modules in that it applies to two levels, whereas the other modules apply to one level only. The claim of Case theory is that "every phonetically realized NP must be assigned (abstract) Case" (Chomsky 1986a:74). Case is assigned uniformly whether or not it is morphologically realized. "Structural Case", such as objective and nominative, is assigned at S-structure, and "inherent Case", which include oblique and genitive case, is associated with θ -marking and so occurs at D-structure.

The transformational component, Move- α , converts D-structures to S-structures. The rule Move- α is currently formulated as "move anything anywhere" (Chomsky 1986a:74), and its output is S-structure. The S-structure contains its own derivational history in the form of chains. A chain consists of an antecedent and those traces indicating its successive positions as it moved from its D-structure position to its S-structure position.

At the level of S-structure, bounding theory specifies the structural conditions which restrict how far an element can move in any one step. Bounding theory "states that a transformation cannot move a phrase 'too far' in a well-defined sense" (Chomsky 1986a:72). Since bounding theory applies at S-structure, it constrains only the output of Move- α . It does not apply at LF, or to the rules that map S-structure to LF.

Binding theory applies at the level of LF (Chomsky 1986a:179). It is "the subtheory of UG that is concerned with the principles that govern the relations between referential dependents such as reciprocals, reflexives, and pronouns on the one hand, and their possible antecedents on the other" (Chomsky 1986a:77). Binding theory describes the possible

structural relations between an anaphor, pronominal or r-expression and its antecedent. In addition, binding theory accounts for the distribution of the null element PRO.

Control theory, which also applies at LF, determines the reference of PRO (the phonologically null subject of an infinitive clause). Control theory determines (a) when PRO must be coindexed with an antecedent, (b) the position of the antecedent, and (c) when it can have an arbitrary interpretation (PRO_{arb}).

The projection principle and the principle of full interpretation are conditions that hold at several levels. The projection principle, which holds at D-structure, S-structure and LF, "requires that complements of heads must be represented at each syntactic level" (Chomsky 1986a:116). Thus, if a complement is represented in the lexicon, it must be represented in D-structure and in all subsequent levels. The projection principle does not require that external arguments (subjects) be represented. A separate principle, the *extended projection principle*, insures that clauses have subjects.

If the Extended Projection Principle is correct, then (a) an EC is present wherever a θ -role is assigned but the corresponding θ -position contains no lexical material, and (b) the category S must always contain an EC as subject if no overt subject is present, as in infinitival clauses or finite clauses in pro-drop languages with missing or inverted subjects. (Chomsky 1981:17)

The principle of full interpretation, which applies to the levels of PF and LF, requires that every element in a string be licensed.

We might express many of these ideas by saying that there is a principle of full interpretation (FI) that requires that every element of PF and LF, taken to be the interface of syntax (in the broad sense) with systems of language use, must receive an appropriate interpretation--must be licensed in the sense in-

licated. None can simply be disregarded. At the level of PF, each phonetic element must be licensed by some physical interpretation. (Chomsky 1986a:98)

The principle of full interpretation ensures that every element in the PF and LF levels has a reason for being there, every element has to be licensed. Chomsky suggests some licensing "options": An operator is licensed by binding a variable, subject to bounding theory; verbs license their complements through θ -assignment; predicates are licensed by their subjects (Chomsky 1986a:93).

The Move- α model just described is a stepwise process from a canonical D-structure through movement via a transformational component to an S-structure containing this historical derivation. The chains created by the operation of Move- α are required for interpretation.

2.2 THE LEXICAL MODEL IN "KNOWLEDGE OF LANGUAGE"

Chomsky's (1986a) Lexical model is a parsing model that assigns structures to strings. When a sentence is given to a parser, the parser has no information about the structure of the string. There is no indication of empty categories -- of their location, of their type, and of their bindings. The parser must, solely on the basis of the input string, assign a syntactic structure to the string, as shown below:

input string \rightarrow parser \rightarrow syntactic structure

Within the framework of transformational grammars, parsers have been proposed using two systems. There are rule based parsers and there

are lexically driven parsers. Rule-based parsers (Correa 1988, Wehrli 1988) use context free phrase structure rules to build phrase structure, onto which lexical items are mapped. Lexically driven parsers, a model of which I develop here, build phrase structure for an input string based upon lexical structure in conjunction with components of the Principles-and-Parameters framework.

Chomsky broadly outlines how a lexical model would work in the following passage, when he describes how a child understands a sentence.

To illustrate how such a system works, consider the sentence:

<92> who was John persuaded to visit

Let us ask what specific knowledge the child must acquire in order to be able to assign to the sentence <92> the structure that underlies its semantic interpretation and use. In other words, what specific knowledge must we have beyond that incorporated in the initial state S_0 in order to understand this sentence, insofar as the language faculty (in our sense) contributes to this end?

We must, first of all, know the lexical properties of the words; otherwise we cannot understand the sentence. We must know, then, that *visit* is a transitive verb that s-selects a category canonically realized as an NP object. By X-bar theory, *visit* must head a VP and, by the projection principle, its NP object must appear in the syntactic representation. The object must be an empty category, because no overt NP is present. One of the values of the X-bar theory parameters for English is that English is a "head-first" language, so that the object is to the right of *visit*. Furthermore, to be licensed, the predicate [*visit* e] must have a subject, the two forming a clause (S); since the subject is not overt, it must be another empty category.

Turning to *persuade*, we know that it is a verb that takes an object and a clausal complement, their order determined by the Case adjacency principle as we have seen. Continuing in this way, we conclude that the structure of <92> must be <93>
...

<93> who was [John [_{VP} persuaded e_i [e_j to [_{VP} visit e_k]]]]

This much of the structure is determined simply on the basis of lexical properties and the value of the head-complement parameter, given principles of UG.

For the structure <93> to be well-formed, each element must be licensed. The *wh*-phrase must bind a variable, and each argument must be assigned a θ -role. For reasons determined by UG, only e_k may be a variable (the other empty categories are not in Case-marked positions and, therefore, are not visible for θ -role assignment). Therefore, *who* must bind e_k . *John* is the subject of a passive, a position to which no θ -role is assigned ... Therefore, *John* must bind some element in a θ -position that can transfer its θ -role to *John* by the general convention regarding chains. Unless bound by *John*, e_j will not be licensed. Therefore, *John* must bind e_j , which, although not in a Case-marked position, is nevertheless visible for θ -marking by *persuade* because it is in a chain headed by the Case-marked element *John*.

This leaves e_j , an element that we will refer to as PRO, an empty category with a restricted distribution ...

Returning to <93>, we know that *persuade* requires that its object control PRO, ... So in <93>, PRO (namely, e_j) is controlled by e_j , the object of *persuade*. In <93>, then, $i=j$, e_j is bound by *John*, and e_k is bound by *who*.

All these connections are uniquely determined by general principles. Spelling them out, we interpret <92> roughly as "for which person x , someone persuaded John that John should visit x ." To achieve this interpretation, the only information required specifically about English is knowledge of the lexical items. This, of course, must be learned, although there undoubtedly are very heavy universal constraints in this system as well, as is clear from the remarkable speed and precision of lexical acquisition, which pose another variant of Plato's problem in quite a sharp form. The remainder is deduced from general principles. (Chomsky 1986a:102-105)

The lexically driven parser described in this passage is flawed. Part of the problem is that it is not a purely lexically driven parser. We will discuss these problems in "'Stubborn' versus 'Illegal'" on page 78. However, in order to convey a sense of how a lexically driven parser operates, I will temporarily overlook the problems that the Chomsky (1986a) Lexical model presents, and go through the parser's operations, step-by-step.

Chomsky's parser begins by finding the most deeply embedded verb in the string³ (in this case, *visit*), and accesses the lexicon for the verb's complement structure. The parser learns from the lexical structure of *visit* that it is transitive, taking an NP object: *visit* [__ NP]. Since *visit* is a verb, by X-bar theory, it must be the head of a VP. The parser looks for an overt NP to fill the object position of the VP it is building. Since no overt object is present, and the Projection Principle requires that arguments be represented at each syntactic level, the parser inserts an empty category in the direct object position³ of the verb phrase. The linear position of the object node in relation to the head is determined by the "head-first" setting of the X-bar parameter. The parser then looks for a subject to license the VP. Since no overt subject is found, the parser inserts an empty category in the subject position of *visit*. When the parser is finished building the most deeply embedded clause, the phrase marker contains the information in (1).

(1) [e to visit e]

The parser then moves the next clause up and finds the verb *persuade*. Accessing *persuade*'s lexical structure, the parser learns that *persuade* takes a direct object NP and a clausal complement: *persuade*[__ NP CP]. In this case, the parser only finds the clausal complement. Finding no overt category in the direct object position, the parser inserts an

³ It is not clear whether the parser is required to start with the most deeply embedded verb (which is obviously problematic in that the parser has no way of knowing which is the most deeply embedded verb until the parse is completed) or whether Chomsky starts with the most deeply embedded verb simply to make his explication clearer.

empty category in the NP object position for *persuade*. The order of the complements for *persuade* are determined by the Case Adjacency Principle which requires that "where Case is not morphologically realized, a Case-marked element must be adjacent to its Case-assigner" (Chomsky 1986a:82). Thus the direct object EC is adjacent to the verb. The parser then looks for a subject to license the VP headed by *persuade*, and finds *John*. The parser has now built the structure in (2).

(2) [John persuaded *e* [*e* to visit *e*]]

Having built this much structure, the parser licenses the remaining constituents with binding and θ -role assignment. To be licensed, the *wh*-phrase *who* must bind a variable, which it finds in the direct object position of *persuade*. The subject of the passive, *John*, must bind an element in a θ -position in order to receive a θ -role. In addition, the object of *persuade* needs to be bound, and *John* is in a legal position to bind it. The subject of an infinitive is the empty category PRO, which receives its θ -role from *to visit*. The lexical structure for *persuade* also requires that its object control PRO. In this way, the S-structure (4) is assigned to the input sentence (3).

(3) Who was John persuaded to visit?

(4) who_i was [$John_j$ persuaded e_j [PRO_j to visit e_i]]

The components of the Lexical model, as described by Chomsky in the preceding quote, are X-bar theory with its head-first parameter, Case theory with its Case adjacency principle, θ -theory, binding theory, control theory, the projection principle, and licensing principles. As in the Move- α model, the phrase structure rule component has been

eliminated. Unlike the Move- α model, all of these principles and parameters are applied to S-structure in the above-quoted passage.

Although Chomsky (1981:102) states that bounding theory plays a role in the above-quoted passage describing a parsing model, bounding theory plays no role in his description of the parser. What Chomsky has, in fact, described is a model with no movement in it. If there were movement, constrained by bounding theory, then the S-structure would contain traces and chains linking a trace to its antecedent. If chains are already extant, then the parser already knows what empty category *who* is the antecedent of. In the model just described, empty categories are assigned on the basis of information about complement structure in the lexicon. This parser does not know, *prima facie*, that *who* is coindexed with an empty category in the direct object position of *visit*. Therefore, in this model, there are no ready-built chains in the S-structure providing a "history of movement". Chains are constructed by the parser.

In Chomsky's Lexical model, phrase markers are built based on the complement structures in lexical representations. The parser is not working with a given S-structure, it is building one for the input string. This Lexical model finds the antecedent of an empty category it inserts in the direct object position of a transitive verb using the parser's licensing conditions. Indexing structures are determined solely on the basis of the S-structures built by the parser. In this model, binding theory accounts for the so-called "movement" of a Move- α -model.

THEORY I: Move- α Model	THEORY II: Parser as Licenser Model		
Transformations:			
Move- α			
Bounding:			
(a)	Wh-movement <i>Ex. Who_i did John see e_i</i>		
(b)	Empty operator movement <i>Ex. The man [O_i [I saw e_i]]</i>		
(c)	NP-movement <i>Ex. John_i was seen e_i</i>		
Binding:			
(d)	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> Principle A reflexives <i>Ex. John_i saw himself_i</i> NP-trace <i>Ex. John_i was seen e_i</i> </td> <td style="width: 50%; vertical-align: top;"> Principle A reflexives NP binding <i>Ex. John_i was seen e_i</i> </td> </tr> </table>	Principle A reflexives <i>Ex. John_i saw himself_i</i> NP-trace <i>Ex. John_i was seen e_i</i>	Principle A reflexives NP binding <i>Ex. John_i was seen e_i</i>
Principle A reflexives <i>Ex. John_i saw himself_i</i> NP-trace <i>Ex. John_i was seen e_i</i>	Principle A reflexives NP binding <i>Ex. John_i was seen e_i</i>		
(e)	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> Principle B pronouns <i>Ex. John_i knew that Bill saw him_i</i> <i>Ex. *John_i saw him_i</i> </td> <td style="width: 50%; vertical-align: top;"> Principle B pronouns </td> </tr> </table>	Principle B pronouns <i>Ex. John_i knew that Bill saw him_i</i> <i>Ex. *John_i saw him_i</i>	Principle B pronouns
Principle B pronouns <i>Ex. John_i knew that Bill saw him_i</i> <i>Ex. *John_i saw him_i</i>	Principle B pronouns		
(f)	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> Principle C r-expressions <i>Ex. He_i saw John_j</i> wh- trace <i>Ex. Who_i did John see e_i</i> empty operator trace <i>Ex. The man_i [O_i [I saw e_i]]</i> </td> <td style="width: 50%; vertical-align: top;"> Principle C r-expressions wh- binding empty operator binding </td> </tr> </table>	Principle C r-expressions <i>Ex. He_i saw John_j</i> wh- trace <i>Ex. Who_i did John see e_i</i> empty operator trace <i>Ex. The man_i [O_i [I saw e_i]]</i>	Principle C r-expressions wh- binding empty operator binding
Principle C r-expressions <i>Ex. He_i saw John_j</i> wh- trace <i>Ex. Who_i did John see e_i</i> empty operator trace <i>Ex. The man_i [O_i [I saw e_i]]</i>	Principle C r-expressions wh- binding empty operator binding		
Control:			
(g)	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> PRO = anaphor <i>Ex. John_i is eager PRO_i to please.</i> </td> <td style="width: 50%; vertical-align: top;"> PRO = anaphor </td> </tr> </table>	PRO = anaphor <i>Ex. John_i is eager PRO_i to please.</i>	PRO = anaphor
PRO = anaphor <i>Ex. John_i is eager PRO_i to please.</i>	PRO = anaphor		
(h)	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> PRO = arb <i>Ex. John is easy PRO_{arb} to please.</i> </td> <td style="width: 50%; vertical-align: top;"> PRO = arb </td> </tr> </table>	PRO = arb <i>Ex. John is easy PRO_{arb} to please.</i>	PRO = arb
PRO = arb <i>Ex. John is easy PRO_{arb} to please.</i>	PRO = arb		

Figure 6. The Two Hypotheses

The components required for licensing empty categories in the two models are outlined in Figure 6 on page 34. The Move- α model requires a transformational component, consisting of the single rule Move- α , Bounding theory which specifies how far a constituent can move, Binding theory which delimits the structural relations between referential dependents and their antecedents, and Control theory which specifies the referential properties of the null subject PRO. The Parser as Licenser model requires only Binding and Control theories to license empty categories. The transformational component and Bounding theory of the Move- α model are replaced in the Parser as Licenser model by binding theory.

2.3 LICENSING STRUCTURES

Based on work by Stowell, Chomsky wants to eliminate the phrase structure rule component from the grammar. The claim is that phrase markers can be projected from the principles, parameters, and lexicon. These are sufficient to determine all the necessary structure in the phrase marker, and therefore phrase structure rules are redundant. All it has at its disposal are the various well-formedness conditions imposed by the principles and parameters of the theory. Within this system, a parser has no recourse to phrase structure rules with which to build structure. Full Interpretation's licensing conditions can be viewed as an attempt to encode information necessary for building an S-structure from a surface string.

The intuitive notion behind licensing is nicely put by Abney:

every element in the structure is licensed by performing a particular function in the structure; the structure is well-formed only if every element in it is licensed.

(Abney 1986:1)

Chomsky (1986) touches on the relation between licensing and parsing, and suggests some possible licensing relations. Modifying Chomsky's proposal somewhat, Abney (1986) proposes a deterministic parser which recovers the licensing structure of a sentence.

I will (a) review the motivation for eliminating the base-rule component from the grammar, (b) review Chomsky's discussion of full interpretation and its licensing principles, and (c) discuss Abney's licensing principles and certain aspects of his proposed parser.

2.3.1 REPLACING BASE RULES WITH LEXICAL STRUCTURES AND X-BAR THEORY

Chomsky (1986a) wants to either constrain phrase structure rules (henceforth PSRs) or, preferably, eliminate them from the grammar entirely, because they are too powerful.

In fact, the phrase structure component poses essentially the same problems as those illustrated with regard to transformations: Far too many possible systems of the permitted format are available, so explanatory adequacy is sacrificed and our variant of Plato's problem is unresolved.
(Chomsky 1986a:81)

An additional problem with rule-based parsers, which is also pointed out by Chomsky, is the computational burden they place on the parsing mechanism:

Rule-based parsers are in many respects implausible. For one thing, complexity of parsing increases rapidly as rules proliferate; for another, since languages appear to differ substantially if viewed from the perspective of rules systems, they will require quite different parsers if the latter are rule-based, an unlikely consequence. The entire question needs substantial rethinking, so it appears.
(Chomsky 1986b:27)

When one looks at the grammar from the point of view of either reducing or eliminating the base-rule component, within a principles and parameters model, there are obvious redundancies (1) between the PSRs and the lexicon, and (2) between the PSRs and X-bar theory.

Redundancy between PSRs and lexical structures occur in the projections of lexical categories (N, V, P, ADJ).

Separation of the lexicon from the syntax permits still further reduction of the phrase structure component, which can be regarded as a kind of "projection" of lexical properties. As we have seen, phrases typically consist of a head (noun, verb, adjective, preposition, and possibly others) and an array of complements determined by lexical properties of the head. The category consisting of the head and its complements is a *projection* of the head (NP if the head is an N, VP if the head is a V, etc.).

(Chomsky 1986a:81)

Lexical structure, in conjunction with the projection principle, which "requires that complements of heads must be represented at each syntactic level (D-structure, S-structure, LF) ... " (Chomsky 1986a:116) ensures that the arguments of a head will always be syntactically represented. For example, (5i) might be a fragment of the lexical entry for *hit* and (5ii) a VP expansion in the base-rule component. These two data structures, which inhere in different components of the grammar, give the same information, basically that a transitive verb takes an NP object as its complement.

- (5) i. *hit* [__ NP]
 ii. VP → V NP

According to the projection principle, if *hit* subcategorizes for an NP object in the lexicon, it must have an NP object at every level of representation. To have the argument structure stipulated only once in the grammar, either (5i) or (5ii) has to be eliminated. Chomsky

claims that "Statement of the lexical properties is ineliminable from the grammar" (Chomsky 1986a:83). However, if the subcategorization information is removed from the verb, then presumably some mapping schema could be devised to make sure that *hit* triggers the appropriate VP expansion. On the other hand, if the PSR (5ii) is eliminated, the appropriate VP expansion follows automatically, with no additional machinery.

The lexical structure, in conjunction with Case theory, predicts the order of arguments. For example, (6i) is part of the lexical structure for *put*. The arguments in the lexicon are presumed to be unordered. If the base rules contain (6ii) and not (6iii), then the PP must follow the NP as in "put the book on the table" but not *"put on the table the book". However, these facts follow independently from the Case adjacency principle, which requires that "where Case is not morphologically realized, a Case-marked element must be adjacent to its Case assigner ..., so that if a verb takes an NP and a PP complement, the former will be closer to the verb" (Chomsky 1986a:82).

- (6) i. *put* [__ NP, PP]
 ii. VP → V NP PP
 iii. * VP → V PP NP

It should be noted that this analysis does not explain the behavior of the arguments of the dative verbs. That is, it does not explain why (7i) is well-formed while (7ii) is ill-formed. The Case adjacency principle, as stated above, predicts they should both be ill-formed. However, having PSRs like (7ii) do not help the situation. With (7iii), both (7i) and (7ii) should be well-formed.

- (7) i. John gave Mary a book
 ii. * John gave a book Mary

iii. VP → V NP NP

Phrase structure rules summarize all possible subcategorizations of lexical items in the language. If a parser is sensitive to lexical structures, then this catalog is superfluous.

The other major overlap of information in the grammar is between the base-rule component and X-bar theory. I will assume the X-bar theory of Chomsky (1986c), his most recent formulation of X-bar theory. The lexical maximal projections are NP (N^{''}), VP (V^{''}), PP (P^{''}), AP (A^{''}). The non-lexical maximal projections are CP (COMP or complementizer, C^{''}) and IP (INFL or sentence, I^{''}). Non-lexical categories have the same structure as the lexical ones. CP, no longer considered to be a defective node, has a structure parallel to all the other maximal projections. The X-bar schema of Chomsky (1986c:3) for English is shown in (8).

- (8) i. $X' = X X''^*$
 ii. $X'' = X''^* X'$

These expand into the generalized maximal projection in Figure 7 where X is the head, the matrix XP (X^{''}) is the maximal projection of the head (X), the XP* directly dominated by the maximal projection is the specifier (SPEC) position, and the XP* dominated by X' is the complement position.

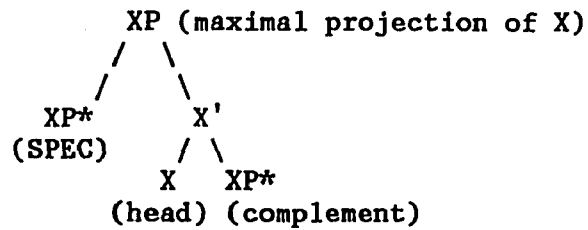


Figure 7. X-bar schema

The S-structure phrase marker for "Who did John hit" is shown in Figure 8. The *wh*-element is the specifier of CP and the helping verb, *do*, is the head of the CP. The head of the sentence (IP) is the inflection, past tense.

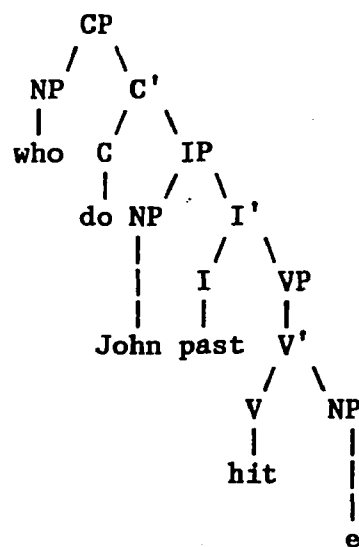


Figure 8. Who did John hit?

This version of X-bar makes the claim that the internal structure of the nonlexical categories parallels exactly the internal structure of lexical categories. Thus the need for PSRs which expand sentences, as illustrated in (9) is obviated. Here, S' represents the "defective" CP category, and S represents the INFL node.

- (9) i. S' → COMP S
 ii. S → NP VP

Ideally, X-bar gives all the internal structure for the maximal projections of all heads. Although Chomsky ultimately wants to eliminate the base-rule component entirely, he stops short of claiming to have achieved this goal.

A number of general principles have been proposed concerning well-formed structures that reduce the recourse to rule systems. One is the projection principle, already mentioned, which states that lexical structure must be represented categorially at every syntactic level. This principle is one of those that contribute to eliminating phrase structure rules--completely, apart from some language-specific idiosyncracies, if the order of various complements and adjuncts is also determined by general principles once the parameters such as head-first, head-last are fixed. (Chomsky 1986a:84)

With all, or almost all, of the base rule component eliminated, there is the problem of how the parser can produce a tree for a string. How can a parser create structure from a collection of principles along with some parametric variation? This is where the licensing conditions of Full Interpretation come in. In a Lexical model, licensing can be used to build structure. Within a Move- α model, licensing conditions act as filters for ill-formed phrase markers. As Abney points out, a principles and parameters parser has to find licensing structures. He

writes, "The parsing problem ... is to determine how each word is licensed" (Abney 1986:2).

2.3.2 THE FUNCTION OF LICENSING PRINCIPLES: CONSTRUCT PHRASE MARKERS OR VALIDATE PHRASE MARKERS

Licensing principles, as discussed in Chomsky (1986a), function as (1) filters on phrase markers and (2) constructors of phrase markers. Chomsky claims that the licensing principles of Full Interpretation apply to the levels of PF and LF. When applied to these levels, licensing principles function as a filter, checking the well-formedness of the phrase marker. However, in his discussion of how a child understands a sentence, Chomsky also uses licensing principles to build S-structure phrase markers from a surface string. The properties of the Lexical and Move- α models are merged by Chomsky in an attempt to construct a parsing model. However, I argue that the properties of these two models are quite different, and a parser that merges the two is not a coherent model.

Within the Lexical model framework, with no D-structure projected by a base component, licensing principles provide the mechanisms necessary to build a phrase marker. They apply to a surface string to build an S-structure phrase marker. Specifically, Chomsky uses licensing to ensure that a VP has a surface subject.

Furthermore, to be licensed, the predicate [*visit e*] must have a subject, the two forming a clause (S); since the subject is not overt, it must be another empty category. (Chomsky 1986a:102)

Within the Move- α model framework, licensing principles serve as filters. After Move- α has applied to the D-structure, the resulting

structures are checked with licensing principles. Phrase markers containing unlicensed nodes are ill-formed.

For the structure ... to be well formed, each element must be licensed. The *wh*-phrase must bind a variable, and each argument must be assigned a θ -role. (Chomsky 1986a:103)

Chomsky (1986a) has licensing serve both functions in a single model. He uses some licensing principles to build an S-structure, and he uses other licensing principles to check the well-formedness of S-structures.

The problem with collapsing the two models in this way is that S-structure is both built and checked by the same principles. In his description of how a child parses a sentence, Chomsky applies the lexical properties of verbs, X-bar theory, the Case adjacency principle, and the licensing principle that predicates must have subjects to the surface string to build an S-structure for (3).

(3) Who was John persuaded to visit?

These assumptions are compatible with a Lexical model. However, he then uses licensing principles to check that the structure is well-formed (Chomsky 1986a:103). This is compatible with the Move- α model. It would be quite surprising if such filters actually found an ill-formed structure -- since, by definition, it will find that everything it has built will comply to its standards. Similarly, Chomsky uses the subcategorization in the lexicon to posit empty categories when no overt arguments are present, and then he uses the licensing principle of θ -role assignment to ensure the well-formedness of the posited structure. Again, the same operation that is used to build the structure is used to test the well-formedness of the structure it has built (assuming that subcategorization and θ -role assignment are represented

together in the lexicon). In order to motivate a set of licensing principles, the Move- α model and Lexical model must be kept distinct.

2.3.3 CHOMSKY'S LICENSING PRINCIPLES

The licensing principles that Chomsky (1986a) mentions follow:

1. Nonmaximal projections "are licensed relative to the maximal projections in which they appear, by X-bar theory" (Chomsky 1986a:100).
2. Arguments or their traces are licensed by θ -role assignment.
 Every complement of a head must be s-selected by it. An element that assigns semantic roles must have recipients in appropriate syntactic positions: For example, the verb *hit* must have an (s-selected) object to receive the role of patient. ... An element that requires a semantic role must be assigned such a role, where this assignment is determined by its grammatical function (subject, object, etc.) and by lexical properties of heads, the grammatical functions being expressed in syntactic configuration. (Chomsky 1986a:93)
3. "A predicate (in particular, a VP) must have a subject, where the notions are syntactically defined ..." (Chomsky 1986a:93). The VP projection of V is licensed by having a subject (Chomsky 1986a:135).
4. An operator is licensed by binding a variable. (Chomsky 1986a:93)
5. A variable is licensed by being strongly bound. (Chomsky 1986a:93)

The claim here is that all maximal projection (CP, IP, VP, NP) have the same internal structure, which is licensed by the X-bar component. The maximal projections themselves remain to be licensed. Chomsky licenses the NP arguments of a verb using the internal θ -role assignment information of the lexicon. The VP is licensed by the presence of a surface subject NP. Binding theory is used to license empty categories and their antecedents. An empty category is licensed by being bound

by its antecedent, and an antecedent is licensed by binding an empty category.

An area for research is the investigation of redundancies among the principles. The logical conclusion of Chomsky's brief sketch of the role of licensing principles is that a single element can be licensed by more than one module. For example, in (10), Chomsky claims that e_i is licensed by being bound by *John*. However, e_i is already licensed by receiving a θ -role from *persuade* which it must transfer to *John*. Similarly, since a variable is licensed by being strongly bound, e_k is licensed by the operator *who*. But since e_k is in the direct object position of *visit*, it is already licensed by receiving a θ -role from *visit*.

(10) who was [John [νp persuaded e_i [e_j to [νp visit e_k]]]]

With this system, all and only non-subject empty categories are doubly licensed.

The licensing principles mentioned in Chomsky (1986a) cover, to some extent, the licensing of CPs (complementizers), IPs (sentences), VPs and argument position NPs. He does not mention how non-argument adverbials and prepositional phrases are licensed. Chomsky (1986a) shows the need for licensing principles, but the details of full interpretation remain to be worked out.

2.3.4 ABNEY'S LICENSING PRINCIPLES

Abney (1986) elaborates on Chomsky's proposals. He claims that "if we could recover licensing relations ... phrase structure would follow immediately" (Abney 1986:2). Abney attempts to formalize a set of licensing conditions by abstracting well-formedness conditions from the

principles and parameters that can be used to build a licensing structure, extending licensing relations to include adjuncts. His licensing structure provides "an intermediate representation which is readily parsed, yet preserves the spirit of the abstract grammar" (Abney 1986:1). Once these intermediate structures are recovered, "phrase structure would follow immediately" (Abney 1986:2).

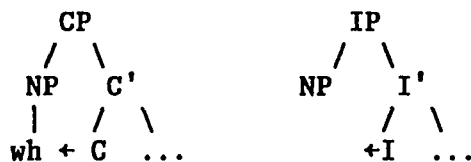
For Abney, the core case of licensing is that of the verb θ -marking its arguments, which he uses as a metaphor for all licensing. Licensing is the generalization of θ -role assignment to all phrase structure, and he develops his licensing relations on the model of θ -role assignment. Just as θ -roles are assigned to sisters by the head, a head licenses all and only its sister nodes. Thus licensing is a strictly local operation. A result of this stipulation is that binding is not a licensing relation. Just as θ -role assignment inheres in the lexicon, so is the ability to assign licensing relations stipulated in the lexicon. Just as the Theta Criterion requires that each argument receive exactly one θ -role, so is licensing a unique relationship. A node can only be licensed once, in contrast to Chomsky's system. Abney defines licensing as a ternary relation. Just as θ -role assignment holds between (1) a θ -role assigner, (2) an argument, and (3) a θ -role, licensing holds between (1) a licensor, (2) a licensee, and (3) a licensing relation. The only place where Abney seems unable to follow his metaphor is that the directionality of each licensing relation is stipulated in the lexicon.

Abney needs the following four licensing relations to get a parse. I also include the directionality of the licensing relations for English.

1. **Functional Selection:** The nodes INFL, COMP, and DET dominate function words which license the maximal projections to their right. INFL licenses VP, COMP licenses IP, DET licenses NP (Abney assumes a framework where a determiner phrase (DP) is a maximal projection which dominates DET NP).



2. **Subjecthood:** Subjecthood is assigned figurationally to a maximal projection to the left. INFL licenses its subject, COMP licenses its "subject" (a fronted *wh*- element), the genitive *'s* licenses its "subject."



3. **Modification:** This is called an "anti-relation." Basically, an adjunct looks for a c-commanding head that can license it.

Modification covers adjuncts, including locative, temporal, and manner adjuncts (ADVP, PP, CP), purpose clauses, pre- and post-nominal AP's in the noun phrase, relative clauses, etc. (Abney 1986:7)

4. **θ -assignment:** This is internal θ -role assignment. Verbs assign θ -roles to noun phrases, preposition phrases, and sentential complements to the right. Nouns and adjectives assign θ -roles to prepositional and sentential complements.

Thus, INFL, COMP, DET (including the genitive marker 's), N, P and V are heads and licensers. Each gives Subjecthood to sister nodes to its left and Functional selection, Modification, and Theta-assignment to nodes to its right.

The most marked contrast between Abney's and Chomsky's system is that binding is not a licensing relation. Abney licenses fronted *wh*-elements locally, so that a *wh*- antecedent is the subject of the COMP node.

Abney explicitly states that licensing must be part of a lexical structure, and he also states that "only words appear in the lexicon" (Abney 1986:7). If INFL licenses the subject of a sentence, then the lexicon will have to contain many elements not usually classified as words. For example, the feature (+ past) will have to be a lexical item in order for *John* to be licensed in "John went to the store".

2.4 CHAIN FORMATION IN THE TWO MODELS

A fundamental difference between the Move- α and Lexical models is how chains are formed in each model. In the Move- α model, chains are the residue of successive applications of Move- α . Thus, in a Move- α model, chains are already fully formed in the S-structure. I will call chains resulting from successive applications of Move- α inherent chains because any S-structure in a Move- α model already contains fully-formed chains.

In the Lexical model, since there is no movement, chains are constructed by the parser as it builds the S-structure. The parser makes use of lexical structures along with independent principles of the

grammar. I will call the chains built by a parser in the Lexical model constructed chains.

2.4.1 FORMING CHAINS WITH MOVE- α

When Chomsky (1986a:95) defines a chain, he defines an inherent chain. Inherent chains are the residue of Move- α , the result of successive operations of Move- α . They are data structures which show their derivational history. These complex data structures are fully formed in the S-structure. Chomsky describes chains as follows:

... a chain is the S-structure reflection of a "history of movement," consisting of the positions through which an element has moved from the A-position it occupied at D-structure; we include here the vacuous case of the single-membered chain of an element that remains in its D-structure A-position. (Chomsky 1986a:95)

In effect, every argument is assigned a distinct index in the D-structure. Every A-position (argument position) in D-structure is associated with a single-membered chain. Upon application of Move- α , the index remains with the trace in the original position and is also copied, along with all the features of the constituent, onto its new position in the phrase marker.

If a constituent is not moved in the mapping between between D-structure and S-structure, then its S-structure chain has a single member. Since both arguments in (11) remain in their D-structure positions, the S-structure contains two single-membered chains: [*Mary_i*] and [*her_j*].

(11) *Mary_i saw her_j*

When Move- α operates on a constituent and it moves, the index on the original position is copied onto the new position. If Move- α has applied

once, the resulting chain structure has two members, as in (12), where the non-single-membered chain indexed structure is $[who_i \dots e_i]$. This chain consists of one link.

(12) $[CP\ who_i\ [IP\ did\ Mary_j\ see\ e_i]]$

This same process is repeated upon each application of Move- α . The non-single-membered chain indexed structure in (13), $[who_i \dots e_i \dots e_i]$, is the result of two successive applications of Move- α . This chain contains two links; the first consisting of *who* and the intermediate trace, and the second consisting of the intermediate trace and the D-structure position of *who*.

(13) $[CP\ who_i\ [IP\ did\ Mary_j\ persuade\ John_k\ [CP\ e_i\ [IP\ PRO_l\ to\ see\ e_i]]]]$

The Theta Criterion, given below, ensures that the lexical structure and the D-structure have exactly the same number and type of arguments:

The Theta Criterion:

- a. Every argument must be assigned a θ -role.
- b. Every θ -role must be assigned to an argument.

(Lasnik & Uriagereka 1988:2)

A possible assumption as to how the Theta Criterion is satisfied is to posit that unique indices are assigned in the mapping between lexical structures and D-structure. If this is the case, the resulting D-structure would have all arguments indexed, including overt anaphors and pronouns. Assuming that these indices remain attached to arguments in the mapping from D-structure to S-structure, the resulting S-structure would likewise have all arguments indexed. The S-structure of (14), with its inherent chains, would be (15).

(14) What did Mary persuade John to buy for himself?

- (15) [CP $what_i$ [IP did $Mary_j$ persuade $John_k$ [CP e_i [IP PRO_l to buy e_i for himself $_m$]]]]]

The S-structure (15) contains five chains: a three-membered or two-link chain (16i) and the four single-membered chains (16ii) through (16v).

- (16) i. [$what_i \dots e_i \dots e_i$]
 ii. [$Mary_j$]
 iii. [$John_k$]
 iv. [PRO_l]
 v. [$himself_m$]

Of these five chains, (16iv) through (16v) require reindexing in the process of determining the antecedents of PRO and *himself*. These re-indexed structures are then filtered by Binding Principle A. In addition, the reindexed structures containing PRO are further filtered by Control Theory. After reindexation and filtering, (14) will have the S-structure (17).

- (17) [CP $what_i$ [IP did $Mary_j$ persuade $John_k$ [CP e_i [IP PRO_k to buy e_i for himself $_k$]]]]]

Thus (16i) and (16ii) are both well-formed chains in the S-structure and are well-formed binding structures too. The remaining well-formed single-membered chains (16iii), (16iv), and (16v) constitute ill-formed binding structures.

An alternative way to view S-structure chains is to posit that indices are only assigned by movement and are the result of feature-copying. If coindexation is the result of movement, single-membered chains receive no index. From this point of view, the S-structure of (14) is (18).

- (18) [CP $what_i$ [IP did $Mary$ persuade $John$ [CP e_i [IP PRO to buy e_i for himself]]]]]

This way, all well-formed inherent chain structures are also well-formed binding structures.

2.4.2 DETERMINING FEATURES OF EMPTY CATEGORIES

There are several possibilities for how the features or identity of empty categories in inherent chains are identified. These include functional determination of empty categories, free assignment of empty categories, and intrinsic determination of empty categories.

Chomsky (1982) explores the possibility that the identity of an empty category can be determined configurationally. This is known as the Functional Determination of Empty Categories. The theory of functional determination assumes that chains are inherent, but that, since the three types of empty categories are in complementary distribution, their type or features can be determined configurationally.

This property of ECs surely implies that there is only one type of EC and that the status of a particular occurrence of an EC as PRO, NP-trace, or variable is functionally determined, that is, determined by the EC's role in derivations and representations. The partitioning, which is otherwise mysterious, will then follow. (Chomsky 1982:34)

With functional determination, the S-structure contains inherent chains, but the inherent chain contains empty categories with unspecified features. An empty category's features are determined by the following algorithm (taken from Chomsky (1982:34-35)):

1. An EC is a *variable* if it is in an A-position and is locally A'-bound.
2. An EC in an A-position that is not a variable is an *anaphor*.

3. An EC [in an A-position] that is not a variable is a pronominal if it is free or locally A-bound by an antecedent with an independent θ -role.

Step 1 determines whether an EC is a *wh*-trace. Step 2 determines whether an EC is either an NP-trace or PRO. Step 3 determines whether an EC is PRO.

The functional determination algorithm poses an insoluble ordering problem. The algorithm relies on the fact that the empty categories are already either bound or free. The bindings required for steps 1 and 2 are chain structures and so are not available to the algorithm before binding applies. The binding required for step 3, however, is not a chain structure. Whether PRO is bound by an antecedent with an independent θ -role or is free is the filtered result of free indexation. That is, binding has already applied to the structure. Since binding structures are required in order for the algorithm to operate, this algorithm must apply after binding. However, as formulated, binding cannot operate without having access to an EC's features. The two systems cannot apply simultaneously because each requires the output of the other in order to operate.

Chomsky claims to have abandoned the functional determination of empty categories.

Note that this conclusion eliminates much of the motivation for functional determination of empty categories ... I have been assuming throughout that the approach is incorrect and that the features assigned to empty categories (like other features) do not change in the course of a derivation; see Chomsky 1985⁴ for further discussion. (Chomsky 1986c:57)

⁴ This reference is Chomsky (1986a) in this bibliography.

The second possibility, which avoids the circularity of Functional Determination is proposed in Brody (1984). In this theory, as in functional determination, Move- α leaves a chain structure but the empty categories contain no features. Instead of determining the features using an algorithm, Brody suggests that features are freely assigned to the empty categories. The claim is that the principles of the grammatical theory will independently force the empty categories to have certain feature structures in order for the string to be well-formed. I follow Lasnik and Uriagereka (1988) in calling this hypothesis the theory of Free Assignment. As with free indexation, the theory of free assignment vastly overgenerates and all ill-formed structures have to be filtered out by independent grammatical principles such as binding.

The third possibility, which Lasnik and Uriagereka (1988) call intrinsic determination of empty categories, is the theory which Chomsky currently claims to follow, as indicated by the above quoted passage in Chomsky (1986c:57). Lasnik and Uriagereka characterize this theory below:

Under this theory, an empty category is what it is: if we pull a PRO out of the lexicon and put it in D-Structure, then it is a PRO; if we generate an empty category by applying *Wh*-Movement, the result is a variable; and so forth.
(Lasnik and Uriagereka 1988:66)

Since Chomsky currently assumes intrinsically determined chains, I too will assume intrinsically determined chains upon discussing inherent chains unless I specifically indicate otherwise.

2.4.3 FORMING CHAINS WITH A LEXICAL PARSER

Constructed chains are chain indexed structures which are licensed by the parser. Chomsky's (1986a) parsing model licenses constructed chains

in his description of a Lexical parser (see "The Lexical Model in "Knowledge of Language"" on page 28). While building the S-structure, Chomsky's parser builds an S-structure containing empty categories based on the lexicon, X-bar theory and its head-first parameter, the projection principle, the licensing principles of full interpretation, and the case adjacency principle. With these, the parser builds the unindexed (chainless) structure (19) (Chomsky's example <93>).

(19) who was [John [νP persuaded e_i [e_j to [VP visit e_k]]]]

The parser then supplies the bindings and identities of the empty categories based on the licensing principles of full interpretation, θ -theory, and case assignment.

The relative advantages and disadvantages of attempting to functionally determine, freely assign, or intrinsically assign features of empty categories are not relevant to a discussion of a lexical model, because all these variants have inherent chains as a premise. I have described them because Chomsky makes use of some form of these various methods for determining the identity of empty categories in his argumentation for the existence of a D-structure in Chomsky (1986a). We will therefore refer back to these methods in "'Stubborn' versus 'Illegal'" on page 78.

2.4.4 TYPES OF CHAINS

A cause of confusion in Chomsky's (1986a) argument for the existence of a level of D-structure is his failure to consistently distinguish among the various types of chains. I have described four possible types of chains:

1. **Intrinsically Determined Chains** are inherent chains in the S-structure. The empty categories in intrinsically determined chains have full feature specification in the S-structure.
2. **Functionally Determined Chains** are inherent chains in the S-structure. The empty categories in functionally determined chains have no feature specifications. Features are determined by recourse to an algorithm.
3. **Freely Assigned Chains** are inherent chains in the S-structure. The empty categories in freely assigned chains have no features specifications. Features are freely assigned to the empty categories. Any resulting ill-formed chains are filtered out by independent principles.
4. **Constructed Chains** are licensed by the parser as it builds an S-structure based on the lexicon and independent principles of the grammar. As the parser builds the chain, it has to determine the location and feature structure of each empty category.

The first three types of chains are variants of the inherent chains of a Move- α model. The fourth type of chain is the constructed chain of a Lexical model.

2.5 CHAIN STRUCTURES AND BINDING STRUCTURES

Chain indexed structures are each associated with exactly one θ -role. Binding indexed structures are associated with one or more θ -roles. When an indexed structure is associated with one θ -role, the binding structure has an identical corresponding chain structure. When they are associated with two θ -roles, then there is no chain indexed structure

which corresponds to the binding indexing structure. Free indexing will also assign indices to PRO, creating a structure which can be associated with a single θ -role (when PRO gets an arbitrary interpretation) or more than one θ -role (where PRO is assigned an antecedent).

In the Move- α model it appears that the similarities between the two kinds of indexed structures are accidental. The theory does not explain why some binding structures are associated with a single θ -role while others are associated with two θ -roles. Since free indexing and binding principles are able to reproduce chain indexed structures, one wonders why the chain building mechanism is posited at all.

2.5.1 THE MOVE- α MODEL -- BINDING AS A FILTER

In the Move- α model, NPs are freely indexed in the S-structure. As a result of free-indexing, ill-formed binding structures can and must be assigned. The role of binding is that of a filter. Ill-formed coindexation structures are filtered out by the Binding Principles.

What then is the relationship between the output of movement and the input to the binding principles? Binding Theory, which determines which overt NPs cannot or must have antecedents, ensures that the antecedents of anaphoric and variable empty categories are within the proper structural configuration. Binding cannot not operate directly on chain indexed structures which are the residue of Move- α . If it did then lexical anaphors, like *himself* in (14), would never be assigned an antecedent. Free indexing, which either occurs on S-structure or LF (or both), supplies the indexed structures on which binding operates.

There are at least three possibilities for how free indexing interacts with pre-existing chains:

A. The indexing process can be blind to S-structure chains and freely index all NPs randomly.

B. Assuming that indices are only assigned where movement has occurred, and that (19) and not (17) is the input to S-structure, the indexing process can assign indices only to NPs which do not already have an index.

C. The indexing process can be sensitive to some other aspect of the grammar. For example, as Lasnik and Uriagereka (1988) suggest, only A-positions are indexed. This would result in the entire chain receiving a single index.

A problem with Theory A, which freely indexes without being sensitive to pre-existing chains, is that all well-formed output will have to reproduce the chains. Thus these complex data structures have to be built twice: once by movement and again by free indexing. If this unconstrained version of free indexing is applied to the simple sentence (20), the resulting S-structures are the five structures in (21). (In these examples, I will use integers to indicate binding structures in order to differentiate them from inherent chain structures which are indicated with lower case letters.)

(20) Who did Mary see?

- (21) i. who₁ did Mary₁ see e₁
 ii. who₁ did Mary₁ see e₂
 iii. who₁ did Mary₂ see e₁
 iv. who₁ did Mary₂ see e₂
 v. who₁ did Mary₂ see e₃

Well-formed chains are rebuilt by (21i) and (21iii). New ill-formed chains are introduced by (21ii), (21iv), and (21v). An additional problem with Theory A is that once the identical structure is reproduced,

the Binding principles would apply to all structures equally. However, Binding Theory, as currently formulated, operates using chain indexed structures. They are explicitly mentioned in Binding Principle C.

Binding Principle C: An R-expression is A-free (in the domain of the head of its chain).

Theory B, which indexes only NPs which have not moved, has the result that the head of a chain can only be an antecedent to its trace since chains are exempt from coindexation. Theory B would build only one S-structure indexed structure for (20), shown in (22).

(22) who_j did $Mary_1$ see e_j

However, by theory B, a chain cannot contain the antecedent for PRO or a lexical anaphor since the chain would be indexed by Move- α and PRO or a lexical anaphor would receive a different index via the process of free-indexation. If theory B were adopted, then PRO in (23) and *himself* would not be coindexed with the moved *wh*- element:

(23) who_j did John tell e_j [PRO₁ to leave]

(24) who_j did John tell e_j [PRO to behave himself]

Theory C differs from Theory B in that it doesn't require that a chain not receive an index. Within Theory C, the head of a chain can be an antecedent to something other than its trace. If we apply the Lasnik and Uriagereka (1988) version of Theory C to (20), the output is the two indexed structures in (25). In this formulation, indices are assigned to all and only A-positions. An entire chain is assigned a single index because well-formed chains contain exactly one A-position.

(25) i. who_j did $Mary_1$ see e_{j1}
 ii. who_j did $Mary_1$ see e_{j2}

Henceforth, when I refer to free indexation within the Move- α model, I am referring to the version of free indexation proposed in Lasnik and Uriagereka (1988:82). This account of free indexation is superior to Theory A because it does not require chains to be reinvented by free-indexation. It is superior to Theory B since chains need to be coindexed with PRO and lexical anaphors and can be coindexed with pronouns. An additional advantage to this version of free-indexation is that it ensures that all and only NPs receive a single index, since every A-position is either filled by an NP or associated with an NP via an inherent chain in the S-structure.

2.5.2 REDUNDANCY BETWEEN INHERENT CHAINS AND BINDING THEORY

In a Move- α model of English, binding theory operates on inherent chains, PRO, and overt anaphora, pronominals, and r-expressions in LF. However, inherent chains are fully formed in the S-structure and, by the time LF is reached, are guaranteed to be well-formed. It is therefore redundant to submit inherent chains to the further scrutiny of binding theory at the level of LF.

According to Chomsky (1986a) Binding theory applies at LF. A definition of binding theory follows:

...binding theory [is] the subtheory of UG that is concerned with the principles that govern the relations between referential dependents such as reciprocals, reflexives, and pronouns on the one hand, and their possible antecedents on the other. Let us express the antecedent-pronoun relation by co-indexing. Then, an element X *binds* an element Y in its domain if X and Y are co-indexed, where we assume that Y is not contained within the category X. The principle <10> can now be understood to state that a pronoun can bind a distinct element X only if X is its trace.

[<10> A pronoun cannot take an element of its domain as its antecedent]

(Chomsky 1986a:77)

Binding theory accounts for the co-indexation of anaphors (both overt anaphors and NP-traces), pronominals (which are always overt pronouns in English), r-expressions (both overt r-expressions and traces of variables), and for the distribution of the pronominal anaphor PRO.

Inherent chains are the output of successive applications of Move- α between the levels of D- and S-structure. These chains are subjected to the independent principles of the grammar at both D-structure and S-structure. At D-structure, single membered or vacuous chains are submitted to the tests of the projection principle, the Theta Criterion, and inherent case assignment. At S-structure, single membered and complex chains are submitted to the tests of the projection principle, bounding theory to filter out subjacency violations, structural case assignment, and government. Although I am assuming intrinsically determined chains, whatever kind of inherent chain theory is chosen, the fully indexed chain is submitted to these tests which comprise all of the filters of ill-formed structures posited in the grammar to date, except for binding, before LF is reached.

In order for a chain to be filtered out by Binding theory, it must disobey one or more binding constraints. Below I show that an otherwise well-formed inherent chain cannot be an ill-formed binding structure. That is, well-formed chains are well-formed binding structures. If this is true, then it is redundant to have Binding operate on chain structures.

- | |
|--|
| <p>Principle A. An anaphor is A-bound in a local domain.</p> <p>Principle B. A pronominal is A-free in a local domain.</p> <p>Principle C. An r-expression is A-free (in the domain of the head of its chain).</p> |
|--|

Figure 9. Binding Theory

Principle A states that an anaphor is A-bound in a local domain. In order to constitute a Principle A violation, an NP-trace (anaphor) would have to be A-free in its local domain. This violation can be manifested in two ways: (a) the NP-trace is not bound in the local domain or (b) the NP-trace is A'-bound in the local domain. The antecedent for an NP-trace can be one of three types of anaphors: a lexical r-expression, an NP-trace, or PRO (since PRO is sometimes anaphoric and sometimes pronominal). I do not cover the theoretical cases in which the antecedent is not an anaphor because I assume, as does Chomsky (1986c), intrinsically determined chains. Thus there are six logical possibilities for Principle A violations.

1. NP-trace is not locally bound
 - a. antecedent = r-expression
 - b. antecedent = NP-trace
 - c. antecedent = PRO
2. NP-trace is locally A'-bound
 - a. antecedent = r-expression
 - b. antecedent = NP-trace
 - c. antecedent = PRO

Principle A violation type 1a is illustrated in (26).

(26) * $John_i$ thinks [CP that [IP I was arrested e_i]]

An S-structure of the form shown in (26) would never reach the binding module in a Move- α model. The inherent chain in (26), [$John_i \dots e_i$] constitutes a subjacency violation (however subjacency is ultimately formulated) because the NP has crossed two bounding nodes: IP and CP (or equivalently, S and S' in many discussions on subjacency). The same ill-formed S-structures would occur if the antecedent is PRO or NP-trace (possibilities 1b and 1c). The structures would also contain subjacency violation. Therefore, all type 1 Principle A violations would be filtered out or blocked by the constraints of Bounding theory before the level of LF.

The second possibility for a Principle A violation in an inherent chain is that the NP-trace is A'-bound. An example of Principle A violation type 2a, with a lexical r-expression as antecedent, is shown in (27).

(27) * [CP $John_i$ [IP I was seen e_i]]

The type of structure illustrated by (27) would be blocked by the Case filter. The requirement of the Case filter is that phonetically realized NPs must have Case. Since neither position in the chain in (27) is a case marked position, this type of ill-formed chain would never reach LF and Binding theory.

Other Principle A violations of type 2 would be structures having an NP-trace A-bound by an NP that does not require Case assignment. That is, the binder is an NP that is not phonetically realized: a trace (violation 2b) or PRO (violation 2c). A chain with PRO as antecedent

to an NP-trace will be blocked by the Theta Criterion, which is formulated as a property of chains, because both elements are assigned independent θ -roles in D-structure.

The final possible violation of Principle A, type 2b, is that the NP-trace is bound by another NP-trace which is in a non-argument position. These include a possible derivation for *tough*-movement predicates through two successive applications of Move- α , as in (28).

(28) * $John_i$ is easy [e_j [PRO to please e_j]]

The chain in (28) violates Principle A because the lower NP-trace, an anaphor, is bound by an intermediate trace in an A'-position. Thus the anaphor, the lower e_j , is A-free in its local domain. However, this derivation is also blocked by the restriction that each chain be associated with only one case. In (28), the lower trace is assigned objective case by *please*, and *John* receives nominative case from INFL. Again, this chain would be blocked before it could be submitted to the binding filter.

Principle B will not eliminate any inherent chains in English because English does not have a pure pronominal empty category. Therefore, in English, no inherent chains can be affected by Principle B.

An inherent chain violates Principle C if a *wh*-trace is A-bound.⁵ Possible binders for *wh*-trace are: a lexical *wh*-element,

⁵ The second clause of Principle C adds a further restriction to the binding of an r-expression. This clause states that *wh*-traces are A-free in the *domain of the head of its chain*. This statement implies that binding is, in fact, sensitive to the differences between inherent chain structures and the results of free indexing. Since we are, at this point, discussing chains without free-indexation, this part of Principle C has no relevance to this discussion.

another *wh*- trace, or an empty operator. If any of these elements appear in a position which is assigned an external θ -role, the derivations will be blocked by the Theta Criterion. Such a violation of Principle C is illustrated by (29).

(29) [*IP* *who_f* kissed *e_f*]

(29) constitutes a Theta Criterion violation because the single chain is associated with two θ -roles: both the internal and the external arguments of *kiss*. The other logical possibility is to have the antecedent of the variable trace move to an A-position when no external argument is assigned, as in (30).

(30) * *who_f* is likely [*e_f* [*PRO* to kiss *e_f*]]

The structure in (30) is ill-formed at S-structure because a single chain is associated with two distinct Cases. *Who* receives nominative Case from INFL and the lower trace is assigned objective Case by *kiss*.

On the other hand, the indices on A-positions are assigned freely in the mapping from S-structure to LF. Freely assigned indices on A-positions are guaranteed to create many ill-formed structures which violate binding theory. Thus binding theory is a very necessary filter for throwing out ill-formed indexed data structures.

Formulations of binding theory define principles which govern antecedent-anaphor relations for both lexical anaphors and empty categories. Binding theory does not differentiate between what is input to it -- between well-formed inherent chains and the possibly ill-formed output of free indexation, treating both data structures equally.

There are various alternatives for eliminating this redundancy from the grammar. The first is to reformulate binding theory so that it only

covers overt anaphors, pronominals, and r-expressions. In so doing, binding would lose the generalization that it works for both overt and empty categories, and the grammar would lose the relation between the distribution of overt and empty anaphora and r-expressions. This does not appear to be a feasible solution to the problem. A second possible solution, one currently followed by some parsing models, is to eliminate inherent chains by freely assigning indices to all possible anaphora. Then binding theory will apply equally to all the resulting data structures. While this approach may be feasible, it is wasteful and there is no reason to follow it. A third approach, the one I take, is to use the binding principles to license constructed chains and to do the coindexing of overt anaphors in the S-structure. In this approach, binding theory is used to license structure, not to filter out ill-formed structures.

3.0 MOTIVATING D-STRUCTURE

Over the years, Chomsky has presented a number of arguments for why there is no motivation for eliminating Move- α and the level of D-structure:

1. The argument in *Lectures on Government and Binding* (Chomsky 1981) has two parts. The first is that the relationship between a model of grammar with Move- α in the syntax and one without Move- α in the syntax is one of "virtual interchangeability." I outline this argument in Figure 10 on page 69. The second part of this argument is that, given the equivalency of the two models, the idea that eliminating Move- α from the syntax will simplify the grammar is fallacious. This is outlined in Figure 11 on page 74.
2. One argument in *Knowledge of Language* (Chomsky 1986a) is that the motivation for eliminating phrase structure rules is not logically equivalent to any motivation for eliminating Move- α .
3. A second argument in *Knowledge of Language* is that there exists a pair of sentences with identical S-structures -- but that one is ill-formed while the other is well-formed. Since the S-structures are identical, then there must be another level where they differ.

Chomsky also asserts that "the choice of levels of representation and their properties is an empirical matter, to be verified by their role in explanation" (Chomsky 1986a:157).

I will argue that Chomsky's arguments do not motivate a D-structure level. I will then argue that if S-structures can be built from the interaction of the lexicon with the components of a principles-and-

parameters framework, then the level of D-structure is not motivated within a principles-and-parameters framework.

3.1 CHOMSKY'S LGB ARGUMENT

The first part of Chomsky's LGB argument (outlined in Figure 10 on page 69) is that a grammar which generates S-structures from D-structures through the transformation Move- α (Theory Ia) is virtually equivalent to a grammar which base generates S-structure directly (Theory Ib).

According to Chomsky, the two theories share (a) interpretive rules of construal binding antecedents to lexical anaphors (pronouns, reflexives, and reciprocals) and binding extraposed phrases and (b) rules of control which determine the referential properties of PRO. The two theories diverge on the issue of where the relation between antecedents and non-pronominal empty categories (anaphors or "NP-traces" and variables or "*wh*-traces") are established. In Theory Ia the relation is established in the syntax by the transformational rule Move- α , and in Theory Ib the relation is established in the logical form⁶ with interpretive rules.

Chomsky does not commit himself on whether there is an empirical difference between Theories Ia and Ib. However, he does assert that if there is an empirical difference between the two, it will be "rather subtle."

⁶ I propose to establish the relation between antecedents and non-pronominal empty categories at S-structure, not LF-structure. However, I do not think that this makes any difference within Chomsky's line of argumentation.

1. Theory I postulates S-structures with empty categories.
The theory has two variants:
 - a. Theory Ia. Move- α forms S-structures from base generated D-structures.
 - b. Theory Ib. S-structures are base generated.
2. Theories Ia and Ib share:
 - a. rules of construal
 - b. rules of control
3. Ia and Ib differ on treatment of antecedent-trace relations:
 - a. Ia establishes the relation in the syntax with Move- α .
 - b. Ib establishes the relation in the LF with interpretative rules.
4. Choosing between Theories Ia and Ib:
 - a. Perhaps there is no empirical difference between them and they are "two realizations of the same more abstract theory" (Chomsky 1981:90).
 - b. If there is an empirical difference, it "will surely be rather subtle."
 - 1) Ia and Ib will differ in
 - a) how they handle idioms.
 - b) how they differentiate PRO and trace.
 - 2) Indirect support for Ia:

The assumption has consequences ... which were established on the basis of a property of D-structure. Insofar as these consequences are conceptually or empirically desirable, we have indirect support for Theory Ia. (Chomsky 1981:91)
5. Conclusion:

The virtual interchangeability of Theories Ia and Ib is clear within the framework of trace theory. While it should not be assumed too quickly that the two theories are empirically indistinguishable variants -- my suspicion is that they are not -- nevertheless the problem of choosing between them (if it exists) is a fairly marginal one ... (Chomsky 1981:90)

Figure 10. Part 1 of the Chomsky (1981) argument

Chomsky indicates how indirect support could be provided for Theory Ia (see 4.b.2 of Figure 10) by alluding to possible desirable consequences which are established on the basis of a property of D-structure

(Chomsky 1982:45-46). As Chomsky points out, the argumentation for choosing between Ia and Ib would be theory internal since the two systems work within the same grammatical theory. The desirable consequence that Chomsky refers the reader to here is that the θ -criterion and the projection principle, taken together, predict that each chain contains exactly one θ -position. This is the result of the combination of (i) if an NP is an argument, then the right-most element in a chain is in a θ -position and (ii) all other elements in the chain are non- θ -positions. (i) is a consequence of the projection principle applying at D-structure where all chains are of length 1. (ii) is a consequence of the θ -criterion--because having more than one position in a chain θ -marked would constitute a θ -criterion violation.

Chomsky concludes there is little or no point in attempting to choose between Theories Ia and Ib -- even though he suspects that the two are not indistinguishable variants.

Just as Chomsky argues that any desirable consequences of Theory Ia can be used for its indirect support, conversely any undesirable consequences of Theory Ia should provide indirect support for Theory Ib. An undesirable consequence of Theory Ia is that it contains an underlying assumption that there is a canonical form for each sentence which is explicitly represented during the course of its derivation. Both a Move- α model and a Lexical model within a principles-and-parameters framework require an S-structure. The S-structure is defined by both models to include the assumption that the order of constituents in the S-structure matches the observed order of constituents in the

sentence. That is, lexical arguments and adjuncts in the phonetic form are in the same linear order as they are in the S-structure.

Unlike the Lexical model, the Move- α model incorporates a D-structure level where the order of constituents differs from that in the perceived sentence. Chomsky defines D-structure as

a direct projection of lexical structure subject to other licensing conditions (e.g., those that guarantee the presence of subjects of predication). (Chomsky 1986a:161)

D-structures, for English, would consist of: subject, predicate, argument(s), and optional adjuncts -- in that order. Let us call this the indicative form of the sentence. Thus an underlying assumption of a system with D-structure has the consequence that there is a canonical form of a sentence. This canonical form is the same as the indicative form. If a canonical form is assumed, then there will be cases (passive, predicates with no external θ -role assignment, *wh*-questions, etc.) in which the observed order of constituents will conflict with the canonical order. The question is, then, is it conceptually or empirically desirable to have a canonical order of constituents? Is there any motivation for another level, D-structure, in which the order of constituents is not the order of constituents observed in the phonetic form?

A system using D-structure assumes that the canonical form of a sentence, the indicative form, has a special status; that of the basic structure from which all permutations are derived by the transformation Move- α . In this way, a theory of movement retains the Harrisian idea of a central kernel (canonical) sentence to which other sentences are related. However, within the principles-and-parameters framework, the

order of arguments in the indicative form is independently motivated by the separate components of the theory. It is generally agreed that lexical structures do not contain ordered arguments, rather they indicate unordered arguments. Case theory, and X-bar theory provide all information about order and hierarchy. If ordering is motivated by independent principles, then there is no reason posit a level where the order of elements deviates from the order of the phonetic string.

In the Lexical model, the basic form is the lexical representation, not the canonical sentence. All sentences would have the same status, in that the lexical structure for each verb is filled, either by phonologically realized or by null elements.

Occam's razor says not to use more descriptive devices than are necessary. Each level, principle, and parameter is motivated to the extent that it plays a role in the description or explanation of some observed phenomena. In order to motivate a level of D-structure, evidence must be provided to prove that there is some reason for the order of elements to differ at an underlying level. If a theory with D-structure is to be permitted, the D-structure theory must provide a range of descriptive or explanatory mechanisms which go beyond a theory which makes no use of the extra D-structure level. These mechanisms must also play a role in the description of some observed phenomena.

If we contrast (a) a parser system which, in its simplest formulation, has no D-structure with (b) a Move- α system, we can choose between the two alternative theories.

(a) The parser system argues that the lexicon has a canonical structure but no canonical order for complements. The arguments of

any lexical item are unordered. All order is motivated independently by general principles of UG.

- (b) The Move- α system maintains that there is a canonical order for arguments which has to be represented during the derivation.

If it is true, as Chomsky seems to claim, that theories Ia and Ib produce grammars which generate the same language, then it is unnecessary to posit an extra level of description, that of D-structure, simply to represent a canonical argument structure which is not made any use of anywhere else in the grammar.

Chomsky's D-structure theory is equivalent to having arguments ordered in the lexicon. Arguments against having ordering of arguments in the lexicon can also be applied to a theory with a level of D-structure.

The second part of the LGB argument, outlined in Figure 11 on page 74, is that if theories with and without D-structure are logically equivalent, then the elimination of Move- α from mapping between D-and S-structures does not simplify the grammar. First Chomsky points out that Move- α is not eliminated from the grammar, as it is still needed to operate between S-structure and LF. Secondly, the operation of Move- α between D-and S-structures would have to be replaced with new interpretive rules "with exactly the properties of the rule Move- α ." Thus the complexity to the grammar introduced by Move- α at that level is equivalent to the complexity introduced by the new interpretive rules.

This argument can only hold if Theories Ia and Ib are logically equivalent. If Ia and Ib generate different languages, then complexity

1. The conclusion of the above argument holds for the question of whether syntactic theory should:
 - a. derive S-structure from D-structure
 - b. express all syntactic properties in a single level of representation
2. The EST

can be understood in either way, given the association between a function chain in S-structure and the derivations of which it can be regarded as a projection.
3. The questions to be asked are:
 - a. What are the properties and scope of Move- α ?
 - b. Do "certain arguments turn on the existence of D-structure"?
4. The argument that Ib is superior because the class of transformations is reduced to zero is fallacious:

While Theory 1b eliminates the rule Move- α from the syntax (leaving it, however, as an option in the PF- and LF- components and presumably in the marked periphery outside of core grammar), it does so by adding a new category of LF-rules, with exactly the properties of the rule Move- α . Thus no argument based on conceptual economy applies, at least if the rule Move- α has different properties from the interpretive rules shared by the two theories, as the preceding discussion indicates. (1981:90-91)

Figure 11. Part 2 of the Chomsky (1981) argument

is not an issue. Thus, for the purposes of this discussion, we will for the moment assume the equivalency of Ia and Ib.

The question to be asked is: Is it true that the only difference between Ia and Ib is the level at which Move- α , or the equivalent interpretive rule, applies?

Chomsky claims that new interpretive rules and Move- α are equivalent in terms of complexity. This is not true. The Move- α model needs more than just Move- α to achieve interpretation of "traces". It uses bounding theory, Move- α , and chain theory. In addition, to interpret overt anaphors and PRO, the model requires a coindexing mechanism and

binding theory. Binding theory, then, not only checks the output of free indexing but also redundantly checks the indexing on chains.

I will illustrate my point below in an analysis of the sentence

(1) Who did John see?

First the base produces a D-structure in (2).

(2) John PAST see who

Move- α applies, constrained by bounding theory to obey subadjacency conditions. A chain is formed by this movement, having the result of coindexing an empty category anaphor to its antecedent, forming the S-structure (3).

(3) who_i PAST+do John see e_i

Then, in either in the mapping from S-structure to LF, or in the LF component, an operation occurs which freely attaches indices to the NPs. Just what operation does this mapping has not been defined. It is not usually included as a component of the grammar, but since it alters structures created by the grammar, it must be part of the grammar and contribute to its complexity. Presumably this indexing operation does not destroy the indices on already existing chains. Thus, (3) would create two chain structures which would then be submitted to the Binding module.

(4) i. who_i PAST+do John_i see e_i

ii. who_i PAST+do John_j see e_i

(4i) violates binding principle C on two counts. Both the variable (e_i) and *John* are illegally bound. Since the indexing in (4ii) is a chain formed by Move- α and, since it has made it to LF, presumably it obeys

all the constraints of Case theory, Theta theory, etc., it is well-formed. Nevertheless, it too has to be submitted to binding principle C. Since we already know that the chain is well-formed, it obviously obeys the constraints of binding.

As presently formulated, binding theory "is concerned with the relations of anaphors and pronominals to their antecedents, if any" (Chomsky 1982:6). The beauty of binding theory is that it captures the relationship between lexical and empty noun phrases. If Chomsky wants to both retain Move- α and not have redundancies among the modules, then Binding theory could only apply to lexical categories and PRO, while Move- α and bounding theory apply to NP and *wh*-traces. The sharing of properties between lexical and non-lexical categories which binding theory captures would then be lost.

The new interpretive rule would replace bounding theory, Move- α , and inherent chains. In addition, binding theory would no longer be redundant. If Move- α is to be eliminated, then so is Bounding Theory. The sole motivation for the Bounding Theory module is to constrain the operation of Move- α at the level of S-structure. Thus, if Move- α does not operate between D- and S-structure, the Bounding theory module is also eliminated. Similarly, if Move- α is eliminated, then so are inherent chains, since inherent chains function solely to keep track of the successive applications of Move- α .

As Theories Ia and Ib are formulated, both retain free indexing and so have equal complexity in terms of the ill-formed data structures which must first be created and then rejected as being ill-formed. A parsing model, however, does not necessarily require free indexing.

It can assign indices as it builds the structure--based on the constraints of binding theory.

3.2 CHOMSKY'S LOGICAL NON-EQUIVALENCY ARGUMENT

Chomsky also argues that Move- α cannot be eliminated on the same grounds as are phrase structure rules:

The device of phrase structure rules is particularly suspect, because these rules so closely reflect lexical properties. Statement of the lexical properties is ineliminable from the grammar: For example, the grammar cannot avoid stating that *claim* takes a propositional complement as part of its lexical entry. Therefore, it is to be expected that the phrase structure rules should be eliminable insofar as they merely restate, in another form, the essential content of lexical entries. In fact, it seems that such rules are eliminable more generally, that there are no rules of this type in language. In the case of transformational rules, we have no comparable reason for skepticism concerning their existence, but it seems that the variety of these rules can be significantly reduced, perhaps to Move- α or Affect- α , with some parametric variation. (Chomsky 1986a:83)

His argument is that the redundancy between phrase structure rules and lexical structures is motivation enough for eliminating the former and there is no "comparable reason" for eliminating Move- α . However, if one assumes a lexically driven parser, then the redundancy between Move- α and the lexicon is logically equivalent to that between phrase structure rules and the lexicon. The function of Move- α is to build chains from argument positions to antecedents in non-argument positions. However, the lexical structure already specifies the arguments which are required. The projection principles insures that these arguments be syntactically represented. The placement of these arguments is independently motivated by X-bar theory, theta theory, and Case theory. If there is no overt NP in one of these argument positions, then the parser fills that position with an empty category. Move- α makes sure

that every argument position is filled, but the parser already knows to fill these positions through information that it finds in lexical entries.

3.3 'STUBBORN' VERSUS 'ILLEGAL'

Although Chomsky claims that any difference between the Move- α and Lexical models of EST is minimal, or perhaps that the two models are simply notational variants of the same model at a more abstract level, he consistently opts for the Move- α model. However, his argumentation for choosing the Move- α model over the Lexical model is marred by his failure to keep the two models distinct.

The most recent argument he presents in favor of the Move- α model, with its transformational component and level of D-structure, is that the respective contrast between (6) and (7) cannot be accounted for at S-structure, because both (6) and (7) have the S-structure (8).

- (6) John is too stubborn to talk to.
- (7) * John is illegal to talk to.
- (8) John_i is AP [e_i [PRO to talk to e_i]]

Therefore, the argument goes, the contrast must be result of the differences between the two sentences at the D-structure level (Chomsky 1986a:109, 113-114).

Chomsky motivates his argument for the level of D-structure as follows. Since (6) and (7) both have the same S-structure, he argues, the difference in grammaticality cannot be decided directly from the S-structure. If the two S-structures are the same, then the inherent chains associated with them must differ. These different inherent

chains can only be the result of the operation of Move- α . The grammaticality of (6) and (7) differ because they have different D-structures, as recorded in the S-structure chains.

If, in (6), *John* is base-generated as the object of a prepositional phrase complement, and (6) is derived by two successive applications of Move- α , the resulting S-structure (9) would violate the theta criterion.

(9) $John_i$ is too stubborn [e_j [PRO to talk to e_j]]

The theta criterion ensures that each chain have exactly one θ -role associated with it. The chain in (9) is associated with two θ -roles. If *John* moves up to the matrix clause, it moves out of one θ -position (the object of *talk*) and into another θ -position (the subject of *is too stubborn*). The fact that the subject of *is too stubborn* is a θ -position is shown by the ungrammaticality of (10), which has the expletive element *it* in the subject position:

(10) * It is too stubborn to talk to John

In Chomsky's analysis of (6), *John* is base generated as the subject of the main clause. The D-structure of (6) contains an empty operator (O) in the object position of the preposition *to*, as shown in (11). (Motivation for the empty operator analysis ([CP O_i [IP . . . e_j]]) will be discussed in "The Empty Operator and Strong Binding" on page 90.) This empty operator moves up to the clause initial operator position in the embedded clause by a single application of Move- α . The D-structure and S-structure of (6) are something like (11) and (12), respectively:

(6) John is too stubborn to talk to.

(11) John is too stubborn [PRO to talk to \emptyset]

(12) John is too stubborn [\emptyset_i [PRO to talk to e_i]]

The S-structure of (6) then contains two well-formed chains, (*John*) and (\emptyset , e), and *John* binds \emptyset . Thus Chomsky posits a base generated empty operator in the object position of the preposition *to* for (6).

In (7), according to Chomsky's analysis, *John* is in the object position of *talk* in D-structure, as illustrated in (13):

(13) *e* is illegal [PRO to talk to John]

By successive applications of Move- α , *John* moves up to the main clause.

First, move *John* to the front of the embedded clause ... leaving the trace e ; then move *John* from this "operator position" to the position of e in ... [(13)], leaving the trace e' . Now e is a variable bound by the operator e' . (Chomsky 1986a:113)

Therefore, the S-structure (14) has only one chain (*John*, e' , e).

(14) John is illegal [e' [PRO to talk to e]]

This chain does not violate the theta criterion, because the subject of *is illegal* is a non- θ -position, as shown by the well-formedness of (15) which allows the expletive *it* in the subject position.

(15) It is illegal to talk to John.

Instead, according to Chomsky, this chain violates binding condition C, which requires "that an r-expression must be A-free in the domain of the head of its chain" (Chomsky 1986a:114). The head of the chain of the variable e in (14) is *John*, which is in an argument position. Since the variable e is A-bound by *John*, the head of its chain, the sentence is ill-formed.

Chomsky claims that the grammar crucially requires a level of D-structure to get this distinction between the grammaticality of (6) and the ungrammaticality of (7).

This analysis of improper movement crucially requires that we regard Move- α as an operation converting D-structure to S-structure. A possible alternative interpretation that regards Move- α as in effect a property of S-structure, so that D-structures is (sic) "abstracted" from S-structure, would require a somewhat different analysis of the phenomenon. (Chomsky 1986a:114)

However, Chomsky's analysis that (6) and (7) have the same S-structure doesn't follow from either the Move- α model or a Lexical model--when they are kept distinct.

In Chomsky's analysis, he combines the Lexical and Move- α models. He generates the S-structure using the Move- α model, then he interprets this S-structure using a Lexical model. Essentially, Chomsky assumes a D-structure and Move- α to generate the S-structures. He then erases the inherent chains from the resulting S-structures and interprets these modified S-structures by building constructed chains with a Lexical model.

A Move- α model would not produce the S-structures for (6) and (7) that Chomsky claims it does. Contrary to Chomsky's assertions, (6) and (7) do not have the identical S-structure, (8), in a Move- α model. Given a Move- α model with movement, (7) has (8) as an S-structure, but (6) does not. In (6), *John* cannot be coindexed with an empty operator at S-structure, because *John* has not moved. *John* strongly binds the empty operator, and binding occurs in the LF. The S-structures differ in that their inherent chain structures differ. As Chomsky writes, (6) has two well-formed chains, and (7) has one ill-formed chain. In the Move- α

model, chains are part of the S-structure. The S-structures for (6) and (7) would be (16) and (8), respectively.

- (6) John is too stubborn to talk to.
 (16) John is AP [0_j [PRO to talk to e_j]]
 (7) * John is illegal to talk to.
 (8) John_{*j*} is AP [e_j [PRO to talk to e_j]]

If inherent chain structures are available, then we know that the S-structure (8) contains two NP traces (anaphors) with *John* at the head of the chain. As such, the anaphors are subject to, and violate, Principle A of Binding, which states that "an anaphor is bound in a local domain" (Chomsky 1986a:166). However, the argument Chomsky uses to explain the ill-formedness of (7) is that, following successive applications of Move- α , (8) contains an operator and a variable which are illegally bound by *John*. Within a Move- α analysis, the only way this could happen is that, when an NP lands on an "operator" position, and then moves again, its trace becomes an operator and the lowest trace becomes a variable. This is where the Move- α and Lexical analyses become entangled. If you have a Move- α analysis, then the S-structure (8) is created by movement. If movement has occurred, then inherent chains connect the traces to their antecedents. These chains are part of the S-structure. Since chains inhere in the S-structure, there is no reason to try to determine, as does Chomsky's analysis, the identity of an empty category on the basis of its location within a structural configuration. And since we know the identity of the empty categories, we know that they are not an operator and a variable, but two anaphors. Since they are two anaphors, the chain is ill-formed because it violates Principle A of Binding--the anaphor is not bound in a local domain. The head of

the chain, *John*, is not in a local domain. In addition, the lowest trace is A'-bound by the intermediate trace. In this movement analysis, both anaphors are in illegal positions: The lowest one is in a Case marked position, and the intermediate trace is in an A'-position. Given all these violations, (7) should sound much worse than it, in fact, does.

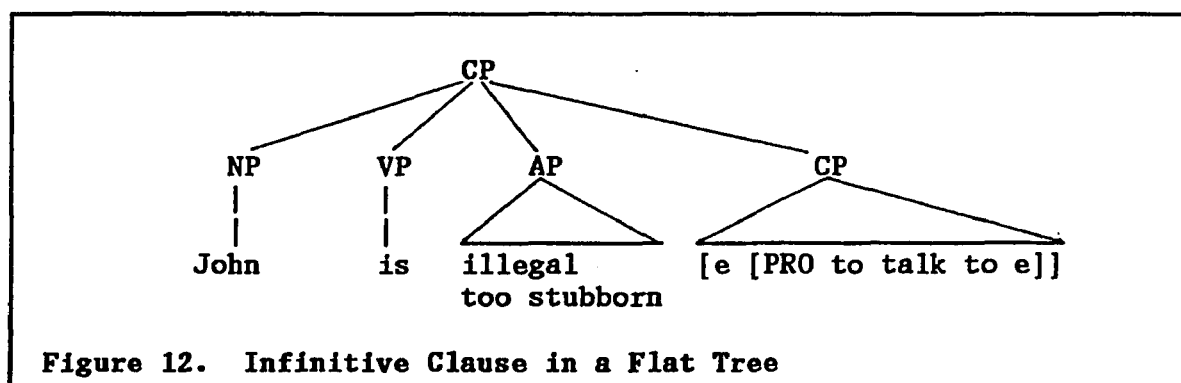
If you have a non-movement analysis within the principles-and-parameters framework, then the identities of the empty categories can be ascertained configurationally. With a non-movement analysis, the empty categories in the subordinate clause would be identified as an operator and a variable.

Let us further examine Chomsky's proposed S-structure for his crucial examples, repeated below:

(8) $John_i$ is AP [e_j [PRO to talk to e_j]]

Since only the infinitive clause is bracketed, it is unclear what common S-structure is being proposed.

If we assume that the bracketing is omitted because there is no structure, then Chomsky is proposing a flat S-structure as in Figure 12.



Since this configuration violates any X-bar theory one cares to appeal to, I will assume that some further bracketing is implicit in (8).

One conclusion about the bracketing in (8) is unavoidable; the infinitive clause is not in the AP. If the infinitive were in the AP, then possible factorizations for (6) and (7) would be (17), not (8).

- (17) i. John is AP
 ii. John is [AP ... [e [PRO to talk to e]]]

Another possibility is that the infinitive is a complement of the copula, as in Figure 13.

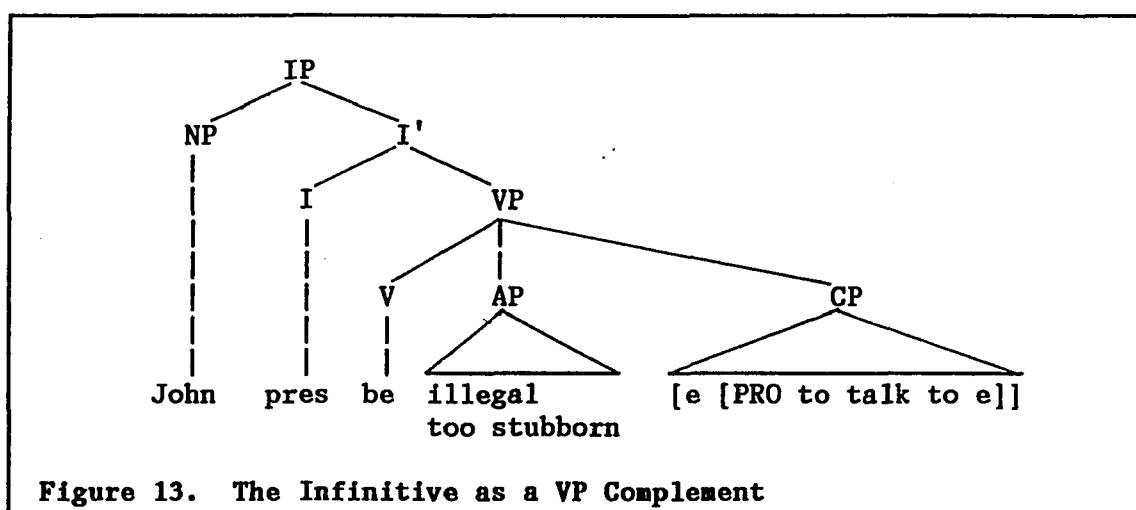
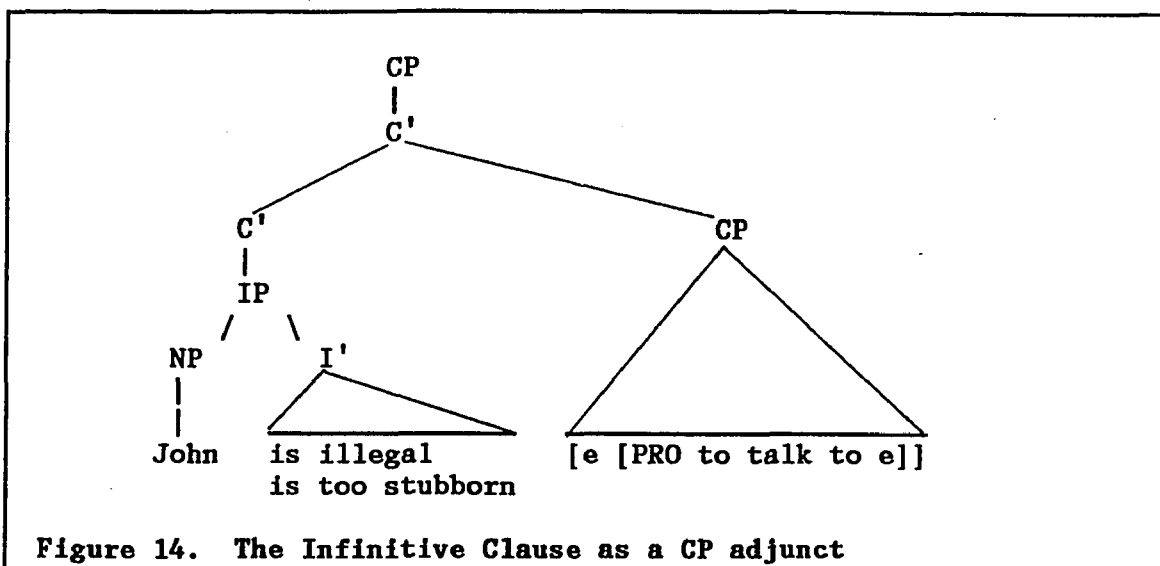


Figure 13. The Infinitive as a VP Complement

This would be hard to motivate as the distribution of the infinitive is restricted to co-occurring with only some adjectives.

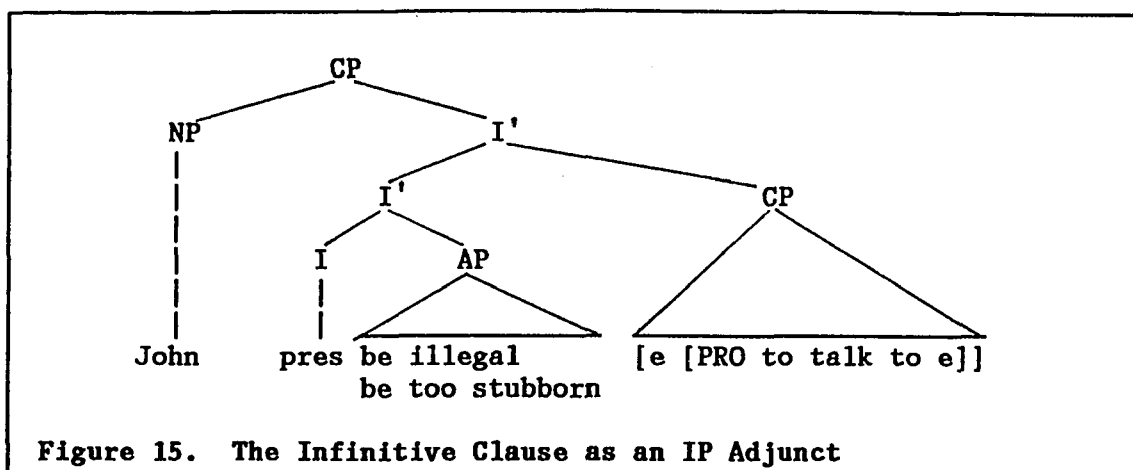
Three possible factorizations remain: (1) the infinitive is adjoined to the CP, (2) the infinitive is adjoined to the IP, and (3) the infinitive is adjoined to the VP.

The first option, that the infinitive is adjoined to the CP, is shown in Figure 14 on page 85.



However, if the infinitive is a CP adjunct, then the antecedent for the trace does not c-command the trace. (7) would be ruled out on syntactic grounds because an antecedent must c-command its trace. However, in order for (6) to be syntactically well-formed, a Strong Binder would not have to c-command the empty operator.

The second option is to assume that the infinitive clause is adjoined to the IP, and that the infinitive is a sentential modifier. This S-structure phrase marker would be as in Figure 15 on page 86.



Infinitival sentential modifiers are *in order to* phrases, as in (18):

(18) John left (in order) to see Mary.

The structure in Figure 15 is ruled out because it is impossible to get such a reading for either (6) or (7) along the lines of (19) as is shown in (20) and (21).

(19) John must be successful in order to impress his parents.

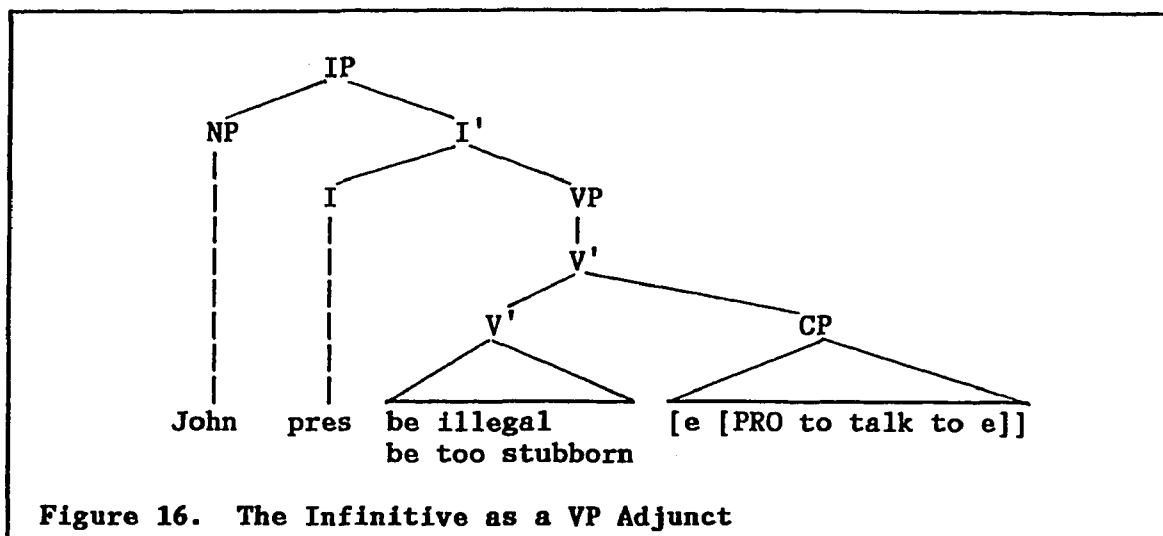
(20) * John is too stubborn in order to annoy his parents.

(21) * John is illegal in order to annoy his parents.

In addition, it is questionable whether the *in order to* phrase is optional in purposive clauses with adjectives, as is illustrated by (22).

(22) ? John is successful to impress his parents.

The third, and last, option is to assume that the infinitive clause is adjoined to the VP, as in Figure 16 on page 87.



This appears to be the only possible position for the infinitive. This structure is proposed in Chomsky & Lasnik (1977) to account for (23).

(23) It is unlikely that John is here.

However, Chomsky & Lasnik note that there are problems with such an analysis:

Then the embedded sentence is not in the domain of the adjective. However, it is not at all clear that this is the correct constituent analysis.

This problem holds for the infinitival clause in (8). In this configuration, the infinitive clause is not in the domain of the adjective.

At one point Chomsky refers to "the predicate 'too stubborn to talk to e_k '". This indicates that the infinitive is, in fact, in the AP. If the infinitive clause is in the AP, then the S-structures for (6) and (7) are not represented by (8). That is, (8) does not represent *too stubborn to talk to* as a predicate.

I will argue in "A Lexically Driven Analysis of 'Stubborn' and 'Illegal'" on page 89 that the differences in grammaticality between (6) and (7) reflects the lexical properties of adverbs and of predicative adjectives, not differences in movement versus non-movement derivations. The fact that (7) is apparently ill-formed while (6) is well-formed can best be incorporated into a grammar which licenses structure using lexical information.

4.0 A LEXICALLY DRIVEN ANALYSIS OF 'STUBBORN' AND 'ILLEGAL'

In this chapter, I propose a purely lexically driven analysis which captures the difference in acceptability between Chomsky's crucial examples concerning the predicative adjectives *stubborn* and *illegal*. In addition, I propose a lexically driven analysis for related constructions. I show that the differences between these predicates can be captured in a lexical analysis. On the other hand, it is not clear how a movement analysis would capture these differences.

Since the empty operator configuration ([CP 0_i [IP... e_i]]) figures prominently in our treatment of adjectives (as well as Chomsky's analysis of *is stubborn* discussed in "'Stubborn' versus 'Illegal'" on page 78), I will first present Chomsky's motivation for an empty operator and strong binding, and introduce the notion of empty operator effects. I then propose lexical structures for *illegal* and *stubborn*. A further property of Chomsky's crucial examples with *too stubborn* and *illegal* is that they exhibit peculiar control properties. The control exhibited by these sentences must also be explained.

Much of the same data that is analyzed here has been discussed by Lasnik and Fiengo (1974), hereafter L&F. My proposal resembles theirs in that we agree that these sentences do not involve movement. In place of a movement analysis, L&F propose an object deletion rule which deletes the objects of the complements in these structures. I propose an alternative analysis, with no deletion rule, which proposes different constituent and lexical structures from L&F in order to explain the phenomena exhibited by the data.

4.1 THE EMPTY OPERATOR AND STRONG BINDING

The empty operator is a non-overt *wh*- element which, like all operators, is in clause initial position at S-structure. There are two methods by which this empty category can be derived in a movement model: (1) a *wh*- element appears in D-structure and is deleted in the mapping between D-structure and S-structure, or (2) an empty operator is base generated in the D-structure.

Chomsky (1986a) motivates the empty operator based, in part, on the optionality of an overt *wh*- element in relative clauses. An overt *wh*- element in relative pronouns is optional at S-structure. That is, (1) and (2) are well-formed noun phrases.

(1) the man I saw

(2) the man who I saw

If one assumes that (1) and (2) have the same D-structure (3), then the optional *wh*- element is deleted from (1). That is, a *wh*- element appears in the D-structure object position, moves to the specifier of COMP, and is optionally deleted under the principle of recoverability of deletion, as illustrated by (4). If *who* is deleted, then (1) results, and if *who* is not deleted, then the result is (2).

(3) the man [I saw *who*]

(4) the man [*who*_{*i*} [I saw *e*_{*i*}]]

↓
ϕ

Alternatively, (1) can be generated with the empty operator as the object of *see*, as in (5), with *0* representing the empty operator. By this means, (1) has the D-structure (5), yielding the S-structure (6).

(5) the man [I saw o]

(6) the man [o_i [I saw e_i]]

The inherent chain in (6), [$o_i \dots e_i$] constitutes a well-formed chain. An r-expression is A'-bound by a variable in a local domain. However, a further principle is needed to explain why the empty operator, as well as overt relative pronouns, obligatorily corefer to the matrix noun phrase. Chomsky calls this principle strong binding.

Strong binding is introduced in this discussion on binding the empty operator in relative clauses such as <53>:

<53> the man [o [I [vp saw e]]]

We may assume that in this case the embedded relative clause is "predicated" of its head so that the index shared by o , e is identified with that of *man*. This property is quite general. A variable must not only be bound by an operator in the sense described earlier but must be bound in a still stronger sense: Either its range must be determined by its operator, or its value must be determined by an antecedent that binds it. Let us call this property *strong binding* as distinct from ordinary binding. Then, a further principle is:

<55> A variable must be strongly bound

(Chomsky 1986a:85)

Chomsky then modifies Binding Principle C, that an "r-expression must be free," yielding:

An r-expression must be A-free (in the domain of its operator)
(Chomsky 1986a:86)

The claim is that a variable requires either a "range" which it acquires from being bound by an overt *wh*- element or a "value" which it acquires from a lexical antecedent. However, points are left unclear. Just how strong binding differs from ordinary binding is not defined. Nor is

the structural configuration of a variable and its possible antecedent defined.

The result of strong binding is that r-expressions can be coindexed with noun phrases in A positions without constituting a Principle C violation. Chomsky later reformulates his modification of Binding Principle C as follows:

<87> An r-expression is A-free (in the domain of the head of its maximal chain)

(Chomsky 1986a:98)

Chomsky redefines the nature of the domain from "domain of its operator" to "domain of the head of its maximal chain" which he subsequently refers to as "domain of the head of its chain."

This reformulation crucially distinguishes between the movement and non-movement models. The domain of the head of a chain is determined from the result of Move- α operating between D-structure and S-structure. The domain of an operator can be structurally determined on an S-structure without recourse to a residue left on the structure by Move- α .

Chomsky refines his definition of Principle C specifically to capture the difference between the grammaticality of (7) and (8).

(7) John is too stubborn to talk to.

(8) John is illegal to talk to.

He claims that (7) is well-formed because in the structure submitted to the binding module, (9), the empty operator is the head of the inherent chain. Since the empty operator position is an A'-position, the variable is A-free in the head of its chain.

(9) John_i is too stubborn [*O*_i [PRO_j to talk to *e*_j]]

In the same vein, (8) is ruled out because in the structure submitted to the binding module, (10), *John* is the head of the inherent chain. Since *John* is in an A-position, the variable is not A-free in the domain of the head of its chain.

(10) John_i is illegal [*e*_i [PRO to talk to *e*_i]]

As I have pointed out in "'Stubborn' versus 'Illegal'" on page 78, it is not clear what constituent structure Chomsky intends to represent in (10). If Chomsky means what he says upon referring to "the predicate 'too stubborn to talk to *e*_k,'" (Chomsky 1986a:109) and that:

both <114> and <115i> have the S-structure form:

<116> John_i is AP [*e*_i [PRO to talk to *e*_j]]
 [<114> John_i is too stubborn [*O*_k [PRO to talk to *e*_k]]]
 [<115> i. * John is illegal [*e*' [PRO to talk to *e*]]]
 (Chomsky 1986a:113)

then *illegal to talk to* is also a predicate. If (10) is marked ill-formed by this version of Binding Principle C, then all *tough*-movement predicates will similarly violate Binding Principle C. Hence, due to the existence of *tough*-movement constructions, Chomsky's result is not descriptively accurate. I discuss this point in more detail in "The Constituent Structure of 'John is easy to talk to'" on page 117.

4.1.1 THE EMPTY OPERATOR IN INFINITIVE CLAUSES

When noun phrases contain infinitival relative clauses, overt *wh*- NPs cannot appear.⁷ Compare (11) and (12) with (13).

⁷ *Wh*- NPs seem to be blocked because they cannot be assigned Case. A *wh*- NP can appear in a prepositional phrase in this position, as in *Mary is the woman to whom to speak*. When the preposition is stranded, the *wh*- element does not get case, and the sentence is

(11) [the man (who) I saw] is in the office

(12) [the man to see] is in the office

(13) * [the man who to see] is in the office

In (12) we find the same effects of operator binding as in Chomsky's example <53>. An empty category in object position is bound by the head of the relative clause. The S-structure of (12) is (14).

<53> the man [O [I [vp saw e]]]

(14) [the man [O_f [PRO to see e_f]]] is in the office

As with (1), there is an empty category in the object position of the relative clause. This empty category is bound by an operator in a complementizer position. This operator is, in turn, strongly bound by *the man*. In addition, since the embedded clause is infinitival, a PRO is in subject position. This PRO is not coindexed with an antecedent..

Let us call this set of effects empty operator effects. Empty operator effects in infinitive clauses consist of: (1) a gap in object position which is (2) strongly bound to an antecedent, and (3) the subject PRO has no antecedent.

There are four infinitive constructions that exhibit empty operator effects:

1. Infinitival relative clauses. These are discussed in "The Control Mechanisms for Adjuncts" on page 278.
2. Infinitival complements of comparative adverbials. These are discussed in "The Constituent Structure of 'John is too stubborn to talk to'" on page 103.

ill-formed, as in * *Mary is the woman whom to speak to* and opposed to *Mary is the woman to speak to*.

3. Infinitival complements of *tough*-movement predicates. These are discussed in "The Constituent Structure of 'John is easy to talk to'" on page 117.
4. Infinitival adjuncts of *pretty*-predicates. This class is defined and discussed in "The Constituent Structure of 'Mary is pretty to look at'" on page 125.

The control of PRO in these four environments is discussed in "Distribution and Interpretation of the Empty Operator" on page 325.

4.1.2 THE EMPTY OPERATOR IN PARASITIC GAPS

Chomsky (1986a:111) also proposes empty operator movement and strong binding for parasitic gap constructions, such as (15), where *which book* appears to bind both e_i and e_j , in violation of the Theta Criterion.

(15) *which book* did you file e_i [without reading e_j]

Chomsky proposes the S-structure (16), where *which book* is bound to e_i by movement and an empty operator is bound to e_j by movement.

(16) *which book_i* did you file e_i [without [0_j [reading e_j]]]

Strong binding subsequently links the operator to the variable *which book*.

4.2 THE LEXICAL STRUCTURES OF 'STUBBORN' AND 'ILLEGAL'

Conceivably a lexically driven parser could be defined to build the S-structure (17) for both (7) and (8), but there are no reasons to do so. On the contrary, below I discuss reasons for not doing so.

(7) John is too stubborn to talk to.

(8) John is illegal to talk to.

(17) John_{*i*} is AP [e_i [PRO to talk to e_i]]

Since the parser has access to lexical structures, it licenses S-structures based on such lexical properties as external θ -role assignment and internal argument structure. I will show that the two adjectives exhibit very different lexical properties and so require different lexical structures. There are no data to indicate that the lexical structures for the predicative adjectives *illegal* and *stubborn* are identical.

4.2.1 EXTERNAL θ -ROLE ASSIGNMENT

As we have already noted, the two predicative adjectives, *stubborn* and *illegal* differ with respect to external θ -role assignment. The lexical structure for *stubborn* has to indicate external θ -role assignment. There are two lexical structures for *illegal*. *Illegal*₁ does not assign an external θ -role, and *illegal*₂ does assign an external θ -role.

The ill-formedness of (18) indicates that the subject position of *stubborn* is a θ -position.

(18) * It is too stubborn to talk to John.

The subject of *illegal*, when it subcategorizes for an infinitive complement, is a non- θ -position. This is illustrated by (19).

(19) It is illegal to talk to John.

The second subcategorization for *illegal* does assign an external θ -role, so an expletive cannot appear in the subject position. The lexical entry for *illegal* which licenses an external θ -role licenses no complements. All of the copular constructions in (20) are well-formed, where the subjects are θ -positions and no complements are licensed by the lexical entry.

- (20) i. Unlicensed pistols are illegal.
 ii. Pornography should be illegal.
 iii. Prostitution is illegal.
 iv. Cocaine is illegal.

If the subject positions were not θ -marked, then the examples in (21), when the interpretation of *it* is non-referential, should be as well-formed as (22) is.

(21) * It is illegal.

(22) It is raining.

4.2.2 COMPLEMENT STRUCTURE

The two adjectives also differ with respect to their internal complement structures.

*Illegal*₁ licenses a clausal complement, as illustrated in (23).

(23) It is illegal (for me) to talk to John.

*Illegal*₂ does not license a clausal complement, as is illustrated by the ill-formedness of the examples in (24).

- (24) i. * John is illegal to talk to Bill.
 ii. * Cocaine is illegal to take drugs.
 iii. * Pornography is illegal to daydream.

Stubborn does not license a clausal complement, as illustrated in (25).

(25) * John is stubborn to die.

Thus the clausal complement in (7) is not licensed by *stubborn*, but by something else. I will discuss what licenses the clausal complement in (7) directly.

4.2.3 COMPARATIVE FORMATION

The two adjectives differ with respect to the formation of comparatives and superlatives. *Illegal*₁ does not form a comparative or superlative except in a restricted class of cases where the context is crucial. This is illustrated by the ungrammaticality of the examples in (26).

- (26) i. * It is more illegal to talk to John than to Bill.
 ii. * It is most illegal to talk to John.

In certain cases, *Illegal*₁ does allow formation of comparatives -- as illustrated by the relative grammaticality of the examples in (27).

- (27) i. ? It is more illegal to sell cocaine than (to buy) marijuana.
 ii. ? It is most illegal to sell cocaine.

*Illegal*₂ forms comparatives more easily, as illustrated by (28).

- (28) Cocaine is more illegal than marijuana is.

Stubborn does take comparatives and superlatives structures, as is illustrated in (29).

- (29) i. John is more stubborn than Bill (is).
 ii. John is the most stubborn person I know.

Formation of comparatives and superlatives either has to be marked in the lexical structure of *stubborn* and *illegal*₂, or non-formation of comparatives and superlatives has to be marked in the lexical structure of *illegal*₁. I will assume the latter, although none of my argumentation bears on this choice.

The comparative construction relevant in example (7) is that formed with *too*. This comparative form is described in Quirk et. al.:

One type of comparative construction contains a word or phrase expressing the notion of sufficiency or excess followed by a *to*-infinitive clause of purpose, result, or condition. The most common word for sufficiency is *enough* (as adverb, determiner, or pronoun) and the most common word for excess is the adverb *too*.

(Quirk, et. al. 1987:1140)

It is the comparative formative *too* which licenses the clausal complement in (7). *Stubborn* takes an infinitival complement when it is in a comparative construction with *too*, as is illustrated by (7), but does not itself license a clausal complement, as illustrated in (30).

(7) John is too stubborn to talk to.

(30) * John is stubborn to talk to.

Since *illegal*₁ does not form comparative constructions, the examples in (31) are ill-formed.

- (31) i. * John is too illegal to talk to.
 ii. * Cars are too illegal to park here.
 iii. * It is too illegal to talk to John.
 iv. ? Cocaine is too illegal to sell.

Lasnik and Fiengo (1974) and Chomsky (1977b) have noted that *too* and *enough* take clausal complements. Lasnik and Fiengo assume that the adverbial as well as its complement are in the specifier position of the AP and surface order is achieved by a rule of extraposition. Chomsky (1977b) also assumes that *too* and *enough* are specifiers which take complements:

Let us now turn to infinitival complements within the category of adjective phrases. Consider first structures of the form <107>, where I assume that S' is a complement of the adjective qualifier *enough*

<107> John is tall [enough [S' for us to see him]]

...

Other complements of adjective qualifiers, as in <120>, have about the same properties as the complements of *enough*, so far as I can see, so I will have nothing to say about these:

<120> Muhammad Ali is too good [(who) for Bill to arrange for John to fight t]

(Chomsky 1977b:101-102).

Since *too* licenses a complement clause, its lexical structure must include this information. Within any framework, the lexical structure should include the information that *too* licenses an adjective phrase and an optional infinitive as complements. The lexical structure for *too* contains the information given in the partial lexical structure (32).

(32) *too* $\left[\begin{array}{l} \text{adverb} \\ \text{— AP CP} \end{array} \right]$

4.2.4 LEXICAL STRUCTURES FOR 'STUBBORN' AND 'ILLEGAL'

The predicative form of *illegal* has two lexical structures which would include the information given in the partial structures (33i) and (33ii). (33i) requires a clausal complement and does not form comparatives and superlatives; (33ii) assigns an external θ -role.

(33) i. *illegal*₁

$$\left[\begin{array}{l} \text{adjective} \\ \text{— CP} \\ \neg \text{comparative} \end{array} \right]$$

ii. *illegal*₂

$$\left[\begin{array}{l} \text{adjective} \\ \text{external } \theta\text{-role} \end{array} \right]$$

The lexical structure for the predicative form of *stubborn* always assigns an external theta-role, as is illustrated by the well-formedness

of the examples in (34) and the ill-formedness of (35), when *it* has a non-referential interpretation.

- (34) i. John is stubborn.
 ii. The lock is stubborn.

(35) * It is stubborn.

In addition, as I have pointed out, *stubborn* does not license a clausal complement. The lexical structure for *stubborn* would then include at least the information in the partial lexical structure (36).

(36) stubborn $\left[\begin{array}{l} \text{adjective} \\ \text{external } \theta\text{-role} \end{array} \right]$

Note that *illegal*₂ and *stubborn* have similar lexical structures. Both assign external θ -roles, as shown in (20iv) and (34i), and neither licenses a clausal complement, as shown in (24ii) and (25).

- (20) iv. Cocaine is illegal.
 (34) i. John is stubborn.
 (24) ii. * Cocaine is illegal to take drugs.
 (25) i. * John is stubborn to die.

4.3 CONTROL VIOLATIONS IN THE CRUCIAL EXAMPLES

The difference in lexical structures between the two adjectives accounts for the well-formedness or ill-formedness of many of the preceding examples, but they do not explain all of the properties exhibited by (7) and (8).

- (7) John is too stubborn to talk to.
 (8) John is illegal to talk to.

An additional property of (7) and (8) is that they both exhibit Control violations. It is still an open question whether control is lexical, configurational (defined in terms of UG primitives), or a mixture of the two. If the first option proves true, and control is lexical, then control has to be explicitly specified in the lexicon. If the second option is true and control is structurally defined, then the antecedent of PRO is determined by structural considerations only. A third possibility is that whether PRO may or must be bound is structurally defined by control theory, but that subject or object control is specified in the lexicon. Chomsky appears to assume some version of this third possibility. According to Chomsky, Control theory is a principle of Universal Grammar:

The question of when PRO may or must be bound or free falls under control theory, another module of UG.
(Chomsky 1986a:124)

However, in his analysis, Chomsky assumes that control is stated in the lexicon: "we know that *persuade* requires that its object control PRO" (Chomsky 1986a:104).

Theories of Control (Chomsky 1980a, Williams 1980, Manzini 1983, Hornstein and Lightfoot 1987) generally assume that the element PRO, the null subject of infinitives and gerunds, looks for the nearest commanding NP, if there is such an NP, as a potential controller. Otherwise reference is arbitrary. If the features of the nearest potential controller are incompatible with the features associated with PRO, that is, the potential controller is not a possible controller, then the sentence is usually ill-formed, as in (37) and (38).

(37) * John told Mary [PRO to hit himself]

(38) * It rained before [PRO going to the movies]

It is this Theta criterion violation which motivates Chomsky's empty operator analysis of (7), yielding the partial S-structure (40).

(40) John is too stubborn [θ_i [PRO to talk to e_i]]

I will adopt Chomsky's empty operator analysis for (7).

However, as I have argued in "Comparative Formation" on page 98, *too* licenses the empty operator construction. Although Chomsky (1977b) adopts this position, it appears to have been abandoned in Chomsky (1986a). In Chomsky's (1986a) discussion of (7), he does not specify how the infinitive clause is licensed. He simply refers to "the predicate 'too stubborn to talk to e_k '" (Chomsky 1986a:109). Although Chomsky does not explicitly say how the infinitive complement in (7) is licensed, it follows from his analysis of (7) and (8) that *too* cannot be the licenser, because he claims that (7) and (8) have identical S-structures. If an adverbial is a licenser in (7), and (8) contains no adverbial, then the sentences would not have identical S-structures.

As was discussed in "External θ -role Assignment" on page 96, *stubborn* assigns an external θ -role, as shown in (34i), but does not license a clausal complement, as shown in (41).

(34) i. John is stubborn.

(41) i. * John is stubborn (for me) to talk to Mary.
 ii. * John is stubborn (for me) to talk to.

The apparent violations in (7), which are typical properties exhibited by the presence of an empty operator, are not the result of external θ -role assignment in conjunction with a clausal complement,

as I will illustrate in "The Constituent Structure of 'John is afraid to talk to Mary'" on page 111.

The problematic structure is a consequence of the lexical entry for the grammatical formative *too*. *Too* does license a clausal complement, as is illustrated by (42).

(42) John is too stubborn (for me) to talk to.

In addition, *too* licenses the empty operator effects. The empty operator binds an empty category in the object position of a transitive verb or, in this case, the object of a preposition. The empty operator is, in turn, strongly bound by *John*.

If *too* licenses an empty operator, then *too*-sentences with optionally transitive verbs in the embedded clause should be ambiguous. This ambiguity would be whether the intransitive form is used and the matrix subject binds an EC in the subject position of the embedded infinitive or if the transitive form is used and an empty operator in the specifier of the CP binds the object position of the embedded infinitive. The class of sentences which contains this ambiguity was discussed in Lees (1960). (43i) has two interpretations, (43ii) and (43iii). The interpretation of (43ii) is that a (live) lamb is too fat to eat any food. Here *eat* is intransitive, and *the lamb* binds an EC in the subject position of *eat*. The interpretation of (43iii) is that some cooked lamb is so fat that someone cannot eat it. Here *eat* is transitive, an empty operator binds the object position of *eat*, and the PRO receives an arbitrary interpretation. *The lamb*, in turn, strongly binds the empty operator.

- (43) i. The lamb is too fat to eat.
 ii. the lamb_{*i*} is too fat [*e_i* to eat]

iii. the λ_f is too fat [0_f [PRO_{arb} to eat e_f]]

The parser inserts the empty operator only when a gap in the complement position licenses it. No Control violation is exhibited by the examples in (44), where the object position is filled by a lexical noun phrase. Their interpretations are that of (45), with PRO being controlled by the matrix subject.

- (44) i. The lamb is too fat to eat carrots.
 ii. John is too stubborn to talk to Mary.

- (45) i. the λ_f is too fat [PRO_f to eat carrots]
 ii. John_f is too stubborn [PRO_f to talk to Mary]

Similarly, when an intransitive verb is used, as in (46), the subject of the infinitive complement of *too* is controlled by the matrix subject.

(46) John_f is too stubborn [PRO_f to die]

In (7), *too* licenses the CP which can contain the empty operator. The empty operator, in turn, binds the EC in the object position of *to talk*. We now have the structure proposed by Chomsky, shown in (40).

(40) John is too stubborn [0_f [PRO to talk to e_f]]

In order to explain why *John* and the variable corefer, Chomsky appeals to the principle of strong binding. With strong binding, r-expressions can be coindexed with NPs in A-positions without violating Binding Principle C. Since (7), like relative clauses, is analyzed as having an empty operator, Chomsky argues that the same principle is operating:

<104> John_f is too stubborn [0_k [PRO_j to talk to e_k]]

But to be licensed, the variable e_k must be strongly bound in the sense defined earlier (see <55>). Since its operator, being empty, does not specify a range, the variable must be associated with an antecedent in a structurally appropriate position that assigns it a value. Only *John* is available as an antecedent,

and it is in an appropriate position, as the subject of the predicate "too stubborn to talk to e_k " for general reasons. Therefore, e_k takes John as value. It follows, then, that $i=k$.

Recall that the binding principle <87> requires that a variable be A-free in the domain of the head of its chain. This condition is satisfied by <104>; the variable is A-bound by *John*, but the latter is not in the domain of its operator O , which is the head of the chain formed by movement of O from its D-structure position to the position it occupies in <104>. (Chomsky 1986a:109)

Since Binding (and presumably strong binding) occur in the LF, and movement or bounding is recorded in the S-structure, Chomsky's analysis predicts that the S-structure of (7) would be (47) and its LF-structure would be (48) (assuming that strong binding reassigns indices).

(47) $John_i$ is too stubborn [O_k [PRO to talk to e_k]]

(48) $John_i$ is too stubborn [O_i [PRO _{j} to talk to e_i]]

As Chomsky points out, his analysis of (7) and (8) with its use of Principle C and Strong Binding depends crucially on a model with movement. In addition, the Binding component in this Move- α model is required to be sensitive to the difference between an inherent chain and data structures which are the output of free-indexing.

A lexical model analysis for (7) will have to be somewhat different. In a lexical model, binding operates on S-structures, imposing its well-formedness conditions on both constructed chains and Binding structures. In addition, the Theta criterion is a well-formedness condition on the constructed chains which are built by the parser.

The lexical model, on the other hand, has the advantage of having access to the lexical structures. Licensing of phrase structure is achieved through three channels: (1) An element can be licensed by the lexicon based on the content of an entry. For example, the lexicon

licenses complements and external arguments. (2) General principles of UG provide licensing principles. For example, by the "extended projection principle" each IP needs a subject; VP is a complement of IP; each sentence needs a CP and an IP. (3) Language particular principles of UG (parameter settings) are licensing principles. For example, English is SVO by the head-first parameter setting of X-bar theory; English is a non-pro-drop language; etc. The parser constructs or builds the tree for a string based on licensing principles. Neither the lexicon nor the principles do the building, they simply regulate the operation of the parser.

A lexical parser would build an S-structure for (7) along the following lines:

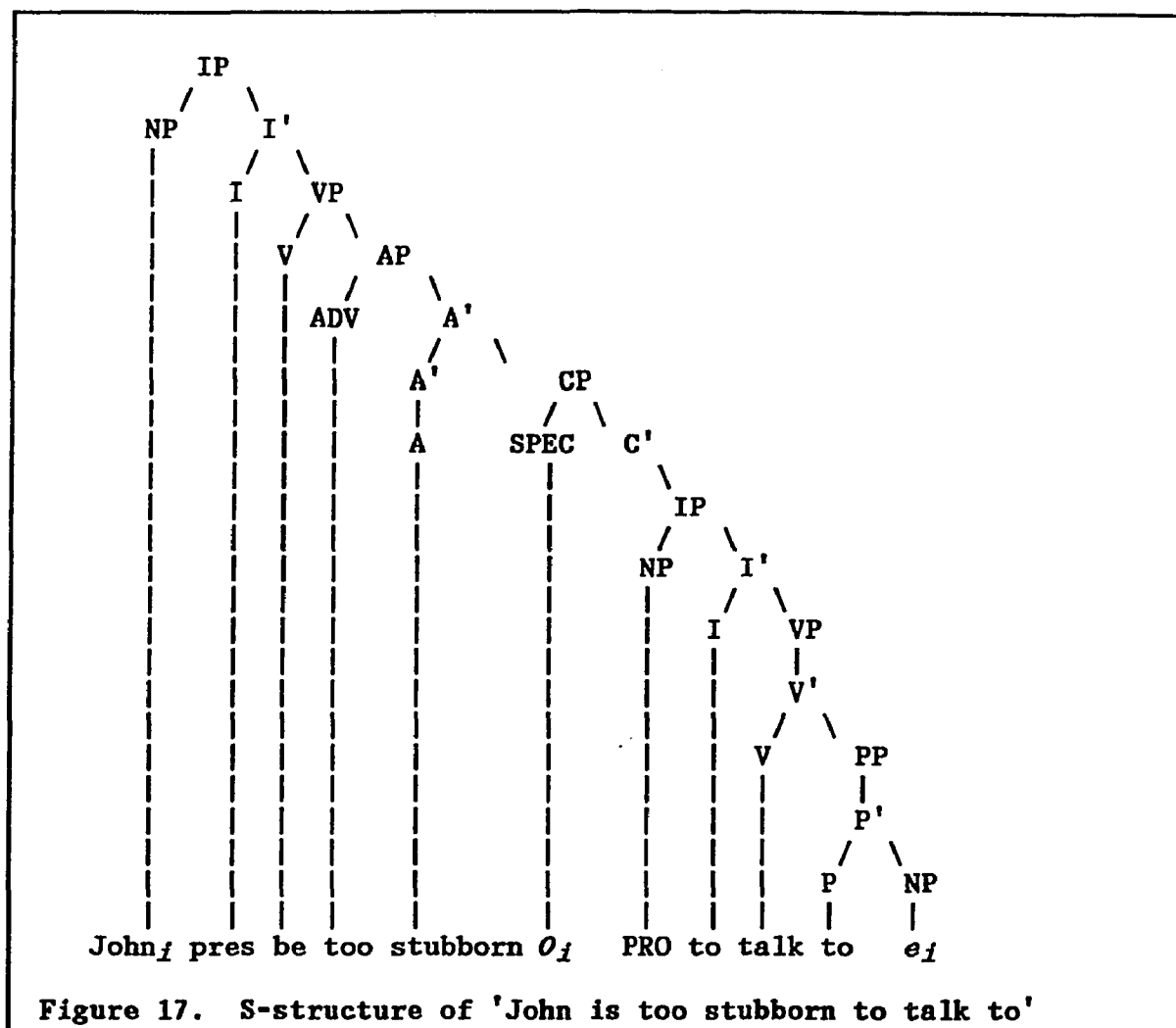
(7) John is too stubborn to talk to.

1. Verbs head VPs. In addition, the lexical item *talk* licenses a PP complement. Hence, based on X-bar theory and the head-first parameter for English, the parser builds a VP, headed by *talk*, with a PP complement. Since the PP has no NP object, the parser inserts an empty category in object position:

[*VP* talk to *e*]

The empty category is determined to be a variable because it is in a case marked position.

2. By UG principles, IPs take VP complements. The parser starts building an IP over the VP. The head of INFL is determined to be *to*. By the extended projection principle, IPs require subjects. In building an IP for the complement clause, the parser looks for



a subject for *talk*. Not finding an overt subject, the parser inserts an empty category in the IP specifier position.

[IP e to [vp talk to e]]

3. The lexical structure for *stubborn* licenses no complements, but it does assign an external θ -role. The parser builds an AP over *stubborn*.

[AP stubborn]

4. The lexical entry for *too* licenses an AP and a CP complement. The parser builds a CP over the IP. Since the subject of the IP is now ungoverned, and the head of INFL is *to*, the subject empty category is determined to be PRO. Since variables must be bound, the parser inserts an empty category in the complementizer position to bind the variable.

[CP e_i [IP PRO *to* [VP *talk to e_i*]]]

The parser builds an adverbial phrase headed by *too* and attaches the AP and CP as its complements.

5. Build VP.
6. Build IP with *tense* as its head and *John* in specifier position. The lexical entry for *stubborn* licenses an external θ -role. Therefore *John*, the subject of *stubborn*, is in a θ -position.
7. Since the parser finds no overt operator to bind the empty category in the operator position of the complement clause, it determines that the empty category in the operator position is an empty operator.

The parser has now licensed the structure in (49).

(49) John is [*too* [*stubborn*] [θ_i [PRO *to talk to e_i*]]]

Now, however strong binding is formulated, it applies, binding the empty operator to *John*, yielding the S-structure in Figure 17 on page 109.

Chomsky claims that PRO in (7) receives an arbitrary interpretation because, if PRO were bound by *John*, the sentence would violate crossover constraints.

Therefore, PRO cannot be bound by *John*. But there is no other binder. Therefore, PRO must be free, hence, arbitrary in interpretation...

(Chomsky 1986a:109).

This analysis of why PRO is allowed to be arbitrary in interpretation is problematic. Usually, as was pointed out in "Control Violations in the Crucial Examples" on page 101, if a potential controller is not a possible controller, then the sentence is ill-formed, as in (37) and (38).

(37) * John told Mary [PRO to hit himself]

(38) * It rained before [PRO going to the movies]

Sentences containing empty operators require an arbitrary interpretation for PRO when a potential controller is not a possible controller. A consequence of my analysis is that the empty operator configuration licenses an arbitrary interpretation for PRO, not the fact that *John* cannot bind PRO without violating crossover constraints.

4.5 THE CONSTITUENT STRUCTURE OF 'JOHN IS AFRAID TO TALK TO MARY'

Predicate adjectives which both assign an external θ -role and license a clausal complement do not exhibit the peculiar properties of *is too stubborn*. These adjectives, which I will call *afraid* predicates, have the partial lexical structure in (50).

(50) afraid $\left[\begin{array}{l} \text{adjective} \\ \text{external } \theta\text{-role} \\ \text{— CP} \end{array} \right]$

Afraid predicates do not exhibit Control violations, as does *is too stubborn*. In the *afraid*-class examples in (51), there is no arbitrary interpretation for the subject of *to talk*; PRO is controlled by *John*.

- (51) i. John_f is afraid [PRO_f to talk to Mary]
 ii. John_f is accustomed [PRO_f to talk to Mary]
 iii. John_f is eager [PRO_f to talk to Mary]

Similarly, *afraid*-class predicative adjectives do not exhibit double θ -role assignment. The examples in (52) are ill-formed because there are two θ -roles associated with *John*, which violates Theta criterion.

- (52) i. * John is afraid to talk to.
 ii. * John is accustomed to talk to.
 iii. * John is eager to talk to.

If (52) were assigned an appropriate structure for strong binding, then the sentences would be well-formed. However, since strong binding does not occur, we can assume that an empty operator is not licensed with this class of adjectives.

Adjectives which assign external θ -roles and license clausal complements do not exhibit any of the features which are associated with empty operators. Empty operators are not licensed in complements of adjectives which assign external θ -roles.

4.6 LEXICAL STRUCTURES IN 'JOHN IS ILLEGAL TO TALK TO'

Chomsky's crucial example (8) poses no problems for the theta criterion. Since the predicative adjective *illegal* assigns no external θ -role, *John* and the empty category object of the preposition *to* are associated with a single θ -role, as illustrated in (53).

- (8) John is illegal to talk to.

(53) John is illegal [PRO to talk to e]
 \uparrow --- \downarrow
 - θ θ

Chomsky's claim that (8) is ill-formed due to a violation of Binding Principle C doesn't hold for a Move- α model (see "'Stubborn' versus 'Illegal'" on page 78), however a Principle C violation could occur in a Lexical Model. But, in order for the Principle C violation analysis to hold for a lexical model, the model would require (a) Chomsky's version of Binding Principle C in which a variable is A-free in the domain of the head of its chain and (b) that ill-formed chain structures are built and then filtered out. In this parsing model all chains must be constructed, no inherent chains are input to the parser. Let us assume that the parser, like Chomsky's, determines the identity of empty categories configurationally. The parser would determine that the lower empty category is a variable because it is in a case-marked argument position. The parser would then insert an empty category in the operator position (specifier of the CP) of the lower clause to bind the variable. Up to this point, the parser has built a well-formed chain. However, in order for the operator to be bound by *John*, the parser would have to build an ill-formed chain. Let us assume that it freely coindexes the empty categories it has inserted with overt NPs, and that *John* and the operator position are thus joined in a chain. The output of this parser can be a tree which has a variable bound to an A-position, yielding the structure (54).

(54) $John_f$ is illegal [e'_f [PRO to talk to e_f]]

The resulting constructed chain (*John*, e' , e) violates the version of Binding Principle C which states that an r-expression must be A-free

in the domain of the head of its chain. *John*, the head of the chain, is in an argument position, and so cannot head a chain which contains variable. With the version of Principle C that states that an r-expression is A-free in the domain of its operator, the chain in (54) would be well-formed because the variable is bound from an A' operator position in its local domain.

However, this Principle C violation analysis of Chomsky's would mark the *easy* (or *tough*) constructions as ill-formed. If the parser builds a structure which violates Binding Principle C for (8), it will presumably build a similar ill-formed structure for *tough* constructions, as I discuss in "The Constituent Structure of 'John is easy to talk to'" on page 117.

I propose that (8) does not constitute a Principle C violation. What, then, is the status of (8)? In "The Lexical Structures of 'stubborn' and 'illegal'" on page 95, I proposed two lexical structures for the predicative adjective *illegal*, repeated in (33) below:

(33) i. *illegal*₁

adjective — CP ¬comparative

ii. *illegal*₂

adjective external θ-role

Two questions arise: If (8) is ill-formed, why doesn't either lexical structure drive the parser to assign a structure to it? If (8) is well-formed, which structure in (33) does it pattern after?

(8) John is illegal to talk to.

To help us answer these questions, let us examine the range of intuitions people get for (8) and related sentences.

Intuitions on sentences with the predicate adjective construction *is illegal* differ. There is not even agreement among native English speakers that (8) is, in fact, ill-formed. An informal survey shows that most informants agree that (8) is ill-formed. However, all those who dislike (8) find the examples in (55) acceptable to varying degrees. Although they find the examples in (55) well-formed, they agree that the sentences with expletive subjects in (56) "sound better." To further complicate matters, a couple of informants find (8) and (55) equally well-formed.

- (55) i. Unlicensed pistols are illegal to own.
 ii. Pornography should be illegal to sell.
 iii. Unregistered aliens are illegal to hire.
 iv. Abortions will be illegal to perform.

It is unanimously agreed that all the sentences in (56), which have expletives in subject position, are fine.

- (56) i. It is illegal to own unlicensed pistols.
 ii. It should be illegal to sell pornography.
 iii. It is illegal to hire unregistered aliens.
 iv. It will be illegal to perform abortions.

The interesting thing about the examples in (55) is that although (8) and (55) have the same structure, many people who like the examples in (55) do not accept (8). If the examples in (55) are syntactically well-formed, then so should (8) be. Why, then, do so many people find

(8) bad? On the other hand, if (8) is ill-formed, then the examples in (55) should also be bad.

Inspection of the sentences in (55) reveals that the matrix subject NPs are those which are pragmatically acceptable for the lexical structure for *illegal*₂ (which assigns an external θ -role but assigns no internal arguments). All of the matrix NPs denote things or actions which either happen to be illegal or whose legal status is currently being debated. In addition, the infinitive clause is not acting like a complement in (55). All of the sentences in (55) are well formed sentences without the infinitive clause, as illustrated by (57). The infinitive clauses in (55) are optional. That is, the infinitives are adjunct clauses.

- (57) i. Unlicensed pistols are illegal.
 ii. Pornography should be illegal.
 iii. Unregistered aliens are illegal.
 iv. Abortions will be illegal.

There are two classes of speakers: A-type speakers like (55) but not (8). They require an expletive element in the subject position of *illegal*₁ (which assigns no external θ -role but does license a complement clause) except for a specific class of NPs. B-type speakers like both (8) and (55). They allow all NPs -- expletive and non-expletive -- in the subject position of *illegal*₁. A-type and B-type speakers will differ in their judgements on (58). A-type speakers find the examples in (58) ill-formed, B-type speakers find the examples in (58) well-formed.

- (58) i. Cars are illegal to park here.
 ii. Dogs are illegal to walk in the park.
 iii. Women will be illegal to perform abortions on.
 iv. Minors are illegal to sell alcohol to.
 v. Flags are illegal to burn.

This range of intuitions is explained when we assume that the two types of speakers are using different lexical structures to drive the parser to assign a structure to a string containing *illegal* with an infinitive clause.

A-type speakers assume the lexical structure of *illegal*₂ and interpret the infinitive as an adjunct. I will discuss this construction in "The Constituent Structure of 'Mary is pretty to look at'" on page 125. Since they are assigning a θ -role to the subject position, then the subjects are interpreted as being predicated of *illegal* and the sentences in (58) are marked as ill-formed. Note the difference in the logical relationship between (55) and (57), and between and between (58) and (59), where the infinitive clauses are omitted from the latter sentence set in each pair.

- (59) i. Cars are illegal.
 ii. Dogs are illegal.
 iii. Women will be illegal.
 iv. Minors are illegal.
 v. Flags are illegal.

B-type speakers use the *tough* predicate lexical structure of *illegal*₁ to drive the parser. Since they assign no external θ -role to the matrix subject of (58), *illegal* need not be predicated of the subjects. I will discuss this strategy directly in "The Constituent Structure of 'John is easy to talk to'".

4.7 THE CONSTITUENT STRUCTURE OF 'JOHN IS EASY TO TALK TO'

Tough-movement constructions, like (60), have been problematic for generative grammarians since the introduction of transformational grammar.

(60) John is easy to please.

Within the Chomsky (1986a) account of the principles and parameters theory, (60) is underivable. The same mechanisms of Chomsky's system which block (8) will also block (60) and all other *tough*-movement constructions. However, *tough*-movement constructions, like those in (61), happen to be perfectly well-formed English sentences.

- (61) i. John is easy to talk to.
 ii. John is amusing to talk to.
 iii. John is impossible to talk to.

Although Chomsky makes no claim that (8) and (60) have the same S-structure, the complement in *tough*-movement predicates will violate Binding Principle A. As I have shown in "'Stubborn' versus 'Illegal'" on page 78, Chomsky neglects to specify what constituent structure he assumes for (8). On the other hand, it is safe to assume that the infinitive in *tough*-movement constructions is an AP complement. Regardless of where the infinitive may be attached, however, the chain formed by movement will still result in a Binding Principle A violation.

According to Chomsky's analysis of (8), *John* has to move up from the object position of *talk* because the subject position of *is illegal* is a non- θ -position.

(8) John is illegal to talk to.

As with (8), all of the subject positions of the adjective phrases in *tough*-constructions, as in (61), are non- θ -positions, as is illustrated by their ability to contain expletives in (62).

- (62) i. It is easy to talk to John.
 ii. It is amusing to talk to John.
 iii. It is impossible to talk to John.

Therefore, by the same argument with which Chomsky motivates his movement analysis of (8), the surface subjects of (61) should also have moved up from the subordinate clause. Thus (60) would have the chain structure in (63).

(63) $John_i$ is [*AP* easy [e_i [PRO to please e_i]]]

The chain structure shown in (63) violates the same binding constraint that (8) does. However, all of the examples in (62) are perfectly well-formed.

The analysis of *tough* movement in Chomsky (1977b) is no longer viable within the Chomsky (1986a) principles-and-parameters framework. Chomsky (1977b) analyzes *John* as being base generated in the subject position of *is easy*. He posits the *wh*-element *who* as the complement of *talk to*. The *wh*-element then moves up to the complementizer position of the subordinate clause. This complementizer then gets obligatorily deleted. Thus (64) has the S-structure (65).

(64) John is easy (for us) to please.

(65) John is easy (for us) [[*who* for] PRO to please *t*]

↓
ϕ

Basically, this Chomsky (1977b) *tough*-movement analysis is the same as the Chomsky (1986a) empty operator analysis for (7), as shown in Chomsky's example <114>.

(7) John is too stubborn to talk to.

<114> $John_i$ is too stubborn [0_k [PRO to talk to e_k]]

In the Chomsky (1986a) framework however, *John* cannot be base generated in the subject position, because the subject position of *is easy* is a non- θ -position. The remaining option is to analyze (60) on

a par with (8), with *John* moving up to the matrix subject position through successive applications of Move- α . This would result in a Binding Principle A violation. The NP-trace, an anaphor, is A'-bound in a local domain.

Chomsky argues that within the Move- α model presented in Chomsky (1986a), (8) is underivable. *John* cannot be base generated as the matrix subject because that is a non- θ -position. *John* cannot be base generated as the subject of the subordinate clause because the empty category in that position has a reference distinct from *John*. The only remaining option is to base generate *John* as the object of the subordinate clause, with *John* moving up from the object position of the subordinate clause to the subject of the matrix clause through successive applications of Move- α . An ill-formed S-structure chain would then link the resulting traces to its antecedent.

However, (60) is equally underivable. The matrix subject position of (60) is an A-position, since all subject positions are A-positions. This matrix subject position is also a non- θ -position because *easy* does not assign an external θ -role. Since the matrix subject is a non- θ -position, *John* cannot be base generated in the matrix clause. However, *John* cannot move up from the complement clause either. Such movement would result in an ill-formed chain. In terms of Binding Principle A, the NP trace, an anaphor, is not A-bound in a local domain. In terms of case theory, the resulting chain is ill-formed because it is associated with two distinct case roles. The empty operator analysis in conjunction with strong binding is barred because then *John* would have to be base generated in a non- θ -position in the matrix clause.

Finally, *John* cannot be inserted after D-structure as this would constitute a violation of the projection principle.

In keeping with tradition, I will use the adjective *easy* to illustrate how my parser can handle *tough* constructions. Like *illegal*₁, the predicative adjective *easy* subcategorizes for an infinitival CP and does not assign an external θ -role. It has the partial lexical structure (69).

(69) *easy* $\left[\begin{array}{l} \text{adjective} \\ \text{--- CP} \end{array} \right]$

Tough predicates exhibit features associated with the presence of an empty operator: (1) a missing argument in the complement clause corefers with the matrix subject and (2) PRO has arbitrary interpretation. The presence of an empty operator is obligatory in *tough* constructions. The embedded verb may not be intransitive, as illustrated by the ill-formedness of (66).

(66) * John is easy to dream.

Nor may the object position of the embedded clause be filled, as illustrated by the ill-formedness of (67).

(67) * John is easy to please Mary.

In order for a lexical parser to assign a well-formed S-structure to (68),

(68) John is easy to talk to.

adjustments must be made to Chomsky's (1986a) partial description of a parser which assigns the ill-formed S-structure (54) to (8) (see "Lexical Structures in 'John is illegal to talk to'" on page 112).

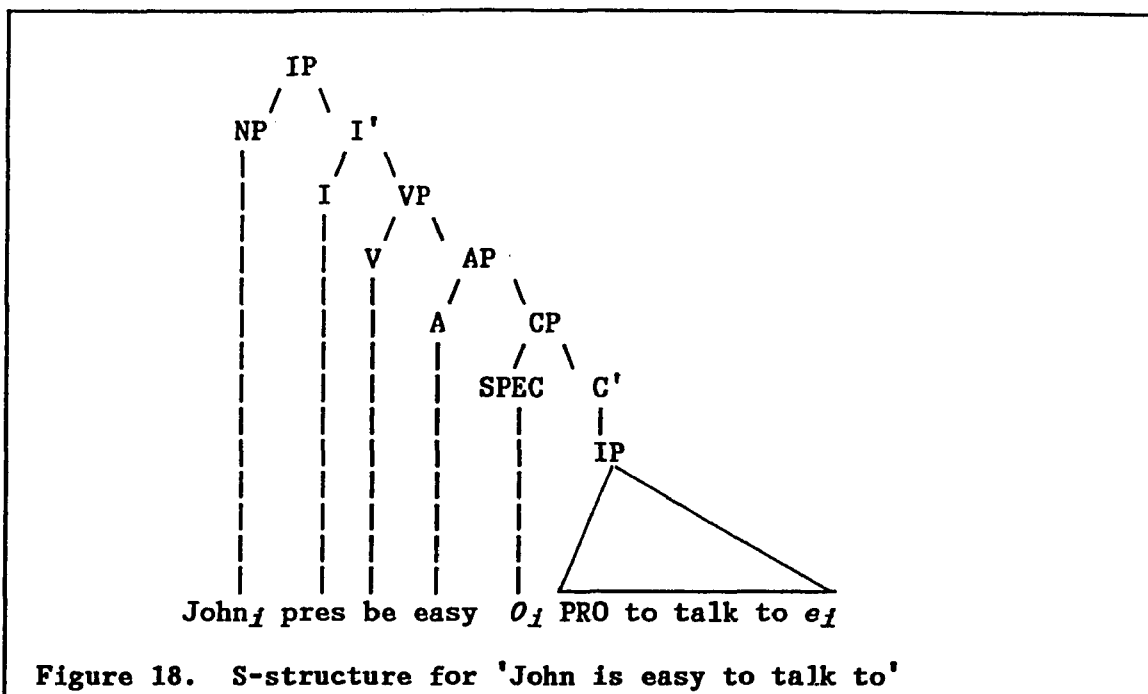
(8) John is illegal to talk to.

(54) $John_f$ is illegal [e_f [PRO to talk to e_f]]

To get a well-formed S-structure for (68), the revised parser requires the version of Binding Principle C which states that an r-expression is A-free *in the domain of its operator*.

I assume that the parser licenses structure for all clauses simultaneously, or in parallel. However, for ease of exposition, I will describe the parse beginning with the embedded clause in (60).

1. The lexical entry for *talk* licenses a PP complement, which the parser finds. Since the PP is missing its object, the parser inserts an empty category as the object of the preposition *to*. As prepositions are case assigners, the object of the preposition is in a case marked argument position, so the empty category is determined to be a variable.
2. The parser then looks for the subject of the embedded clause. Not finding an overt subject, the parser inserts an empty category in the subject position of *to talk*.
3. Since the subject position is ungoverned, the empty category is determined to be PRO. Since there is no overt wh-element in the operator position of the CP to bind the variable object of *to*, the parser inserts an empty category in this position which binds the variable.



4. Looking at the lexical structure of *easy*, the parser sees that it takes a clausal complement, and attaches the subordinate clause to the VP.
5. Since *easy* does not assign an external θ -role when it has a clausal complement,⁸ the subject *John* must be part of a chain which carries a θ -role. There is no θ -role-carrying empty category available for *John* to bind. However, the constructed chain [CP e_i [IP... e_i]] is available for strong binding by *John*. Thus the chain is determined to be an empty operator chain ([CP O_i [IP... e_i]]). The resulting structure (70) does not violate Binding Principle C because the

⁸ Perhaps *easy* does assign an external θ -role when it is intransitive as in *The test was easy*. However, the subject position is not assigned a θ -role by the adjective when there is a clausal complement as illustrated by: *It is easy to talk to John*, **John is easy to talk to Bill*, and **Who is John easy to talk to*.

variable is A-free in the domain of its operator. In addition, the variable is now strongly bound, as required, by *John*.

(70) $John_i$ is easy [O_i [PRO to talk to e_i]]

The resulting S-structure is shown in Figure 18.

Note that this proposed S-structure for *tough* predicates is essentially that proposed in Chomsky (1977b). However, with my parsing model, the question of where *John* was base generated does not arise. The question of base generation makes no sense from the point of view of a lexically driven parser, since the input to the parser is the surface string. Since the parser is not deriving the S-structure from a D-structure, information about base generation is inaccessible and therefore has to be irrelevant to the parser's construction of S-structure. The relevant issue is whether the NP in subject position of *easy* is licensed at S-structure. If we require that each chain (or set of indexed structures) has one overt member (which is the effect of strong binding), then the structure in (70) satisfies this requirement. Such a requirement is, in effect, a condition against vacuous movement.

However, we are left with the problem of why strong binding in (70) appears to satisfy the Theta Criterion while strong binding in (71) does not result in a Theta Criterion violation after identical indices are assigned to two θ -positions: *John* and the empty operator.

(71) $John_i$ is too stubborn [O_i [PRO to talk to e_i]]

This problem of a chain being associated with two θ -positions via strong binding arises for all adjunct constructions where strong binding has been proposed, including binding of relative clauses and parasitic gaps.

As we have noted, Chomsky does not address this issue in his discussion of the empty operator analysis and strong binding.

The difference between the structures in (70) and (71) is that the former has an empty operator within the argument structure of the predicate, while the empty operator of the latter is in an adjunct position. It appears that strong binding can transmit a θ -role when the operator is in the argument structure of the predicate. We can assume that the parser's goal is to license each constituent in the S-structure. When it encounters an overt NP, the parser tries to license the NP by associating it with, among other things (like case), a θ -role. In (70), *John* is not in a θ -position. In this case, the parser associates *John* with a θ -role through the strong binding relation. In looking for a θ -role to associate *John* with, the parser is limited to the complement structure of the verb or adjective. It cannot get the θ -role via a strong binding relation between a noun and an adjunct. Thus an extra θ -role is not associated with *John* in (71).

4.8 THE CONSTITUENT STRUCTURE OF 'MARY IS PRETTY TO LOOK AT'

The class of predicates illustrated in (72) are discussed in Lasnik and Fiengo (1974).

(72) *Mary is pretty to look at.*

This class of predicates, which I call *pretty* predicates, exhibits the same empty operator effects that we have seen with *tough* predicates and *too* complements. There is no overt argument in object position, the object position is bound by the matrix subject, and PRO is free.

L&F argue that the infinitival clause in (72), as with *tough*-movement constructions, is a complement to the adjective. They propose an obligatory object-deletion rule which applies to the infinitival clause in both structures. Thus L&F call this class of predicates "pure OD" (object deletion) structures, and offer some examples of *pretty*-type adjectives: *pretty, delicious, fragrant, graceful, melodious, tasty, cacophonous, beautiful, slippery, pungent, scratchy* (L&F 1974:566).

Pretty predicates assign an external θ -role, as illustrated by the ill-formedness of (73), which has an expletive in subject position of the matrix clause.

(73) * It is pretty to look at Mary.

They also require that the infinitive clause contain a transitive verb, as shown in (74), with an empty category in object position, as in (75).

(74) * Mary is pretty to dream.

(75) * Mary is pretty to look at Jane.

In addition, PRO is never controlled by the matrix subject, as in (76).

(76) * Mary_{*i*} is pretty [_{*e*} [PRO_{*i*} to look at _{*e*}]]

This distribution for data is the diagnostic of the presence of the empty operator. However, the infinitive clause in (72) differs from the complement of *too* and *enough* in that the presence of the empty operator is not optional. The empty operator, which is licensed only by a missing r-expression, is required.

Although the infinitive in (72) resembles the complement in *tough* predicates in some respects, it differs in other respects. Unlike *pretty*

predicates, *tough* predicates do not assign an external θ -role, as in (77).

(77) It is easy to look at Mary.

It follows from the fact that *pretty* predicates require an external θ -role that they require an empty operator in (73). *Tough* predicates show no empty operator effects when they have expletive subjects, as in (77), where the object position is lexically filled.

Pretty predicates are more restricted, both syntactically and semantically, than *tough* predicates. *Tough* predicate complements can appear in subject position, as in (78ii), whereas the infinitive cannot be in the subject position of *pretty* predicates, as in (79i).

- (78) i. Mary is pleasant to visit.
 ii. Visiting Mary is pleasant.

- (79) i. Mary is pretty to look at.
 ii. * Looking at Mary is pretty.

Chomsky (1977b) notes that the infinitive clauses of *pretty* predicates have to be adjacent to the adjective. With *tough* predicates there can be an intervening clause. He also notes that the range of verbs in the infinitive of *pretty* predicates are more restricted than the range of infinitives allowed with *tough* predicates. Chomsky, like I&F, suggests that the constituent structure of *pretty* predicates is the same as that of *tough* predicates. He adds, however, that *pretty* predicates may be interpreted as idioms:

Other structures similar to <121> are much more restricted in scope, e.g., <167>:

[<121> John is easy (for us) [to please]]

<167> Mary is pretty to look at

In this case, we do not have the full range of properties [of *wh*-movement]. Thus there is no form <168>, analogous to <127b>:

- [<127> b. John is easy (for us) to convince Bill to do business with t
John is easy (for us) to convince Bill to arrange for Mary to meet t]

<168> Mary is pretty to tell Bill to look at

furthermore, in such cases as <167> there are very narrow restrictions on the choice of the matrix adjective and embedded verb. We may propose the same analysis as in the *easy-to-please* cases, but with idiom interpretation rules associated with the adjectives in question. Note that there are structures such as <169>, but in this case the embedded complement is not associated with the adjective but with the adjective qualifier, *too*:

<169> Mary is too pretty to expect anyone to look at (her)

(Chomsky 1977b:109)

An example of the restrictiveness of the relationship between the matrix adjective and the embedded verb is in (80), as compared to (81).

(80) * Mary is pretty to dance with.

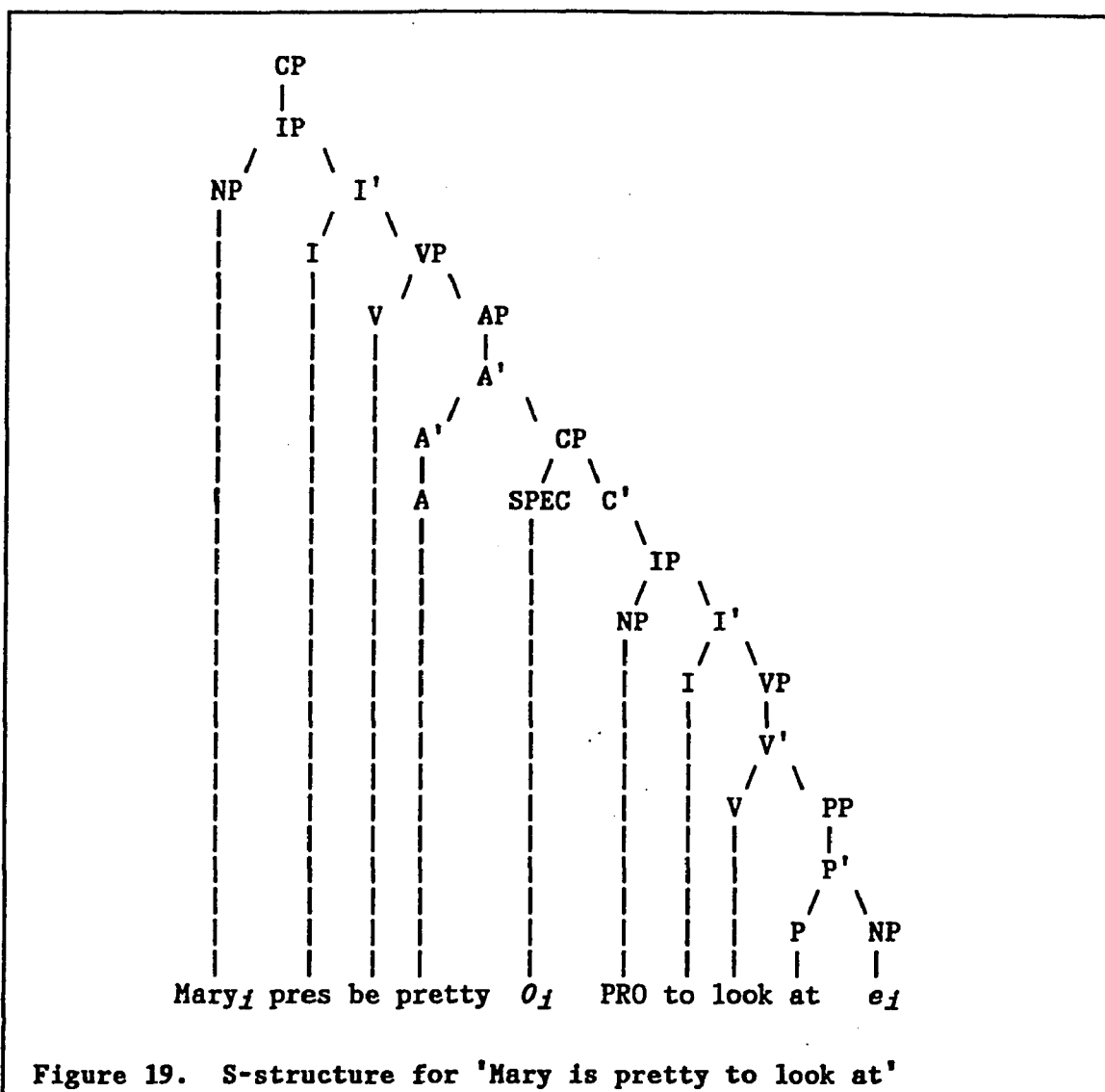
(81) Mary is easy to dance with.

The choice of verb in the infinitive clause with *pretty* is restricted to one of *looking*.

I propose that the infinitive of *pretty* predicates is an adjunct clause, and that (72) has the constituent structure shown in Figure 19 on page 129. Like an adjunct, it is always optional. The truth value of the matrix sentence is in no way altered when the infinitive is omitted. In the examples below, (ii) is always implied by (i):

(82) i. Mary is pretty to look at.
ii. Mary is pretty.

(83) i. The apple is delicious to eat.
ii. The apple is delicious.



This is not the case with the complements in *tough* predicates. Often the complement is obligatory, as in (84). When the complement is optional, as in (85), the meaning of the matrix clause is modified; (85i) does not imply (85ii).

- (84) i. John is easy to please.
 ii. * John is easy.

- (85) i. John is difficult to understand.
 ii. John is difficult.

Since adjuncts are, by definition, non-lexically specified, they ought to occur freely. However, the distribution of the infinitive with *pretty* predicates is so highly constrained that Chomsky suggests that they may be idiomatic.

If adjuncts are licensed by the thematic structure of the predicate, then *pretty* predicates must form a thematic class. External θ -role assignment has already been identified as one aspect of lexical structure which has affects outside of the domain of the head. External θ -role assignment also affects the distribution of this adjunct. *Pretty* predicates, which take this AP adjunct assign an external θ -role. The thematic role of the external argument is that of *patient*. In addition, although the subjects are usually non-animate, as in (87), they can be animate, as in (86ii).

- (86) i. The painting is pretty to look at.
 ii. Mary is pretty to look at.
- (87) i. The music is melodious to listen to.
 ii. Mary's voice is melodious to listen to.
 iii. * Mary is melodious to listen to.

I will refer to the thematic role of the subject of *pretty* predicates as that of passive-patient.

With respect to meaning, L&F describe the adjectives as denoting "subjective evaluation":

The adjectives and nouns in the list above denote a subjective evaluation; the subjective evaluation is almost always that of the speaker. The adjectives and nouns are subjective in that nonobjective criteria determine their sense. Thus adjectives such as *rectangular* and *black*, which are objective in this sense, cannot occur in this subcategorization frame:

<B.1> * This room is rectangular to live in.

<B.2> * Coal is black to look at.

(Lasnik and Fiengo 1974:567)

However, a semantically similar adjunct is attached to adjectives denoting objective criteria. *Black* and *rectangular* both take a PP adjunct which serves much the same purpose as the infinitive in *pretty* predicates:

(88) This room is rectangular in shape.

(89) Coal is black in color.

Chomsky (1977a) calls them "adjective qualifier constructions". However, the infinitive does not appear to be a qualifier. It does not limit or modify the meaning of the adjective phrase. As the examples in (82) and (83) show, the infinitive does not change the truth value of the matrix clause.

The infinitive that appears with *pretty* is an appositive adjunct rather than a complement clause, which serves as a kind of intensifier. An appositive adjunct on APs can be introduced with phrases such as *that is*, or *namely*, as in (90)-(92):

(90) i. Mary is pretty, that is, to look at.
ii. Mary is pretty, namely, to look at.

(91) i. This room is rectangular, that is, in shape.
ii. This room is rectangular, namely, in shape.

(92) i. Coal is black, that is, in color.
ii. Coal is black, namely, in color.

The adjunct answers the question: "How?"

With respect to distribution, this adjunct is restricted to adjectives, as in (93), and nouns, as in (94), and does not appear in verb phrases, as in (95).

(93) This apple is delicious to eat.

(94) Nureyev is a marvel to watch.

(95) * This apple tastes delicious to eat.

A curious property of predicates which license the appositive adjunct is that they do not license complements. The class of *pretty* adjectives, which licenses the appositive adjunct, does not assign complement clauses. This is not a general property of adjectives which assign external- θ roles, as in (96).

- (96) i. John is afraid to go outside.
 ii. John is afraid that the cat will go outside.

Nor is it a general property of predicates which have thematic patients as external arguments, as in (97):

(97) John is unable to go outside.

However, *pretty* predicates never take a clausal complement, either with an adjunct, as in (98), or without an adjunct, as in (99).

(98) * Mary is pretty to expect John to look at.

(99) * Mary is pretty that John look at her.

Lasnik and Fiengo point out that *like* also forms *pretty* predicate constructions:

The determiner *like* serves to form a productive class of OD constructions with predicate nominals, as in the following constructions.

- <B.6> a. This paint is like concrete to work with.
 b. * This paint is concrete to work with.

- <B.7> a. This coffee is like tar to drink.
 b. * This coffee is tar to drink.

- <B.8> a. John is like Nixon to talk to.
 b. *John is Nixon to talk to.

(Lasnik and Fiengo 1974:567)

I propose a different analysis for the sentences <B.6a>-<B.8a>. These Sentences appear to be predicate adjective constructions with passive-patient adjuncts just like (72).

(72) Mary is pretty to look at.

The adjective *like* assigns an external θ -role and an internal θ -role to an NP. It does not license a clausal complement.

(100) *like*:

$$\left[\begin{array}{l} [+V, +N] \\ P\text{-external } \theta\text{-role} \\ NP \end{array} \right]$$

Thus in (101), *tar* is the object of the adjective *like*, and *to drink* is an AP adjunct.

(101) This coffee is like tar to drink.

As is to be expected with this type of adjunct, it answers the question: "How?" With this analysis, one would expect phrases with *like NP* to exhibit empty operator effects.

4.9 THE TWO CONSTITUENT STRUCTURES OF 'JOHN IS ILLEGAL TO TALK TO'

Due to the two lexical structures for *illegal*, the parser is presented with an ambiguity when presented with (8):

(8) John is illegal to talk to.

Both lexical structures for *illegal* license this string. A parse for (8) can be driven by both *illegal*₁ and *illegal*₂.

According to the lexical structure *illegal*₁, *illegal*₂ is the head of a *tough* predicate which assigns no external θ -role and takes a clausal complement, as in (102):

(102) It is illegal to rob banks.

Since *illegal*₁ does not assign an external θ -role, the *tough* predicate analysis is also assigned to predicates where the external argument is filled with things which are not inherently illegal, such as in (103):

(103) Flags are illegal to burn.

According to the lexical structure *illegal*₂, *illegal* is intransitive, assigning an external θ -role to the subject, as in (104):

(104) Unlicensed pistols are illegal.

Since *illegal*₂ is intransitive, an infinitive clause is assigned as an AP adjunct, as in (105).

(105) Unlicensed pistols are illegal to own.

However, as we have seen, the distribution of the AP adjunct is very restricted. The only examples we have encountered thus far are those involved with the five senses: taste, touch, smell, hearing, and sight. What do these *pretty* predicates have in common with *illegal*?

*Illegal*₂ assigns the same thematic role to its subject that *pretty* predicates do--that of passive-patient. The θ -marked subject exhibits the properties typical of this role: It is a patient which can be inanimate. As with the other *pretty* predicates, the truth value of the sentence does not change when the adjunct is omitted, as in (106).

(106) Unlicensed pistols are illegal.

Similarly, the matrix sentence and adjunct can be paraphrased as an appositive in (107).

(107) Unlicensed pistols are illegal, that is, to own.

When the implication relation does not hold, the adjunct analysis is not permitted, as in (108).

(108) * Unlicensed pistols are illegal to lose.

Now we can explain the two sets of intuitions for sentences containing *illegal* described in "Lexical Structures in 'John is illegal to talk to'" on page 112. Those speakers who accept both (8) and (103) are parsing them as *tough* predicates, with a non-thematic subject and complement clause.

(8) John is illegal to talk to.

(103) Flags are illegal to burn.

Those speakers who dislike both (8) and (103) are parsing them as *pretty* predicates with thematic subjects and appositive adjuncts. However, the adjunct structure is bad because the implication relation does not hold. By the *pretty* analysis, (8) and (103) have θ -marked subjects. However, (8) and (103) do not imply (109) and (110), whose subjects should be identically θ -marked.

(109) John is illegal.

(110) Flags are illegal.

Those speakers who do not accept (8) but do accept (103) are assigning different lexical structures to the two sentences. They assign the *tough* analysis to (103) and the *pretty* analysis to (8).

The question of why more predicative adjectives do not pose this kind of ambiguity remains. Many *tough* predicates also have lexical structures which assign an external θ -role. We have already discussed one such predicate, repeated below:

- (111) i. John is difficult to talk to.
 ii. John is difficult.

However, with *difficult*, as with other *tough* predicates, the subject is not thematically a *patient* when they θ mark an external argument.

PLEASE NOTE

**Page(s) not included with original material
and unavailable from author or university.
Filmed as received.**

page 136

University Microfilms International

when either the subject or the infinitival clause is not specified in the lexical structure of the adjective.

Afraid predicates assign an external θ -role and never license an empty operator. They pattern after (114):

- (114) i. * it is afraid [PRO to talk to John]
 ii. John_i is afraid [PRO_i to talk]
 iii. * John_i is afraid [O_i [PRO_{arb} to talk to e_i]]
 iv. John_i is afraid [PRO_i to talk to Mary]

The infinitives in (112) are complements and not adjuncts. The selectional restrictions for the subject position for *afraid*-class predicates are [+ANIMATE], or sometimes [+HUMAN]. For example, the predicates are well-formed in (115) with human subjects, but ill-formed in (116) with inanimate subjects.

- (115) i. John is afraid to watch the film.
 ii. John is eager to watch the film.
 iii. John is qualified to criticize the film.
 iv. John is welcome to watch the film.

- (116) i. * The film is afraid to be watched.
 ii. * The film is eager to be watched.
 iii. * The film is qualified to be criticized.
 iv. * The film is welcome to be shown.

The thematic role of *afraid* predicate subjects does not license the appositive adjunct.

Tough predicates do not assign an external θ -role, and optionally license an empty operator. *Tough* predicates pattern after (117):

- (117) i. it is easy [PRO_{arb} to talk to John]
 ii. * John is easy [PRO to talk]
 iii. John is easy [O_i [PRO_{arb} to talk to e_i]]
 iv. * John is easy [PRO to talk to Mary]

The parser inserts the empty operator only when the operator can be licensed by an empty category in the clause. The complements in (117i),

(117ii) and (117iv) are not missing any arguments, therefore the empty operator is not licensed. The reason that (117i) is well-formed while (117ii) and (117iv) are not is that the latter sentences constitute Theta Criterion violations. In neither sentence is *John* associated with a θ -role.

Like *pretty* predicates, *tough* predicates take either animate subjects, as in (118), or inanimate subjects, as in (119).

- (118) i. John is easy to talk to.
 ii. John is impossible to dance with.
 iii. John is stimulating to talk to.
 iv. John is exciting to visit.
- (119) i. The music is easy to listen to.
 ii. The music is impossible to dance to.
 iii. The music is stimulating to listen to.
 iv. The music is exciting to listen to.

Unlike *pretty* predicates, *tough* predicates do not assign an external θ -role. Since *tough* predicates do not assign an external θ -role, they do not select for a subject, and thereby cannot be in the thematic class which takes the appositive adjunct. The infinitive in *tough* predicates is a clausal complement.

The comparative formatives *too* and *enough*, also license the empty operator construction. As with *tough* predicates, the parser licenses an empty operator only when an internal argument is missing in the embedded sentence.

- (120) i. John_i is [too [stubborn] [PRO_i to talk]]
 ii. John_i is [too [stubborn] [O_i [PRO to talk to e_i]]]
 iii. John_i is [too [stubborn] [PRO_i to talk to Mary]]

Pretty predicates assign the passive-patient external θ -role which licenses an adjunct. They do not take complements. These adjuncts obligatorily contain an empty operator. They pattern after (121):

- (121) i. * it is pretty [PRO to look at Mary]
 ii. * Mary is pretty [PRO to smile]
 iii. Mary_i is pretty [O_i [PRO_{arb} to look at e_i]]
 iv. * Mary is pretty [PRO to look at John]

As with relative clauses, the adjunct with *pretty* predicates is predicated of its head and requires an empty operator.

I have suggested that a predicative adjective which both assigns an external θ -role and a clausal complement does not license an empty operator. However, *ready* appears to pose a counter-example to my hypothesis. *Ready* assigns an external θ -role, as in (122), and takes an infinitive clause, as in (123).

(122) * It is ready to leave.

(123) John is ready to leave.

According to my analysis, this infinitive must either be (1) a complement or (2) an adjunct. If the infinitive is a complement, then *ready* is an *afraid* predicate and the empty operator is not licensed. If the infinitive is an adjunct, then the empty operator analysis is obligatory. However, *ready* appears to optionally get the empty operator effects in (124).

- (124) i. The lamb is ready to eat.
 ii. the lamb_i is ready [PRO_i to eat]
 iii. the lamb_i is ready [O_i [PRO_{arb} to eat e_i]]

The reading of (124ii), where *the lamb* controls PRO, is that a (live) lamb is ready to eat something. The reading of (124iii), where an empty operator binds the object EC, a (cooked) lamb is ready to be eaten by someone.

Just as *illegal* is ambiguously a *tough* and *pretty* predicate, *ready* is ambiguously an *afraid* and *pretty* predicate. When *ready* has a human

subject, it acts like an *afraid* predicate, patterning after (114), as in (125).

- (125) i. * it is ready [PRO to talk to John]
 ii. John_i is ready [PRO_i to talk]
 iii. * John_i is ready [O_i [PRO_{arb} to talk to e_i]]
 iv. John_i is ready [PRO_i to talk to Mary]

In (125) the infinitive is a complement. Notice that the thematic role assigned to *Mary* in these sentences is not the passive-patient role.

When *ready* has an inanimate subject, it acts like a *pretty* predicate, patterning after (121), as in (126).

- (126) i. * it is ready [PRO to move into the house]
 ii. * the house is ready [PRO to live]
 iii. the house_i is ready [O_i [PRO_{arb} to move into e_i]]
 iv. * the house is ready [PRO to move into the city]

In (126) the infinitive is an adjunct. As with *pretty* predicates, the subject of these sentences is usually inanimate and thematically a passive-patient.

5.0 CONTROL FACTS AND CONTROL THEORIES

Control Theory is the component of UG which is concerned with how the reference of PRO is determined. In this chapter, the control theories proposed by Chomsky (1980a), Williams (1980), Manzini (1983), Koster (1984, 1987), and Hornstein and Lightfoot (1987) are outlined. I present the control properties of the verb *ask*, and show how each theory would handle these properties.

Typically, theories of control try to capture control facts between PRO and a possible antecedent through structural relations such as government or c-command. Facts I will introduce indicate that, in certain cases, government and c-command cannot determine the distribution and interpretation of PRO.

5.1 THE CONTROL PROPERTIES OF "ASK"

The verb *ask* has object control, as in (1), unless the infinitival clause is headed by a *wh*- element. When the infinitive clause is headed by a *wh*- element, *ask* has subject control, as in (2) and (3). The subject of *put* is *Mary* in (1) and *John* in (2) and (3), as illustrated by the indexed S-structures (4), (5), and (6).

- (1) John asked Mary to put the book on the table.
- (2) John asked Mary how to put the book on the table.
- (3) John asked Mary where to put the book.
- (4) John_{*i*} asked Mary_{*j*} [PRO_{*j*} to put the book on the table]
- (5) John_{*i*} asked Mary_{*j*} [how [PRO_{*i*} to put the book on the table]]
- (6) John_{*i*} asked Mary_{*j*} [where_{*k*} [PRO_{*i*} to put the book e_{*k*}]]

These control properties of *ask* are described in Chomsky (1977a). He writes:

Summarizing, *ask* takes an S complement or an NP S complement. In the latter case, a lexical property indicates that the embedded COMP can be -WH only if the subject of the embedded sentence is PRO. Where the subject of the embedded S is PRO, PRO is controlled necessarily by the subject of *ask* if the complement is merely S; PRO is controlled by the subject of *ask* if the complement is NP S and the COMP of the embedded S is +WH; PRO is controlled by the NP following *ask* if the COMP of the embedded S is -WH. (We might seek a further explanation of these facts, but that is another matter). Given the idiosyncracies noted, the distribution of grammatical and ungrammatical sentences and their interpretations follows from the rules already given.
(Chomsky 1977a:150)

I represent this description below.

<p>ask: __ S CONTROL=SUBJECT __ NP S if COMP=[-WH] then S=infinitive and CONTROL=OBJECT if COMP=[+WH] then CONTROL=SUBJECT</p>
--

Figure 20. Chomsky (1977a) description of *ask*.

Among the "lexical idiosyncracies" he notes is that *ask* requires the subjunctive in the environment [__ S [-WH COMP]] as in (7) but takes no -WH COMP when it has a direct object as in (8).

(7) John asked that Bill leave early.

(8) * John asked me that Bill leave early.

However, as we will see, the control properties exhibited by *ask* appear to be more complex. In the next section, I propose a lexical structure for *ask* which accounts for the control properties of *ask* as exhibited in (1)-(3) and in other configurations.

5.1.1 THE LEXICAL STRUCTURE OF "ASK"

There are two senses of *ask* which take sentential complements. These two senses are (1) a request for information or permission and (2) a request for action. I will call the former *ask_{info}* and the latter *ask_{act}*. Below I give the subcategorizations for these two senses of *ask*, when they take sentential complements:

1. *ask_{info}*: To request information or permission.
 - a. — CP
 - 1) CP=indirect question

I asked who he was.
I asked where to go.
He asks if you are ready.
 - 2) CP=active infinitive

They asked to go for a walk.
 - 3) CP=subjunctive

I asked that I (should) be allowed to see her.
 - b. — NP CP: CP=indirect question

I asked him who he was.
I asked him where to go.
 - c. — NP (PP(for)) CP : CP=infinitive

She asked the teacher for permission to go.
She asked the teacher to be excused.
John asked the teacher to be allowed to leave.
2. *ask_{act}*: to make a request for action
 - a. — CP : CP=passive infinitive

She asked to be woken at 6 o'clock.
Mary asked to be assigned a seat.
 - b. — NP CP : CP=infinitive

She asked him to wake her at 6 o'clock.
Mary asked the teacher to assign her a seat.

When *ask* is understood as a request for action, it manifests object control; when it is understood as a request for permission or information, it manifests subject control. This describes for control alternation exhibited in (9) and (10):

- (9) John asked the guard_i [PRO_i to admit him]

(10) John_i asked the guard [PRO_i to be admitted]

In (9), the request is for action and the guard is the controller. In (10), the request is for permission, and *John* is the controller. Thus *ask_{info}* exhibits subject control and *ask_{act}* exhibits object control.

When *ask_{info}* has no direct object, it takes indirect questions, infinitives and subjunctives. When *ask_{info}* takes a direct object, it usually only takes indirect questions. When it has a PP(for) such as *for permission* or *for leave*, it also takes an infinitive. Sometimes, when the context makes the choice of antecedent obvious, this PP can be omitted. Compare, for example, the interpretation of (11) where, on one reading, *John* is the one to leave the room, and (12) where only *Mary* can be the one to leave.

(11) John asked the teacher to leave the room.

(12) John asked Mary to leave the room.

In the case of (11) and (12), the two senses of *ask* have the same syntactic realization. However, the *ask_{act}* interpretation is preferred, as in (12), except in a very explicit context, as in (11). *Ask_{act}* takes only infinitive complements whether it has a direct object or not. It never takes an indirect question or a subjunctive.

5.2 CHOMSKY'S CONTROL THEORY IN "ON BINDING"

The control theory proposed in Chomsky (1980a) is motivated, in part, by the control properties of *ask*. I will outline the theory and illustrate how it correctly predicts many of the control properties of *ask*. However, it provides no account for sentences like (13) which have an arbitrary interpretation when the COMP node is unfilled:

(13) It is fun PRO_{arb} to swim.

In addition, it does not account for the double-passive constructions such as (14).

(14) John was promised [PRO to be allowed [PRO to leave]]

This control theory is a procedure which assigns indices to unindexed NPs; it is not a filter on the output of free-indexing:

We assume that Control applies in the course of a systematic indexing procedure for NPs to which we return in the appendix. Assume the procedure to be "top-to-bottom", in the sense that an index is assigned to an NP only if all NPs *c*-commanding or dominating it have been indexed. NPs not yet indexed by a movement rule will be indexed by this procedure; the procedure assigns but does not change indices. A controller, then, can be a lexical NP, trace, or PRO already assigned an index by Control ... (Chomsky 1980a:33)

The theory assumes that the only indices on the input string are the residue of Move- α . That is, only inherent chains are indexed. In this theory, NPs that have not moved (i.e. single-membered chains) are not indexed in the D-structure.

The theory requires three lexical features: [SC], [Control] and [F]. [+SC] indicates that the verb assigns subject control. For example, *promise* has [+SC] in its lexical entry. [\pm Control] indicates whether a verb is a control verb or a non-control verb.

Evidently, verbs with infinitival complements fall into two categories in English with regard to control: noncontrol verbs such as *believe* and control verbs such as *try*, *persuade*, and *promise* (Chomsky 1980a:24)

The environment for obligatory control is when a [+Control] verb has an infinitive complement. If a verb is marked with [+F] then Case is assigned across a clause boundary. [+F] verbs, such as *believe*, have since become known as S'-deletion verbs. Verbs marked with [-Control]

are also [+F]. Chomsky (1980a) proposes the redundancy rule (15) which makes the relationship between [Control] and [F] explicit.

- (15) a. If [+Control] then [-F].
 b. If [-Control] then [+F].

Verbs with the features of (15a) are obligatory control verbs, and verbs with the features of (15b) are not.

Chomsky's (1980a) control theory only covers control when the verb has the features in (15a). If a verb takes a *for*-complementizer, then it is not a control verb. That is, verbs that take *for*-complementizers are described by (15b). Chomsky writes:

Only [+F] verbs take such complements as <85> ...

<85> [S'[COMP for][S NP to VP]]

(Chomsky 1980a:30)

Since [+F] verbs are [-Control], verbs which take *for*-complementizers are not subject to control. Thus, presumably, the infinitives in (16) are not subject to control because they optionally take *for*-complementizers, as in (17).

- (16) i. I would hate [PRO to leave]
 ii. I would prefer [PRO to leave]

- (17) i. I would hate for John to leave.
 ii. I would prefer for John to leave.

Chomsky does not indicate how PRO gets indexed when it appears in the complement of [+F] verbs.

The theory also makes use of the notion "properly related," which Chomsky describes as follows:

The notion "properly related", still to be clarified, includes subjects, direct objects, and certain complements.
 (Chomsky 1980a:33)

Presumably the notion of being "properly related" is meant to distinguish arguments and complements from adjuncts, and can be correlated with what is now called θ -role assignment.

Finally, the theory contains the notion of "nearness". In this sense "nearness" is not determined simply by linear order but also by *c*-command: a *c*-commanding NP in the matrix clause is always "nearer" than a non-*c*-commanding NP:

A controller *c*-commanded by S' of <93> is "nearer" to [NP *e*] than one not *c*-commanded by S' ... (Chomsky 1980a:33)

Below in <95> is the control theory proposed in Chomsky (1980a). <93> is the structural description for control. It ensures that control applies only when the matrix verb does not take a *for*-complementizer. The structural description <93> also requires that control be a strictly local process. The mutual *c*-command restriction insures that the controller is in the clause immediately dominating the infinitive.

Consider now the theory of control for the context <93>:

<93> ... V ... [S' COMP ... [NP *e*] ...]
 where V = [-F] and V and S' *c*-command one another

(Chomsky 1980a:33)

<94> specifies which NPs in the controlling clause can be controllers. Assuming that the notion of being "properly related" is a precursor of θ -role assignment. Possible controllers, in present terminology, are NPs which have been assigned a θ -role by the verb. According to <94b> however, if the verb has the feature [+SC], then the subject NP is the only possible controller.

<94> NP is a *controller* for V in <93> if
 a. NP is an indexed NP properly related to V:
 b. If V = [+SC] then NP is the subject of V.

(Chomsky 1980a:33)

<95> is the rule of control which applies once <93> is satisfied and <94> has determined which NPs, if any, are possible controllers.

The rule of Control for the context <93> can be given as <95>:

<95> In <93>,

- a. if COMP ≠ null and V has no controller,
then [NP e] is assigned *arb*;
- b. [NP e] is assigned the index of the nearest controller.

(Chomsky 1980a:33)

This control algorithm is ordered, <95a> applies before <95b>.

Assume the order (a), (b) as given in <95>. We continue to assume that <95> is an index-assigning and not an index-chaining rule, like other rules of indexing, so that if *arb* is assigned by case (a), then case (b) is inapplicable; [NP e] must be without index for <95> to apply. (Chomsky 1980a:34)

The control algorithm also applies obligatorily when the structural description is matched.

Like other rules of the interpretive component, rule <95> is obligatory. In the context <93>, the rule of Control <95> must apply or the construction is assigned *. (Chomsky 1980a:34)

Chomsky's (1980a) control theory can do one of two things: Either <95a> applies and control assigns *arb* to PRO, or <95b> applies and control determines the antecedent for PRO. Since <95> is sensitive to two items, whether COMP is filled and whether V has a controller, four possible structures are defined by the control algorithm:

1. filled COMP, no controller: <95a> is satisfied and PRO is assigned *arb*:

(18) It is unclear [what [PRO_{*arb*} to do e]]

2. filled COMP, + controller: <95a> is not satisfied, so <95b> applies.

PRO is coindexed with the nearest controller:

(19) John_f wondered [who [PRO_f to visit e]]

(20) John told Bill_f [what [PRO_f to do e]]

If the verb is [+SC], then PRO is coindexed with the subject:

(21) John_f asked (Bill) [what [PRO_f to do e]]

3. empty COMP, no controller: This case is not covered by <95>: <95a> assigns *arb* only if COMP is filled and <95b> requires a controller. However the construction exists with predicate adjectives, as in (22).

(22) It is fun [PRO to swim]

Since neither <95a> nor <95b> apply to (22), (22) would be marked as ill-formed. Thus sentences like (22) present a counterexample to Chomsky's (1980a) control theory.

4. empty COMP, + controller: <95a> is not satisfied, and so <95b> applies, coindexing PRO with the nearest controller:

(23) John_f tried [PRO_f to leave]

(24) John persuaded Bill_f [PRO_f to leave]

If the verb is marked [+SC], then PRO is coindexed with the subject:

(25) John_f promised (Bill) [PRO_f to win]

To summarize, Chomsky's (1980a) control theory has three ordered strategies: First it assigns PRO the index *arb* based on (a) the lack of lexical control features and (b) the structural configuration. Second it finds an antecedent for PRO based on (a) lexical control

features and (b) the structural configuration. Third, and last, it starts any sentence which contains PRO that has not been assigned an index by either of the first two strategies.

5.2.1 CHOMSKY (1980A) CONTROL THEORY AND "ASK"

The control theory proposed in Chomsky (1980a) correctly predicts the control properties of (1) through (3), repeated below, and their corresponding passives (26) through (28).

- (1) John asked Mary to put the book on the table.
- (2) John asked Mary how to put the book on the table.
- (3) John asked Mary where to put the book.

- (26) Mary was asked to put the book on the table.
- (27) Mary was asked how to put the book on the table.
- (28) Mary was asked where to put the book.

Chomsky (1980a) claims that *ask* has the feature [+SC] only when it subcategorizes for an indirect question:

But certain verbs (e.g. *promise* or *ask* with an indirect question complement...) are marked in the lexicon with the feature [+SC] indicating "assigns subject control". (Chomsky 1980a:33)

(1) triggers <95b>. Since *ask* in (1) does not have the feature [+SC], PRO is coindexed with the nearest controller, as in (29).

- (29) John asked Mary_{*f*} [PRO_{*f*} to put the book on the table]

If no direct object appears, then the closest controller is the subject as in (30).

- (30) John_{*f*} asked [PRO_{*f*} to put the book on the table]

In (2) and (3), *ask* takes an indirect question, and so is marked [+SC]. Thus <95b> coindexes PRO with the subject as in (31) and (32).

- (31) John_{*f*} asked Mary [how [PRO_{*f*} to put the book on the table]]
- (32) John_{*f*} asked Mary [where [PRO_{*f*} to put the book e]]

(26) does not meet the criteria of <95a> as the COMP node is empty and the V has a controller. Therefore, by <95b>, PRO is coindexed with the controller, resulting in (33).

(33) Mary_{*f*} was asked e_{*f*} [PRO_{*f*} to put the book on the table]

The requirements of <95a> are met by (27) and (28). In both cases, COMP is filled by a *wh*-element which is base generated in the case of (27) and coindexed with the indirect object position of *put* in the case of (28). Since *ask* is [+SC] when it has an indirect question complement, and the passive clauses in (27) and (28) have no subjects, then there is no controller in either sentence. Therefore PRO is assigned *arb* by <95a>, as in (34) and (35).

(34) Mary_{*f*} was asked e_{*f*} [how [PRO_{*arb*} to put the book on the table]]

(35) Mary_{*f*} was asked e_{*f*} [where_{*f*} [PRO_{*arb*} to put the book e_{*f*}]]

Thus the control rule <95> accounts for the passive of *ask*:

We can therefore accommodate in a natural way the well-known resistance of verbs of subject control to passivization; there is no subject under passive, and hence no way for control to be assigned. Note that it would be incorrect to say that verbs of subject control do not passivize, as we can see from <92e>, contrary to what is stated in Chomsky (1977a, 14). Rather, such verbs do not passivize unless arbitrary control is assigned in the complement. That is, the operative principle seems to be that the Control rule <95> is obligatory where applicable, as is generally the case with interpretive rules.

<92e> John was asked t [what [PRO to do t]]

(Chomsky 1980a:35)

By this analysis, (36) is blocked because, COMP being empty, <95a> is not triggered. Since *promise* is [+SC], and there is no subject in (36), the requirements for <95b> are not met either. Since the control rule does not apply to the control structure, (36) is assigned *.

(36) * Mary was promised e [PRO to leave]

The Chomsky (1980a) theory is criticized because it (a) assigns control incorrectly for (37) (Ruzicka 1980, Manzini 1983), and (b) would predict that (38) is well-formed.

(37) John asked Bill [PRO to be allowed [PRO to shave himself]]

(38) * Bill was asked to be allowed to shave himself.

In (37), the two PROs should be coindexed with the subject of *ask*, as in (39). However, <95b> would coindex PRO with the nearest controller--which is the object of *ask*, yielding (40).

(39) John_i asked Bill [PRO_i to be allowed [PRO_i to shave himself_i]]

(40) * John asked Bill_i [PRO_i to be allowed [PRO_i to shave himself_i]]

Likewise, <95b> applies to (38), and PRO is assigned the index of the nearest controller, yielding (41).

(41) * Bill was asked t [PRO_{arb} to be allowed [PRO to shave himself]]

Since the control rule has been triggered, (41) should be well-formed.

These are valid criticisms and, although Chomsky notes the control problem in double passive constructions, he suggests no solution for them:

There are some curious exceptions to this principle. Hust and Brame (1976) and Solan (1977) cite such double-passives as *John was promised to be allowed to leave*, which should be blocked but is not for reasons that are unclear. It seems that some syntactic property is involved, since substitution of a near synonym for the embedded passive changes the grammatical status to the expected *: e.g. **John was promised to get permission to leave*. (Chomsky 1980a:35 footnote 40)

Finally, Manzini (1983) criticizes Chomsky's theory on the basis that it would coindex *John* and PRO in (42).

(42) John asked [how [PRO to behave oneself]]

In (42), Manzini claims that PRO should have an arbitrary interpretation, because it can be the antecedent of *oneself*. I will discuss this criticism in "Manzini's Control Theory" on page 168.

5.3 WILLIAMS ON PREDICATION AND CONTROL

Williams (1980) argues that control is part of a more general theory of predication. He claims that there are two distinct control structures: obligatory control (OC) and nonobligatory control (NOC). Obligatory control occurs when an antecedent and a predicate containing PRO meet all five of the obligatory control criteria. All other cases fall under nonobligatory control. In cases of obligatory control, a predicate and its antecedent are assigned identical indices at the level of predicate structure. In cases of nonobligatory control, the node containing PRO is assigned an arbitrary index in the predicate structure. Under certain circumstances, this arbitrary index is rewritten in the logical form.

Williams introduces a new level to the grammar called predicate structure (PS), which falls between S-structure and logical form. In Williams' scheme, since a predicate requires a subject, the entire predicate is coindexed with an antecedent NP. Simple predicates are NP, AP, VP, and PP. Complex predicates are S and S'.⁹ At the level of predicate structure the only structural relationship defined is that of c-command restrictions. Therefore, the relationship between a predicate and its antecedent is one of c-command. Indices are assigned

⁹ Williams defines the complex predicates S and S' as [PRO VP]_S and [{PRO, WH} S]_{S'}, respectively. If an S' can take a lexical NP in its subject position, then it is not a predicate and therefore subject to nonobligatory control.

to OC at the PS level. All other instances of S containing PRO are assigned the index *arb* at PS. A set of two ordered *Arb* Rewriting rules, which apply disjunctively (the second being an optional rule), are responsible for rewriting indices on non-obligatorily controlled PRO in LF. If the *Arb* Rewriting rules do not apply, then PRO keeps its *arb* index and is interpreted generically.

The requirements for obligatory control (coindexation in PS) are quoted in (43):

(43)

- OC1. Lexical NP cannot appear in the position of PRO.
 - OC2. The antecedent precedes the controlled PRO.
 - OC3. The antecedent *c*-commands the controlled PRO.
 - OC4. The antecedent is thematically or grammatically uniquely determined.
 - OC5. There must be an antecedent.
- (Williams 1980:209)

If structures containing PRO do not meet all five of these conditions in PS, they are NOC, and PRO is assigned an *arb* index in PS.

The test for OC1 is whether or not the infinitive clause can take a *for*-complementizer which assigns case to the lexical NP in subject position. If the verb allows a *for*-complementizer, as illustrated in (44), then control is NOC.

(44) I would *hate* (for John) to leave.

The test for OC2 is whether the infinitive can appear in subject position, thereby preceding the antecedent. If the infinitive can appear as the matrix subject, preceding its antecedent, as in (45), then control is NOC.

(45) [PRO to leave] is nice

The test for OC3 is whether the antecedent can appear in a non-c-commanding position. If the antecedent can be in a non-c-commanding position, such as in a PP, then control is NOC. Williams defines c-command as follows:

A c-commands B iff every branching node which dominates A dominates B. (Williams 1980:204)

Williams gives (46) to illustrate a sentence which fails the test of OC3.

(46) I am counting [on Bill]_i [PRO to perjure himself]_i

Williams introduces thematic government and grammatical government in OC4. Thematic government occurs when the predicate is in the VP. The predicate is a lexically specified argument of the verb. If the predicate is in the VP, the controller is the verb's theme. For grammatical government, the predicate is a sister to the antecedent. The predicate can be the VP itself ([NP VP]), an adjunct ([NP VP PRED]), or a copular construction ([NP *be* PRED]). Williams illustrates thematic government with (47) and grammatical government with (48).

(47) John_i promised Bill [PRO to leave]_i

(48) John_i [died]_{VP} [PRO waiting for a bus]_i

The condition for OC4 is that the antecedent must be uniquely determined. If there is more than one possible thematic and/or grammatical governor, then control is NOC. Williams gives (50), where either the matrix subject or object can be the controller, as a violation of OC4.

(49) John told Mary that it would be important [PRO to leave early]

Williams writes that in (49) "either *John*, or *Mary*, or neither (in which case a generic interpretation is assigned) can be taken as the

controller" (Williams 1980:217). Williams also cites (50) as a violation of OC4.

Its [*count on*'s] not being OC is consistent with the fact that the antecedent of the complement of *count on* is not uniquely determined; in <48> either the matrix subject or object can be construed as the controller:

<48> I am counting on Bill to get there on time.

(Williams 1980:214):

Although Williams points out an ambiguity in <48>, he does not indicate that one of the two readings is an *in order to* reading. The reading shown in (50i) is that Bill gets somewhere on time. The reading shown in (50ii) is that Bill has to do something *in order for* me to get somewhere on time. Hence in (50i) the infinitive is a complement, but in (50ii) the infinitive is a sentential adjunct whose subject is always coreferential with the matrix subject.

- (50) i. I am counting [on Bill]_f [PRO to get there on time]_f
 ii. I_f am counting [on Bill] [PRO to get there on time]_f

The ambiguity, or non-unique determination of an antecedent, with a sentential *in order to* adjunct in (50) can be forced with what Williams' sets forth as a clear case of OC an verb, *persuade*.

- (51) John_f persuaded Bill (in order) [PRO to get there on time]_f

In (51), *John* persuades *Bill* of something so that *John* can get there on time. This example is parallel to the ambiguous reading Williams proposes for (50).

A fundamental problem with Williams' OC4 test is that whenever a transitive verb has an infinitival complement, there is always a possible *in order to* reading in which the subject of the infinitive is coreferential with the matrix subject. Thus it appears that Williams'

OC4 test needs to be revised to exclude ambiguous readings where *in order to* sentential adjuncts are involved. However, Williams argues that this kind of ambiguity shows that NOC properties cluster with *count on*. If OC4 is revised to omit the *in order to* reading, then the only NOC property exhibited by *count on* is that it has a non-c-commanding antecedent, as in (46). If OC4 remains sensitive to the *in order to* reading, then all transitive verbs which take an infinitival complement are NOC due to an OC4 violation involving *in order to* readings. Williams' theory requires, by definition, that OC controllers c-command PRO, since predicate structures are defined by-command. Therefore, since the antecedent for PRO in the complement of *count on* is a non-c-commanding NP, *count on* has to be a clear case of NOC. However, the only OC test *count on* appears to fail is the c-command condition of OC4.

It should be noted that a section of Williams (1980) is devoted to *purpose clauses*. What a *purpose clause* is is not defined. They appear to be *in order to* adjuncts which are in the VP. Williams cites Chomsky's (1980a) discussion of them, which I quote below. Purposive clause, a term used by Chomsky (1980a:42) refers to the infinitival clauses in these sentences:

- <116> a. John bought the dog [α to play with Mary]
 b. John bought the dog [α for Mary to play with]
 c. John bought the dog [α to play with]

Chomsky says, of the examples in <116>:

These structures resemble infinitival relatives but differ from them in some respects. For one thing, α is a constituent of VP rather than a relative clause with the head *the dog*. For another, there is no corresponding form such as *with which to play*, etc., for α of <116>. (Chomsky 1980a:41)

Chomsky continues, in a footnote:

For <116a> there is a possible structure with α a constituent of S rather than VP, so that the only possible controller for the missing subject is *John*, giving the interpretation *John bought the dog so that he could play with Mary*, more natural in such examples as *John bought the dog to please Mary*. We ignore this possibility here. ... (Chomsky 1980a:41).

Thus it appears that *in order to* sentential adjuncts are specifically not treated in Williams' discussion. However, Williams argues that since these VP adjuncts appear in predicate positions, they are both predicates. If this is the case, then sentential purpose clauses, which are also adjuncts that appear in predicate positions, are predicates.

According to OC5, if the matrix can appear without an NP subject to serve as an antecedent, then control is NOC. For example, the passive of *arrange* can have an expletive in the S-structure subject position, as in (52). Since *arrange* can take an antecedentless passive, it is NOC.

(52) It was arranged to have dinner at 6.

If a verb fails any of these five OC tests in any construction, then that verb is NOC.

The lexicon also plays a role in Williams' theory of predication. The lexicon (I) specifies syntactic and semantic subcategorization, including whether a complement is a predicate and (II) specifies whether the antecedent is the subject or object.¹⁰

¹⁰ Williams (1980), like Chomsky (1980a) and Koster (1984, 1987), assumes that, in clear cases of obligatory control, for example *promise* and *persuade*, the controller is stipulated in the lexicon. However, no proposal has been made for how control is stipulated in the lexicon. Since subjects do not appear in the lexicon (only external θ -role assignment does), there is the problem of how to specify subject control. Perhaps a default mechanism can be devised. For example, if object control is not specified, then the default is subject control. For now, I, like all those mentioned

In the next section, we will apply Williams' control theory to *ask*.

5.3.1 "ASK" AND THE THEORY OF PREDICATION

In Williams' Theory of Predication, *ask* is NOC. However, sentences (1) through (3) (repeated below) are NOC for different reasons. *Ask* in (1) is NOC because it fails the OC4 unique determination test, and *ask* in (2) and (3) is NOC because of its lexical features.

- (1) John asked Mary to put the book on the table.
- (2) John asked Mary how to put the book on the table.
- (3) John asked Mary where to put the book.

That *ask* in (2) and (3) are NOC is due to Williams' theory of how the distribution of verb-governed predicates is specified. Williams, who bases his theory of lexical structure on Grimshaw's work, includes in the lexical structure of verbs both the syntactic type and the semantic type of a complement. The semantic specification includes at least two types of complements: ___Q for direct and indirect questions and ___Pred for predicates. A complement can only be one semantic type. Since some verbs, such as *know*, subcategorize for both questions and predicates, Williams states that any instance of a complement must be *either* a predicate *or* a question, not a predicate *and* a question.

Of course, a given verb may specify that its complement is either A *or* B (*know*, for example, takes either a proposition *or* a question as its complement); what is forbidden is a specification which says that a particular complement must be simultaneously an A *and* a B. (Williams 1908:214).

above, will assume assignment of an obligatory controller in the lexicon without proposing how the specification is made.

It follows from this account that Williams must assign *ask* in (1) a different lexical structure from *ask* in (2) and (3). *Ask* in (1) does not subcategorize for a question. *Ask* in (2) and (3) does subcategorize for a question. Thus the two instances of *ask* must be treated separately in determining whether the verb is OC or NOC. Since a question is not a predicate and cannot fall under OC, (2) and (3) are NOC because the complements are of the semantic type ---Q .

Now it remains to be determined whether *ask* in (1) subcategorizes for a predicate. If it does subcategorize for a predicate, it is an OC verb. If it doesn't, it is an NOC verb. Williams does not offer a lexical specification for propositions which are not predicates. Since a predicate seems to be an OC proposition, then for purposes of reference let us call propositions which are not predicates NOC propositions or NOCprop. Following Williams, I will call OC propositions Pred.

A question remains unanswered: If there is a different semantic specification for OC and NOC propositions, what is the semantic difference between the two? The way in which control is assigned to the two seems to be the only difference that Williams' theory can offer.

If we translate the lexical structure of *ask* into Williams' scheme, *ask* subcategorizes for either a questioned or a propositional complement when it takes a direct object.

(53) *ask* ---NP , ---S' { ---Q , ---Prop }

The partial lexical structure for *ask* (53)¹¹ can be expanded out to questioned *ask* (54) and propositional *ask* (55).

¹¹ The lexical structure (53) isn't entirely satisfactory because it doesn't reflect that [-*wh*] complementizers are not allowed.

(54) [__ NP S', __ NP Q]

(55) [__ NP S', __ NP Prop]

The lexical structures (53) and (55) remain neutral as to whether the proposition is OC (a predicate) or NOC (not a predicate).

I will now go through Williams' OC tests for propositional *ask* to determine whether it is OC or NOC.

Propositional *ask* passes the no-lexical-subject condition of OC1. The inability of propositional *ask* to take a *for*-complementizer and subject NP is illustrated by (56).

(56) * John asked Mary (for) Bill to put the book on the table.

Propositional *ask* passes the linear precedence condition of OC2. The clause containing PRO cannot precede its antecedent, as is illustrated in (57).

(57) * To put the book on the table was asked by John.

The c-command condition of OC3 obtains for propositional *ask*, as is illustrated in (58).

(58) i. Mary_f was asked by John [PRO to put the book on the table
] _f
ii. * Mary was asked by John_f [PRO to put the book on the table
] _f

Propositional *ask* fails the unique determination condition of OC4 due to an ambiguity in (1). Like *count on*, propositional *ask* has an *in order to* reading. This ambiguity is clearer if we change the complement, as in (59).

(59) John asked Mary to get there on time.

(60) i. John asked Mary_f [PRO to get there on time] _f
ii. John_f asked Mary [PRO to get there on time] _f

In (60i), it is *Mary* who should get somewhere on time, while in (60ii), *John* asks *Mary* something *in order for John* to get somewhere on time. Therefore, according to Williams, the antecedent is not uniquely determined with propositional *ask*.

Another possible ambiguity with *ask* is the kind exhibited in (61).

(61) John asked Mary [PRO to go to the movies]

In (61), either *Mary* or *John and Mary* can go to the movies. But the *Arb* Rewriting Rules predict that the ambiguity should be that either *John* or *Mary* can go to the movies. There is no reading of (61) where *John* but not *Mary* goes to the movies except for a contrived *in order to* reading where *John* asks *Mary* something so that *John* can go to the movies.

In addition to the *in order to* ambiguous reading, propositional *ask* exhibits another kind of ambiguity. There are two possible S-structures for (62), which are shown in (63).

(62) John asked the teacher to leave the room.

- (63) i. John_f asked the teacher [PRO to leave the room]_f
 ii. John asked [the teacher]_f [PRO to leave the room]_f

In the reading (63i), John asks the teacher's permission to leave the room and in the reading (63ii), John requests that the teacher leave the room. This ambiguity is peculiar to *ask* and, as I have argued in "The Control Properties of "Ask"" on page 141, is due to a lexical structure where two different senses with different control properties have the same syntactic realization. It should be noted that this ambiguity does not hold for questioned *ask*. In neither of the sentences in (64) can the teacher be construed as being the one who goes anywhere.

- (64) i. John asked the teacher how to leave the room.
 ii. John asked the teacher where to go.

Finally, the obligatory antecedent condition of OC5 holds for propositional *ask*. It cannot take an antecedentless passive.

- (65) * It was asked to put the book on the table.

Thus *ask* in (1) is determined to be NOC because it fails OC4's unique determination condition, and *ask* in (2) and (3) is NOC because questions cannot be predicates.

Thus propositional *ask* is NOC because it fails just one OC test. This is an undesirable result for Williams' system. Williams observes that NOC verbs should fail several of the OC tests:

It is unsound to say that if members of X have properties P, then nonmembers of X will not have properties P, or will have the converses of properties P. But that to some extent is what we will do here. Specifically, the cases of NOC will not characteristically have the properties of OC (the properties of predication, as we have seen). Thus, if any property of OC is lacking for a particular case of control, we should expect the other properties of OC to be lacking for that case as well. (Williams 1980:212)

Propositional *ask* does exactly what Williams claims should not happen. It is a verb which fails only one OC test, it does not fail a cluster of the tests. Such a result suggests that the distinction between OC and NOC is an arbitrary distinction which is neither particularly useful nor significant.

Williams' theory of predication does allow, in theory, there to be a difference in how control is assigned when a verb takes or doesn't take a *wh*- complement. In the case of *ask*, however, *arb* gets assigned to both types of complement in PS, because the semantic types of its

complements are ___Q and ___NOCprop. Both instances of *ask* are then subject to the *Arb* Rewriting Rules.

It remains to be seen whether there is any verb which requires the distinction between OC and NOC depending upon whether its complement is of the semantic type ___Pred or ___Q. A theory that allows this possibility is weakened if the need for a distinction which it allows is never manifested. If the distinction is not required, then the theory is too rich, permitting classification of items which do not exist.

Williams proposes two *arb* Rewriting Rules, quoted below as (66) and (67):

(66) *Arb Rewriting Rule I:*

Rewrite *arb* as coindexed with an NP which commands it, and which it commands. (Williams 1980:216)

(67) *Arb Rewriting Rule II:*

If an NP commands *arb*, coindex NP and *arb* (Williams 1980:217)

Williams (1980) does not define *command* explicitly. However, in his discussion of the rewriting rules he distinguishes *command* from *c-command* as follows:

Arb Rewriting is not subject to *c-command*, *command* stated in terms of NP, or A-over-A. (Williams 1980:217)

Judging from his examples, a node commands any other node in the minimal sentence that contains it. The relevant nodes are NP antecedents and nodes which have received an *arb* index in PS (NP, VP, PP, S, and S'),¹² which I will refer to as X. If X and NP are both dominated by the same S', and there is no S' which dominates one but not the other,

¹² AP's cannot have an *arb* index because "AP is the only category obligatorily interpreted as a predicate." (Williams 1980:206)

then they command each other. This configuration triggers *Arb* Rewriting Rule I (66). Williams' examples of sentences where indices get rewritten by *arb* Rewriting Rule I are given in (68), where "*arb* is rewritten as a coindex with the italicized NP" (Williams 1980:217).

- (68) i. I want [PRO to leave]*arb*
 ii. [PRO to leave]*arb* would be *my* pleasure.
 iii. it is important to *me* [PRO to leave]*arb*

If X is dominated by an S' which does not dominate the NP, but an S' dominates both the NP and the S' containing X, and every S' dominating NP also dominates X, then the NP commands X but X does not command the NP. This configuration triggers *arb* Rewriting Rule II (67). Williams' example of a sentence where the index is rewritten by *arb* Rewriting Rule II is given in (69).

- (69) John told Mary that it would be important [PRO to leave early]*arb*

Arb Rewriting I applies first, and "is more or less obligatory". If two antecedents meet the structural description of (66), then either of them can be the controller, and the sentence is ambiguous. Williams gives (70) as an example.

- (70) John counted on *Bill* [PRO to get there on time]*arb*

When one of the structurally defined possible antecedents cannot be construed as the controller, as in Williams' example (71), Williams says that "pragmatic or semantic factors choose one controller over another."

- (71) i. I bought a book [PRO to read]*arb*
 ii. I bought *Bill* a book [PRO to read]*arb*

Arb Rewriting Rule II is an optional rule. Thus (69) has three possible readings: PRO is indexed with *John*, PRO is indexed with *Mary*, or *arb*

Rewriting Rule II does not trigger, and PRO keeps its *arb* index. If neither rule is triggered, then there is no controller and the *arb* index remains on the predicate.

In examples (1), (2), and (3), *John*, *Mary*, and the clausal complement indexed *arb* all command each other. Thus they all trigger *arb* Rewriting Rule I. The input of (1), (2) and (3) to *arb* Rewriting Rule I would be (72), (73) and (74), respectively.

(72) *John_i* asked *Mary_j* [[*s*PRO to put the book on the table]*arb*]

(73) *John_i* asked *Mary_j* [*s*'how [*s*PRO to put the book on the table]*arb*]

(74) *John_i* asked *Mary_j* [*s*' where_{*k*} [*s*PRO to put the book *e_k*]*arb*]

In all three cases, *Arb* Rewriting Rule I assigns ambiguous readings to the sentences. Williams claims that "in some cases, pragmatic or semantic factors choose one controller over another" (Williams 1980:217), but he does not indicate how this would work. He offers (75) as an example of a case where control is influenced by pragmatic or semantic factors:

- (75) i. I bought a book [PRO to read]*arb*
 ii. I bought Bill a book [PRO to read]*arb*

I will return to this problem upon discussing the interpretation of PRO in adjunct clauses.

Example (1) does have an ambiguous *in order to* reading. However, examples (2) and (3) have no ambiguous reading. Furthermore, it is difficult to imagine what pragmatic or semantic factors would influence the choice of antecedent in (2) and (3). If in (75) PRO is assigned reference by pragmatic or semantic factors, which I doubt, one can imagine what they would be. People can read and inanimate things cannot

read. But what pragmatic or semantic factor would choose between John and Mary in (2) and (3)?

The *arb* Rewriting Rules incorrectly assign antecedents in (2) and (3). When (1) is passivized, *arb* Rewriting Rule I correctly assigns control to *Mary*, as shown in (76). However, the rewriting rule incorrectly assigns control to *Mary* in (2) and (3), as shown in (77) and (78).

(76) *Mary*_i was asked [PRO to put the book on the table]_i

(77) * *Mary*_i was asked [how [PRO to put the book on the table]]_i

(78) * *Mary*_i was asked [where_j [PRO to put the book e_j]]_i

The *arb* Rewriting Rules assume that if there is an NP in the adjacent matrix clause, then that NP is the controller. When (2) and (3) are in an agentless passive construction, the NP in the matrix clause is not the controller. In (2) and (3), PRO should not receive an anaphoric index. However, (2) and (3) match the structural description of the obligatory *arb* Rewriting Rule I, and are thereby coindexed with the surface subject.

To summarize, Williams' theory of predication cannot correctly assign a controller to PRO in the complement of *ask*. When the sentential complement for *ask* is [+wh], *ask* is unambiguously controlled by the subject. By Williams' theory of predication, the antecedent of PRO is ambiguous, and the preferred controller is decided by unspecified pragmatic or semantic factors. In addition, when *ask* is passivized, Williams' *arb* Rewriting Rules will assign control incorrectly.

5.4 MANZINI'S CONTROL THEORY

Manzini (1983) wants to reduce control theory to binding theory.

If we are correct, control is an essentially configurational phenomenon and control theory an essentially configurational theory; in particular, control theory is constructed on essentially the same notions on which binding theory is constructed in Chomsky (1981). (Manzini 1983:421)

If control is a purely configurational phenomenon, then control cannot be subject to lexical exceptions. Therefore Manzini rejects the possibility that the lexicon specify subject or object control. Instead she claims that, in a control environment, either subject or object control is syntactically correct and that the choice of one over the other is not in the domain of control. That is, according to Manzini, the goal of control theory is to specify *if* PRO has a controller but not which constituent that controller is. The theory predicts whether PRO must be bound or can be free--not which possible controller is *the* controller. She holds that all of the following examples are syntactically well-formed when PRO is controlled from either the subject or the object position of the matrix clause:

- <12> John asked Bill [PRO to shave himself]
- <13> John asked Bill [PRO to be allowed to shave himself]
- <14> John promised Bill [PRO to shave himself]
- <15> John promised Bill [PRO to be allowed to shave himself]
- ...
- <21> Bill was asked [PRO to shave himself]
- <22> * Bill was asked [PRO to be allowed to shave himself]
- <23> * Bill was promised [PRO to shave himself]
- <24> Bill was promised [PRO to be allowed to shave himself]

... it appears quite possible to assume that on syntactic grounds all of <12>-<15> and <21>-<24> are indeed well formed with both object and subject control; and that either subject or object control is excluded for other reasons--say, on semantic grounds. (Manzini 1983:423)

What Manzini wants to determine configurationally is whether PRO must be bound (or "cosuperscripted") or whether it can be free (or "(co)refer freely").

Manzini's test for whether PRO must be bound is when PRO can be the antecedent of *himself* or *herself* but not of *oneself*. Her test for (co)referring freely is whether PRO can be the antecedent of *oneself* (as well as of *himself* or *herself*). I will call this diagnostic for distinguishing between PRO_{anaphor} and PRO_{arb} the reflexive test.

Manzini claims that control facts are correctly described by the following two generalizations:

- (79) PRO in an object sentence of a sentence S is bound in S.
- (80) PRO in a subject sentence (co)refers freely.

According to (79), PRO in object infinitives is bound. In the counterexample <25>, where PRO is not bound by the matrix subject, Manzini assumes the presence of a null object which acts as the binder of PRO:

<25> John said [PRO to behave oneself]

...it appears quite possible to assume that in <25> *say* has a phonologically null indirect object and that the PRO is bound by it. (Manzini 1983:423-424)

According to (80), PRO in subject sentences, including extraposed sentences, (co)refers freely.

Manzini's claim is that these and other control facts follow from (81) and (82).

- (81) PRO is an anaphor without a governing category.
- (82) An anaphor without a governing category (PRO) is bound in its domain-governing category.

Thus, if PRO has a domain-governing category, it is bound. If PRO has no domain-governing category, it can (co)refer freely. I will not go into the details of Manzini's binding theory or definition of *domain-governing category*. Instead let us consider her predictions and discuss whether these predictions are correct.

According to Manzini, it follows from (81) and (82) that if COMP is filled, either with a trace or with an overt complementizer, PRO has no domain-governing category and therefore (co)refers freely. She offers <35> through <37> to illustrate how PRO does in fact (co)refer freely when COMP is filled.

<35> John asked [how [PRO to behave himself]]

<36> John asked [how [PRO to behave oneself]]

<37> John was asked [how [PRO to behave oneself]]

However, as she points out, <39> and <40> are counterexamples to this claim. PRO in these examples does not (co)refer freely, even though COMP appears to be filled with traces. To get around these counterexamples, Manzini assumes that some kind of adjunction has occurred, and that the traces are not in a COMP node. She does not, however, offer any further clues as to what the structures in <39> and <40> could be.

Notice that under the present approach, if the higher *t* in sentences like <39>-<40> is in Comp, ... the PRO is predicted to (co)refer freely; if on the other hand the higher *t* is adjoined to *S/S'*, ... the PRO is predicted to be bound in the matrix sentence:

<39> What did John ask Bill [*t* [PRO to do *t*]]

<40> What was Bill asked [*t* [PRO to do *t*]]

Since in both <39> and <40> PRO must be coreferential with *Bill*, we are led to the conclusion that the higher *t* in <39>-<40> is not in Comp, but adjoined somewhere.

Finally, in the case of modification, PRO has a domain-governing category and is predicted to be bound. When the infinitive is in the VP, then PRO is bound by either subject or object.

... in the case in which the modifier control sentence is attached to VP, the prediction is that PRO is coreferential with either an object or the subject of S; this can be the case, for instance, with the purpose sentence <52>:

<52> John hired Mary [PRO to fire Bill]

When the infinitive is attached to the S, PRO is bound by the subject.

In the case in which the modifier control sentence is attached to S, on the other hand, since only the subject c-commands (hence, can bind) the PRO in S, the prediction is that the PRO is coreferential with the subject of S; this must be the case, for instance, with the purpose sentence <53>:

<53> John fired Mary [PRO to hire Bill]

Manzini cites the passive constructions in <54> and <55> as counterexamples to her predictions for modification. In these sentences, PRO cannot be bound by the matrix subject, although her theory predicts that PRO in this configuration must be bound in the matrix clause. In these cases, Manzini again assumes that PRO is bound by a null agent.

The case in which PRO has arbitrary reference, as in <54> and possibly <55>, can be assumed to be the result of control by a phonologically null agent:

<54> Mary was fired [PRO to hire Bill]

<55> Mary was hired [PRO to fire Bill]

(Manzini 1983:428)

Manzini offers the following argument, based on the reflexive test for arbitrary interpretation, against Chomsky's (1980a) position:

A second problem with [the Chomsky (1980a) control theory] is the existence, next to examples like <37>, of examples like

<36>, where V has a controller but arbitrary reference is possible; an example analogous to <36> is in fact pointed out in Chomsky (1980a) outside the discussion of control:

<36> John asked how to behave oneself.

<37> John was asked how to behave oneself.

This suggests that the distinction between the case in which Comp \neq null and V has no controller and the case in which Comp = null or V has a controller has to be given up, as in the present theory, in favor of a distinction between the case in which Comp \neq null and the case in which Comp = null. (Manzini 1983:429)

These examples (<36> and <37>) are crucial in two respects: First, they indicate that the reflexive test for arbitrary interpretation, i.e. that the PRO can be the antecedent of *oneself*, is an inadequate diagnostic for arbitrary reference. The indefinite reflexive *oneself* in <36> is understood to include *John*, but *oneself* in <37> does not include *John*. Notice that (83) is understood by most people to mean that *John behaves*, and (84) is understood to mean that someone, not *John*, *behaves*. Also, (85), but not (86) is well-formed with *himself* understood as *John*.

(83) John asked [how [PRO to behave at a black-tie dinner]]

(84) John was asked [how [PRO to behave at a black-tie dinner]]

(85) John asked [how [PRO to behave himself]]

(86) * John was asked [how [PRO to behave himself]]

Second, if *oneself* includes *John* in <36> but not <37>, this would argue for Chomsky's position about the lexical structure of *ask*, repeated below:

But certain verbs (e.g. *promise* or *ask* with an indirect question complement...) are marked in the lexicon with the feature [+SC] indicating "assigns subject control". (Chomsky 1980a:33)

Since *ask* in <36> has only an S complement, PRO is necessarily controlled by the subject. When PRO is the antecedent of the indefinite reflexive *oneself*, one would expect the subject of *ask* to be included as a possible antecedent. Since <37> is specified for subject control in the lexicon, and the underlying subject does not appear in this passive construction, then one would expect *oneself* and the surface subject be disjoint in reference, as in <37>.

5.4.1 "ASK" IN MANZINI'S CONTROL THEORY

Manzini's theory predicts that PRO in (1) is bound and that PRO in (2) and (3) can (co)refer freely.

- (1) John asked Mary to put the book on the table.
- (2) John asked Mary how to put the book on the table.
- (3) John asked Mary where to put the book.

In the case of (1) PRO is in a domain-governing category and therefore must be bound. Manzini's theory makes no claims about whether *John* or *Mary* is the antecedent. The claim is that syntactically either NP can be the antecedent for PRO. If one is excluded from antecedenthood, this is due to "semantic grounds."

As for (2) and (3), since COMP is filled, PRO does not have a domain-governing category, and PRO (co)refers freely. If PRO (co)refers freely, then either the subject or the object or neither can be the antecedent for PRO. Manzini's theory predicts, then, that (87) through (89) are all syntactically well-formed.

- (87) John asked Mary how to behave himself.
- (88) * John asked Mary how to behave herself.
- (89) ? John asked Mary how to behave oneself.

Presumably (88) would have to be ruled out on semantic grounds.

However, when *ask* has an indirect question with *what* in COMP, PRO must be bound in the matrix sentence. This property of *ask* is exhibited in a version of Manzini's own example <39>. In (90), *what* has not moved up to the matrix COMP, as it has in <39>.

<39> What did John ask Bill [*t* [PRO to do *t*]]

(90) John asked Bill [what [PRO to do *e*]]

Manzini's theory predicts that PRO in (90) (co)refers freely because the COMP node is filled. However, PRO must be coreferential with *John* in (90). There is no arbitrary interpretation for (90), as is shown by the ill-formedness of (91).

(91) * John asked Mary what to do with oneself.

Manzini cannot argue, as she does with <39> and <40>, that (90) is an adjunction structure. The lower clause in (90) is the indirect question complement of *ask*.

Manzini's theory cannot account for this switch in control exhibited by transitive *ask* depending upon whether its complement is an indirect question. A basic claim of Manzini's theory is that the existence of control can be predicted solely from the structural configuration of the sentence. However, the facts for *ask* indicate that neither the existence of control nor the antecedent of PRO can be predicted on a purely configurational basis.

5.5 KOSTER ON GOVERNED PRO AND CONTROL

Koster, like Williams, has two processes which assign an antecedent to PRO. Whereas Williams proposes that the two processes occur at separate

levels, Koster proposes that coreference is assigned by two separate components of the grammar -- those of Binding and Control.

In defending his thesis of *radical autonomy*,¹³ Koster (1984, 1987) argues that the overlap between control of obligatory PRO and anaphor binding (binding principle A) indicates that obligatory PRO is, in fact, an anaphor. Koster assumes that control facts for obligatory PRO are captured if obligatory PRO is treated as an anaphor that must be bound in its governing category. Since the key relation for binding anaphora is government, and obligatory PRO is an anaphor, then obligatory PRO, which is subject to binding principle A, is governed.

In distinguishing between obligatory and nonobligatory control, Koster refers to instances of PRO as obligatory PRO and optional PRO. He offers a new test for differentiating between the two types of PRO. Obligatory PRO, which occurs in obligatory control configurations, appears in complements of verbs that *do not* take the complementizer *for* or a gerund. Optional PRO, which occurs in nonobligatory control configurations, appears in complements of verbs that *do* take the complementizer *for* or a gerund. The environment of obligatory PRO is attributed to a lexical property of the matrix verb. The verb selects a reduced S' which contains no COMP node. This is an alternative account for S'-deletion.¹⁴ The result of S'-deletion is that the matrix verb governs the subject of the complement clause. Similarly, when a verb

¹³ *Radical autonomy* is "the idea that locality principles for movement and anaphora are of the same type" (Koster 1984:417).

¹⁴ Koster claims that when PRO is obligatory, instead of a mechanism of S'-deletion, the matrix verb subcategorizes for a reduced clause, an S' which lacks COMP: "I believe that the phenomenon referred to as S'-deletion *is* in fact the absence of Comp" (Koster 1984:427).

selects a reduced clause, the embedded subject is governed by the matrix verb. By this mechanism, obligatory PRO is governed. Government of optional PRO is blocked by the presence of the COMP node.

Koster defines the properties that are shared by binding and obligatory control in (92), "with instances of γ as arguments and unique α 's as arguments":

(92)

- a. *Obligatoriness*: α and γ are obligatorily linked.
 - b. *Uniqueness of α* : There is only one α for each γ .
 - c. *Prominence*: α c-commands γ .
 - d. *Locality*: α and γ are in the same domain β .
- (Koster 1984:418)

Koster observes that the antecedent-PRO relations which deviate from (92) occur only with the environment of optional PRO. Therefore, he argues, control theory is required only for ungoverned (optional) PRO. When PRO is governed (obligatory), it obeys the constraints of (92), as do antecedent-trace relations. When PRO is ungoverned, it deviates from (92). The behavior of ungoverned PRO is exemplified in (93) (these examples are from Koster 1984:426):

- (93) i. It is impossible [PRO to help Bill]
 ii. John proposed to Mary [PRO to go to the movies]
 iii. It is difficult for Mary [PRO to help Bill]
 iv. John thinks [_S it is impossible [_S PRO to shave himself]]

Ungoverned PRO in (93i) has no antecedent, ungoverned PRO in (93ii) has a split antecedent, ungoverned PRO in (93iii) has a non-c-commanding antecedent, and the antecedent of ungoverned PRO in (93iv) is non-local.

Once Koster has limited the domain of control theory to that of ungoverned PRO, he defines the properties of control. He claims that control has two properties: a locality condition and a lexical specification. The locality condition requires the controller to be in the

adjacent matrix clause. The lexical specification is that subject or object control is stipulated in the lexical structures of control verbs:

The controller for an embedded subject (PRO) is a *designated* argument of the *minimal* argument structure containing the control complement. (Koster 1984:430)

Koster introduces *implicit controllers* to explain how instances of long distance control, as in (93iv), and PRO_{arb}, as in (93i), do not violate the locality condition. He claims that control theory does not distinguish between implicit and explicit arguments, and that PRO can be controlled by an implicit argument. Koster illustrates his argument with the adjective *difficult*. He claims that *difficult* has an obligatory PP object; that it is difficult *for someone* to do something. When the object of *difficult* is explicit, this object controls PRO, as in (94). When the object of *difficult* is implicit, it too controls PRO, as in (95). In (95), the subject of a non-adjacent clause, *Mary* "can be interpreted as the controller *because Mary* can be interpreted as the one *for whom* the particular action is difficult" (Koster 1984:429).

(94) Mary said it was difficult for Bill to take another topic.

(95) Mary said it was difficult to take another topic.

(96) It is difficult to take another topic.

Similarly, in (93i) and (93iv), Koster would claim that ungoverned PRO is controlled by an implicit object in the adjacent matrix clause. Binding theory, unlike Control, "requires an explicit, c-commanding antecedent" (Koster 1984:433).

If Koster's analysis is correct, then ungoverned PRO in (93i), (93ii) and (93iv) is controlled by a non-c-commanding antecedent. According to Koster, c-command is not relevant to control. He claims that

control theory requires only the locality condition and lexical designation.

The relevant argument qualifies as a controller, no matter how it is structurally expressed. In general, this means that the controlling arguments can be left implicit, or can be couched in a characteristic PP. (Koster 1984:429)

Thus, Koster resolves the locality problem by assuming the presence of implicit arguments, represented as empty categories, in the phrase marker.

5.5.1 "ASK" AND THE THEORY OF GOVERNED PRO

According to Koster, when a verb subcategorizes for a full S' (an S' with a COMP node) PRO is optional, or ungoverned, and subject to control theory. When a verb subcategorizes for a reduced S' (an S' without a COMP node) PRO is obligatory, or governed, and subject to binding theory.

Sentences (1) through (3), repeated below, divide with respect to whether they are subject to binding or control theories.

- (1) John asked Mary to put the book on the table.
- (2) John asked Mary how to put the book on the table.
- (3) John asked Mary where to put the book.

Ask does not take a *for*-complementizer in (97), and therefore governs PRO.

- (97) * John asked Mary [for Bill to go]

Since, by Koster's test, PRO in the complement of *ask* is governed, it is an anaphor, and subject to Binding Principle A.

In sentences (2) and (3), the complementizer position is filled with *wh*- elements. Koster (1984) does not discuss indirect questions, but one can only assume that since the complementizer position is

lexically filled in (2) and (3), *ask*, in these cases, subcategorizes for a full S'. Like *wonder*, when *ask* subcategorizes for a full clausal complement, the complement is a *wh*- clause. When *ask* subcategorizes for a question, it does not govern PRO, and ungoverned PRO is subject to Control theory.

Koster's lexicon would require at least the following subcategorizations for *ask*:

- (98) i. ___ NP S'(-Comp)
 ii. ___ NP S'(+*wh*) Control=subject

(98i) is subject to Binding, (98ii) is subject to Control.

Given this information in the lexicon, the *ask* facts fall out nicely. Since *Mary* is the nearest c-commanding NP in the governing category of PRO, binding principle A will assign *Mary* as the antecedent in (1). Control theory, in conjunction with the lexical specification of (98ii) will assign *John* as antecedent in (2) and (3).

The facts for *ask* also follow from his assumptions and lexical entries when there is no direct object NP. According to Koster's theory, the direct object in the sentences in (99) is represented in the phrase markers as implicit arguments: someone asks *someone else* (to do) something. In (99), PRO is governed by *ask*, therefore the antecedent must be an explicit c-commanding NP in the governing category of PRO. *John* fits this structural description for binding principle A, and *John* is the antecedent in (99). In (100i) and (100ii), the implicit argument is not the controller because the lexicon specifies subject control when the complement is a *wh*- clause. Again, *John* is the controller.

(99) John asked to go to the movies.

- (100) i. John asked how to tie a knot.
 ii. John asked where to go.

A problem with Koster's theory is that although PRO in (2) and (3) is ungoverned, these sentences do not fall into the pattern that Koster's theory predicts in control cases. Koster predicts that only ungoverned PRO can (I) be antecedentless, (II) allow a split antecedent, (III) have a non-c-commanding antecedent, and (IV) have an implicit antecedent. When *ask* subcategorizes for a *wh*- complement clause, and PRO is ungoverned, it has none of the four properties just listed. It requires an antecedent, as illustrated in (101).

- (101) i. * It was asked how to tie a knot.
 ii. * How to tie a knot was asked.
 iii. * It was asked where to go.
 iv. * Where to go was asked.

When PRO is not governed by *ask*, it does not allow an interpretation with a split antecedent. In (102), the interpretation is that *John* goes somewhere. There is no interpretation in which *Mary* goes anywhere.

- (102) i. John asked Mary how to go to the movies.
 ii. John asked Mary where to go.

Contrary to Koster's claims however, inasmuch as any verb allows a split antecedent, *ask* allows a split antecedent when it *does* govern PRO, illustrated in (103), where the interpretation can be either *Mary* or *John and Mary*.

- (103) John asked Mary to go to the movies.

Ungoverned PRO in the complement of *ask* requires a c-commanding antecedent as illustrated by (104).

- (104) i. * How to tie a knot was asked by John.
 ii. * Where to go was asked by John.

It also requires an explicit antecedent, thereby not allowing apparent nonlocal antecedents, as in (105).

- (105) i. * John thinks it was asked how to tie a knot.
 ii. * John thinks it was asked where to go.

Although the facts for assigning an antecedent to the clausal complement of *ask* follow from Koster's proposal, it is an anomaly for his diagnostics that when *ask* governs PRO, PRO exhibits some of the behavior of ungoverned PRO. In (103), governed PRO can take a split antecedent. In addition, when *ask* does not govern PRO, the ungoverned PRO exhibits none of the behavior which Koster's theory predicts.

Were one to argue that *ask* always subcategorizes for a full S', and that there is a null complementizer in (1), matters would only become worse. In this case, (1) would be ungoverned and subject to control. The lexicon would have to specify object control when the clause is [-wh], as in (1). Since Koster's theory predicts that the phrase marker for (99) contains an implicit object NP, and if (99) is subject to control theory, then the controller in (99) would be the implicit direct object, yielding (106), where PRO has an arbitrary interpretation.

- (106) * John asked e_i [PRO_{*i*} to leave]

If one argued that (99) is derived from a different lexical structure for *ask*, an intransitive subcategorization, then (99) would be ambiguous between (106) and (107), for the transitive reading with an implicit argument would still be possible.

- (107) John_{*i*} asked [PRO_{*i*} to leave]

Koster's theory of governed PRO correctly predicts the antecedent for PRO in the complement of *ask*. However, the theory also predicts

that the complement clause should follow certain patterns of behavior, one peculiar to governed clauses and the other peculiar to ungoverned clauses. Neither the governed nor ungoverned complements of *ask* follow their predicted pattern. This suggests that Koster's diagnostics do not make the correct distinctions in the data.

5.6 HORNSTEIN AND LIGHTFOOT'S THEORY OF PRO

Hornstein and Lightfoot (1987), hereafter H&L, appear to adopt a system similar to that of Williams for assigning control except, like Koster, H&L argue that the crucial relationship is government, not predication. They claim PRO can be governed and that control is determined based solely upon whether or not PRO is governed. Thus, they argue, there is no need to introduce an extra level of predication to the grammar.

In adopting [these] analyses¹⁵ ... and in treating PRO as an anaphor when governed and as a pronominal when ungoverned, we derive the basic properties of W's 'predication' structures.
(H&L:36)

The basic system of H&L is that in some configurations PRO is an anaphor and in some configurations PRO is pronominal. Government is the relation by which it is determined whether PRO is anaphoric or pronominal. If PRO is governed, it is an anaphor and subject to Binding Principle A. If PRO is ungoverned it is pronominal and "ungoverned PRO receives the arbitrary interpretation" (H&L:26).

H&L determine whether a complement clause containing PRO is governed or ungoverned based upon whether the COMP node is lexically filled. When COMP is lexically filled, then the clause has S' as its maximal

¹⁵ A major part of the H&L analysis, which I will not discuss, is that small clauses have subject PROs.

projection, and PRO is ungoverned. In a footnote, H&L also point out that if COMP is filled by an empty operator, then PRO is ungoverned.

Similarly, adjectives may delete S', as in (a) below; therefore the subject of an infinitival may be governed, and thus anaphoric (in a-b). But if the S' is not deleted, as in (c), PRO will not be anaphoric:

- (a) John_i is likely [*t_i* to lose]_S.
- (b) John is too stubborn [PRO to visit New York]_S
- (c) John_i is too stubborn [(WH_i) PRO to talk to *t_i*]_{S'}.

(H&L:43 footnote 20)

Thus, if there is *something* (either a lexical or an empty category) in COMP, then PRO is ungoverned. If COMP is empty, then S'-deletion occurs, and PRO is governed.

An example of governed PRO is given in (108): PRO is in a subordinate clause and the minimal clause containing PRO has S (not S') as its maximal projection. When governed, PRO is in an obligatory control structure, and *is* an anaphor.

(108) John tried [PRO to leave]_S]_{vp}

When PRO is ungoverned, it is in a non-obligatory control structure, and is pronominal. Examples of ungoverned PRO, and therefore non-obligatory control structures, are given in (109)-(111). Since S' blocks government, PRO is ungoverned when the clause containing PRO which has S' as its maximal projection. The three "major contexts" under which PRO is ungoverned (i.e. has S' as its maximal projection) follow: The first is when COMP is lexically filled, as in (109). The second is when the clause containing PRO "seems to act as a subject" as in (110). The third is when the clause containing PRO is in an extraposed position, as in (111).

(109) It is unclear [how [PRO to behave oneself in public]]

(110) (John knows that) [PRO to behave oneself in public] is a social requirement.

(111) It is a social requirement [PRO to behave oneself in public]

Non-obligatory control, or pronominal, PRO may be subject to further analysis: Although ungoverned PRO "does not require an antecedent of any kind," if another NP is available for coindexing, PRO "may or may not" be coindexed with it. If no NP is available for coindexing, PRO "must be" understood as arbitrary in reference:

An ungoverned PRO is not an anaphor, and therefore does not require an antecedent of any kind. If no other NP is available for coindexing, then PRO must be understood as arbitrary in reference, as in ...[(109)-(111)]. If another NP is available, PRO may or may not be coindexed with it... (H&L:29).

However, H&L do not specify what mechanisms or structural conditions determine which NP is available for coindexing, nor what properties make coindexing possible.

How non-obligatory control PRO is to be interpreted is unclear ... For a good description of the interpretation of non-obligatory PRO, see Williams 1980. His 'Arb' Rewriting rules seem to us pretty accurate. However, these rules describe rather than explain the relevant cases, as he notes. (H&L:42 footnote 19)

Although H&L agree that Williams' *arb*-rewriting rules are descriptively accurate, H&L are unequivocal in their claim that PRO is a pronominal and, like a pronoun, does not require an antecedent. If this is the case, coindexing an ungoverned PRO is always optional. However, Williams' *arb*-rewriting rule I is obligatory, requiring some instances of non-obligatory control PRO to have an antecedent in LF. Thus, in the predication system of Williams, some instances of PRO which are free in predicate structure cannot be free in LF. It is not at all obvious how H&L can incorporate these facts into their system.

H&L qualify what is meant by an arbitrary interpretation, adopting the reflexive test for arbitrary interpretation used by Manzini. Upon referring to the arbitrary reading in examples (109)-(111), they add:

"Arbitrary" is in fact somewhat misleading; W's "non-obligatory control" is more accurate. The relevant distinction is the following. A governed PRO is an anaphor, and thus must have a c-commanding antecedent within its governing category. In [*John tried/wanted*[PRO to leave]S]S, therefore, PRO must be coindexed with a c-commanding antecedent in the matrix clause; hence the ungrammaticality of *It was tried/wanted (by John) to leave*. An ungoverned PRO is not an anaphor, and therefore does not require an antecedent of any kind. If no other NP is available for coindexing, then PRO must be understood as arbitrary in reference ... If another NP is available, PRO may or may not be coindexed with it; thus *It is unclear to John how to behave* may be continued with *himself* or *oneself*. Adding *himself* entails that PRO has definite reference, hence is coindexed with *John*; but adding *oneself* entails that PRO has indefinite reference, i.e. the arbitrary interpretation. In *Mary knows that PRO to behave would help John*, similarly, PRO may be coindexed with *Mary* or *John* or nothing; so after *behave* one may insert *himself*, *herself*, or *oneself*. (H&L 1987:29-30 footnote 5)

Apparently "arbitrary reference" means that PRO can either be the antecedent to a definite reflexive pronoun, in which case it is obligatorily coindexed with a c-commanding NP in the matrix clause, or be the antecedent to the indefinite reflexive pronoun, in which case PRO obligatorily has indefinite reference. Like Manzini, H&L do not take into account the further fact of whether a c-commanding NP is within or disjoint from the scope of the indefinite pronoun *oneself*.

Just as the mechanisms for coindexation are left unspecified for ungoverned PRO, H&L do not specify how control is assigned in obligatory control environments. They agree with Williams that c-command is crucial, and that only NPs can be controllers, but leave open the question of how the controller is chosen when there are two competing c-commanding NPs:

In each case, PRO is governed in the matrix clause ..., and therefore must be coindexed with a c-commanding NP in its governing category ... Note that, in 7b, both *John* and *the meat* c-command PRO; so a tighter restriction is needed. This problem is common to many theories, and several ideas have been put forward; thus Rosenbaum's (1967) Minimal Distance Principle, which would make *the meat* the only possible (or at least the "unmarked") antecedent. Alternatively, one might claim that PRO must normally c-command its antecedent--i.e., that there be no maximal node dominating PRO which does not also dominate its antecedent, where S' but not S is maximal; or one might require a PRO in a structure like 7b to be normally coindexed with a theme. The latter two ideas are adaptations of restrictions that Williams 1980 imposes on his predication rule. Under our approach, whatever device makes *John* an impossible or marked antecedent for PRO in 7b will also make *John* an impossible antecedent for PRO in *John persuaded Bill [PRO to leave]*.

<7> b. John INFL [[ate the meat]v' [PRO INFL₀ [raw]AP]S]VP.

(H&L 1987:27)

H&L explicitly state that governed PRO is an anaphor and also say that governed PRO acts like an anaphor. However, they never state that ungoverned PRO is a pronoun. Rather, they state that ungoverned PRO "is pronominal" or "acts like a pronoun". The status of governed PRO is clearly defined by H&L, but the status of ungoverned PRO is often qualified. There are some passages, however, which seem unequivocal, for instance:

Since a governed PRO is anaphoric, we can now generalize and say that all anaphors require overt antecedents. Pronouns, of course, do not require overt antecedents ...; nor does an ungoverned PRO ..., which is not an anaphor. ... In adopting analyses like 7, and in treating PRO as an anaphor when governed and as a pronominal when ungoverned, we derive the basic properties of W's 'predication' structures. (H&L:36)

We have developed an alternative treatment which avoids the problems identified, involves a new analysis of small clauses, and assumes that PRO may be governed. When governed, a PRO is anaphoric; when ungoverned, it is pronominal. (H&L:50)

H&L say nothing about not adopting the typology of empty categories, presented in Chomsky (1982) and elsewhere, into the features [anaphor] and [pronominal]. Therefore, I assume that they adopt these as the features which drive Binding Theory. If ungoverned PRO is not an anaphor, then it has the feature [-anaphor]. If ungoverned PRO is, as H&L claim, pronominal, then it has the feature [+pronominal]. Thus ungoverned PRO has the features: [+pronominal], [-anaphor]. That is, ungoverned PRO has the same features as a pronoun does. If ungoverned PRO and pronouns have the same features, they would have to obey the same binding constraints. Therefore, just as governed PRO obeys Principle A, if ungoverned PRO is not an anaphor and is pronominal, then it has to obey binding Principle B.

However, binding Principle B states that "a pronominal is free in a local domain". "Local domain" is defined as a governing category. If ungoverned PRO obeys Principle B, then (112) follows:

(112) Ungoverned PRO is free in its governing category.

At best, (112) makes no sense. Although H&L claim that ungoverned PRO is pronominal, the implications of this claim remain unclear. In particular, the following questions remain unanswered: What is the relation between ungoverned PRO and Principle B? What constraints, if any, does the binding of ungoverned PRO obey? The only constraints on ungoverned PRO offered by H&L is that it does not act like an anaphor. Anaphors are defined by van Reimsdijk & Williams (1986) as follows: "Anaphors are NPs that require an antecedent." Thus, if ungoverned PRO does not act like an anaphor, then it should always be able to have an arbitrary interpretation. In fact, H&L make just this claim:

An ungoverned PRO is not an anaphor, and therefore does not require an antecedent of any kind

However, in their description of the behavior of governed PRO, H&L show it acting like an anaphor. Consider the examples offered by H&L; bearing in mind that PRO is an *anaphor* if it requires an antecedent, and a *pronoun* if it need not have an antecedent.

(113) John_i tried/wanted [PRO_i to leave]

(114) It is unclear to John [how [PRO to behave]]

(115) It is unclear to John_i [how [PRO_i to behave himself]]

(116) It is unclear to John [how [PRO_i to behave himself]]

In sentence (115), the ungoverned, "pronominal" PRO is not acting like a pronoun. In (115), PRO is an anaphor: PRO must have an antecedent of the sentence is ill-formed, as in (117):

(117) * It is unclear [how [PRO_i to behave himself_i]]

A PRO coindexed with a definite reflexive (*himself*, *herself* ...) must have an antecedent, and therefore, is by definition an anaphor. Consider the other examples cited by H&L. In each case, the ill-formedness suggests that PRO requires an antecedent, i.e. PRO is an anaphor.

(118) Mary knows that PRO to behave would help John.

(119) It was known that PRO to behave would help John.

(120) *Mary* knows that PRO to behave *herself* would help John.

(121) * It was known that PRO to behave *herself* would help John.

(122) Mary knows that PRO to behave *himself* would help *John*.

(123) * Mary knows that PRO to behave *himself* would help.

H&L incorrectly define the basic problem when they say:

If another NP is available, PRO may or may not be coindexed with it; thus *It is unclear to John how to behave* may be continued with *himself* or *oneself*. (H&L 1987:29-30 footnote 5)

When continued with *himself*, yielding (117), PRO *must* be coindexed with an NP which is singular and masculine. That is, just as the anaphor *himself* requires a masculine singular antecedent; PRO in (5) requires a masculine singular antecedent, and is therefore, by definition, an *anaphor*.

A fundamental difficulty in the H&L analysis derives from their use of the indefinite reflexive (*oneself*) as a diagnostic for ungoverned, or non-obligatory control, PRO: PRO_{arb}. The sentences in (124) contain no possible antecedent for PRO. In these cases, H&L and Manzini predict that PRO receives the arbitrary interpretation, i.e. is PRO_{arb}.

PRO can be the antecedent of *oneself* as well as the antecedent of *myself*, *yourself*, and *ourselves*, but not as the antecedent of *himself*, *herself*, *yourselves*, or *themselves*, as is illustrated in the sentences in (124).

- (124) i. It is impossible PRO to make myself at home here.
 ii. It is impossible PRO to make yourself at home here.
 iii. It is impossible PRO to make oneself at home here.
 iv. * It is impossible PRO to make himself at home here.
 v. * It is impossible PRO to make herself at home here.
 vi. It is impossible PRO to make ourselves at home here.
 vii. * It is impossible PRO to make yourselves at home here.
 viii. * It is impossible PRO to make themselves at home here.

Consider now the examples in (125). PRO can be the antecedent of *oneself* as well as the antecedent of *himself*. It cannot be the antecedent of *myself*, *yourself*, *herself*, *ourselves*, *yourselves* or *themselves*.

- (125) i. * It is unclear to John how PRO to behave myself.
 ii. * It is unclear to John how PRO to behave yourself.
 iii. It is unclear to John how PRO to behave oneself.
 iv. It is unclear to John how PRO to behave himself.
 v. * It is unclear to John how PRO to behave herself.
 vi. * It is unclear to John how PRO to behave ourselves.

- vii. * It is unclear to John how PRO to behave yourselves.
- viii. * It is unclear to John how PRO to behave themselves.

How can analyses which assign PRO_{arb} as the subject of the infinitives in (124) and (125), permit the range of reflexive forms in (124), but block the range of reflexive forms in (125)?

In order to see more clearly what is going on in (124) and (125), we need to differentiate two questions: (a) What is the distribution and interpretation of the indefinite reflexive *oneself*? and (b) What is the distribution and interpretation of PRO_{arb}? We will discuss these questions at length in "The Distribution and Interpretation of PRO" on page 204. These two independent questions are merged into one by using *oneself*, as a diagnostic test for PRO_{arb}. If one agrees with the theories that predict that all instances of PRO in (124) and (125) are assigned PRO_{arb}, then there are two problems: (c) Why is (125iii) understood to assert that PRO includes *John*? and (d) Why are reflexive forms which are good in (124) blocked in (125)?

5.6.1 "ASK" IN THE HORNSTEIN AND LIGHTFOOT THEORY OF PRO

H&L would determine that PRO in (1) is an anaphor and PRO in (2) and (3) is pronominal. However, the governed PRO in (1) does not act more anaphoric than ungoverned PRO in (2) and (3). This suggests that a *wh*-element in COMP, which blocks government, does not prevent PRO from being an anaphor.

The verb *ask* provides evidence against the claim that a *wh*-element in COMP prevents PRO from being an anaphor. Since COMP in (1) is neither lexically filled nor filled by an empty category, S' is deleted and the governed PRO is an anaphor.

- (1) John asked Mary to put the book on the table.

By Principle A, governed PRO must have an antecedent. H&L do not offer a mechanism for choosing between subject and object control, they simply assume such a mechanism. In the case of *ask*, presumably this mechanism would bind PRO to *Mary*.

Since COMP is lexically filled in (2) and (3), PRO is ungoverned.

- (2) John asked Mary how to put the book on the table.
 (3) John asked Mary where to put the book.

H&L would predict that PRO in (2) and (3) gets the arbitrary interpretation. However, by their test, PRO should be able to be the antecedent of *himself* or *herself* or *oneself*. That is, all of the examples in (126) should be equally well-formed.

- (126) i. John asked Mary how to behave himself.
 ii. * John asked Mary how to behave herself.
 iii. John asked Mary how to behave oneself.

In addition, ungoverned PRO in a context with no available antecedent should be well-formed. However, the sentences in (127) are, at best, odd.

- (127) i. * It was asked how drive there.
 ii. * It was asked how to behave (oneself)

The PRO subject of *put* must be *who* in (128) and must be *John* in (129) and (130). *Himself* refers to *who* in (131) and *John* in (132). The system of H&L maintains, contrary to the facts, that the PRO subjects in (129), (130) and (132) are (or perhaps, *can* or *may* be) PRO_{arb}.

- (128) *Who* did John ask to put the book on the table?
 (129) Who did *John* ask how to put the book on the table?
 (130) Who did *John* ask where to put the book?

(131) *Who* did John ask to wash *himself*?

(132) Who did *John* ask when to wash *himself*?

Another problem with the H&L system is that sometimes governed PROs, which they claim are anaphors, do not act like anaphors. An anaphor must have a c-commanding antecedent in its governing category. However, there are counterexamples where governed PRO is controlled by a non-c-commanding antecedent. Consider the sentences in (133) and (134). H&L predict that both sentences in (133) should be well-formed, because governed PRO is coindexed with the only available c-commanding NP. However, (133ii) is ill-formed. Similarly, both sentences in (134) should be ill-formed, because governed PRO is coindexed with a non-c-commanding antecedent. However, (134ii) is well-formed.

- (133) i. *Mary* swore to Bill [PRO to make up *her* mind by morning]
 ii. * *Mary* shouted to Bill [PRO to make up *her* mind by morning]

- (134) i. * *Mary* swore *to Bill* [PRO to make up *his* mind by morning]
 ii. *Mary* shouted *to Bill* [PRO to make up *his* mind by morning]

The H&L analysis analysis would predict that PRO in (135i) is assigned an antecedent by whatever device is used to choose between subject and object control. This same device would block (135ii). However, since the antecedent is a PP in (136), their theory predicts, contrary to fact, that (136i) is well-formed because PRO is coindexed with the only available c-commanding NP, while (136ii) should be ill-formed because governed, anaphoric, PRO is coindexed with a non-c-commanding NP.

- (135) i. *Mary* signaled *Bill* [PRO to make *himself* at home]
 ii. * *Mary* signaled Bill [PRO to make *herself* at home]

- (136) i. * *Mary* signaled to Bill [PRO to make *herself* at home]
 ii. *Mary* signaled to *Bill* [PRO to make *himself* at home]

5.7 CHOMSKY'S DESCRIPTION OF CONTROL PROPERTIES IN "KNOWLEDGE OF LANGUAGE"

Chomsky (1986a) appears to have abandoned trying to formulate a comprehensive control theory. Nor does he appear to adopt anybody else's control theory, including those just discussed. He continues to maintain, however, that PRO is a pronominal anaphor: sometimes exhibiting properties of pronouns and sometimes exhibiting properties of anaphors.

The Chomsky (1986a) discussion of Control is, in effect, a catalog of the various control phenomena. His description of control properties ends inconclusively, with the following passage.

These and many similar examples indicate that factors of a rather complex kind enter into control theory, not all of them well understood. PRO is anaphorlike in its interpretation and in that it is in a certain structural relation to its binder: either in its domain or with the binder sufficiently prominent in the construction K where PRO is the subject of the subject of K. PRO is pronounlike in that it can take a split antecedent, a "remote" controller, and an object controller when it is not locally bound and, we will continue to assume, in that it can be free as well as bound. Of the potential binders, the most prominent is the obligatory binder in certain constructions. Questions remain open about how to make these notions precise and how to explain the range of observed phenomena. (Chomsky 1986a:130-131)

I will outline what he presents as the range of observed phenomena in terms of properties of pronouns and anaphors. A pronoun is (a) free in a local domain, (b) may be either free or non-locally bound, and (c) may have split antecedents. An anaphor is (a) bound in a local domain (b) has to be in the domain of the antecedent -- the relation is usually taken to be that of c-command, and (c) does not have split antecedents.

5.7.1 CONTROL IN DECLARATIVE SENTENCES

The structure for the core case of control, that is control in declarative sentences, is represented in (137). In this configuration "PRO must be bound if α has a potential binder as its subject or object."

(137) [α α (NP) [C PRO to VP]]

PRO is like an anaphor in the following four respects:

1. In declarative sentences, represented in (137), when there is an "appropriately related" potential controller, PRO is bound by an antecedent:

... its structural relation to its controller is essentially the same as that of an antecedent-anaphor pair, and its interpretation is very much like that of a reflexive.
(Chomsky 1986a:125)

Chomsky gives, as examples, (138)-(140):

(138) John decided [PRO to vote twice]

(139) John persuaded Bill [PRO to vote twice]

(140) the students asked the teacher [PRO to leave the room]

Subject control is exhibited by (138), object control by (139), and subject or object control in (140). Chomsky concludes that "Evidently, lexical and perhaps other factors are involved in the choice of controller" (Chomsky 1986a:125).

Thus in the configuration of (137), when a possible controller is present, PRO is obligatorily controlled by an antecedent which is, at least to some extent, specified in the lexicon.

2. PRO is like an anaphor "in that it (typically) lacks independent specific reference". PRO either gets its reference from an antecedent or is interpreted as arbitrary.

3. When PRO is in the declarative construction (137) it, like reflexives, cannot take split antecedents, illustrated in (141).

(141) i. * Bill wanted [Tom to decide [PRO to swim across the pond together]]

ii. * Bill wanted [Tom to feed themselves]

4. PRO is anaphorlike "in that it must be in the domain of its antecedent (if any), as distinct from a pronoun".

(142) i. * John's friends think it is illegal [PRO to feed himself]

ii. John's friends think it is illegal [for him to feed himself]

In (142i) PRO is required to be in the domain of its antecedent. As (142ii) shows, a pronoun does not have to be in its antecedent's domain.

In the same declarative configuration (137), PRO is like a pronoun in the following three respects:

1. PRO may be unbound when there is no antecedent "appropriately related to the head α ". In this case, PRO can be free as in (143) or bound with a more remotely related antecedent as in (144):

(143) i. it is illegal [PRO to vote twice]

ii. it was decided [PRO to vote twice]

(144) i. John announced the decision [PRO to feed himself]

ii. John thinks it is illegal [PRO to feed himself]

iii. John thought Mary said that the decision [PRO to feed himself] was foolish.

2. PRO is also like a pronoun in that "it may take a split antecedent when not locally bound":

(145) Bill wanted [Tom to approve the decision [PRO to swim across the pond together]]

However, Chomsky continues, PRO is also anaphoric in that PRO must be in the domain of both of its antecedents. He claims that the antecedent in (146) cannot be *Bill* and *Tom*.

(146) Bill's mother wanted [Tom to agree that it was time [PRO to swim across the pond together]]

3. Whereas non-locally bound anaphors have to be bound by subjects, specifically the nearest subject, when PRO is not locally bound, it can be bound by the subject or object, as in (147).

(147) i. they told me that the decision [PRO to feed themselves was foolish]

ii. they told me that the decision [PRO to feed myself was foolish]

Chomsky notes that the facts for control are somewhat different when PRO is (a) interrogative complements; (b) in adjunct phrases; and (c) subject sentences.

5.7.2 CONTROL IN INTERROGATIVE COMPLEMENTS

Chomsky uses the reflexive test to illustrate how interrogative complements, or indirect questions, manifest properties of both pronouns and anaphors.

Chomsky claims that indirect question complements differ from declarative complements in that an arbitrary interpretation is possible for PRO even when a possible controller exists. He illustrates this point with (148) and (149) saying that PRO is pronoun-like in that it can be either bound or free in these constructions.

(148) they asked me [how PRO to feed β]

(149) they told me [how PRO to feed β]

In (148) PRO is bound when $\beta = \textit{themselves}$ and is free when $\beta = \textit{oneself}$. Likewise, in (149) PRO is bound when $\beta = \textit{myself}$ and free when $\beta = \textit{oneself}$.

PRO is anaphor-like, he continues, in that the choice of antecedents for PRO is local. That is, (150) is good if $\beta = \textit{themselves}$ or *oneself* but not when $\beta = \textit{myself}$.

(150) I thought they wondered [how PRO to feed β]

The choice of antecedent is also more structurally restricted for PRO than for a pronoun. In (151), β can be *herself* or *oneself* but not *himself*.

(151) John's mother asked me [how PRO to feed β]

A further anaphor-like property, which Chomsky does not point out, is that PRO in (148), (149) and (151) cannot freely be assigned an antecedent in either subject or object position. When *ask* is in the matrix clause, the antecedent cannot be the object; when *tell* is in the matrix clause, the antecedent cannot be the subject. Thus the options for PRO are even more restricted than Chomsky suggests.

5.7.3 CONTROL IN ADJUNCTS

In adjuncts, such as purpose clauses, Chomsky notes that, as with declarative complements, PRO is like an anaphor in that when local binding is possible it is obligatory. Thus (152) is good when $\beta = \textit{myself}$ but bad when $\beta = \textit{themselves}$ or *oneself*.

(152) they thought I bought the presents [PRO to amuse β]

Chomsky assumes also that (153) is an adjunct phrase.

(153) John is too stubborn [PRO to talk to β]

He points out that "PRO will be either bound or free, depending on the choice of β (as *Bill* or *trace*, respectively)." In (153) PRO is anaphor-like in that it is obligatorily bound by *John* when β is lexically filled. When $\beta=e$, then PRO has to be free. Presumably this is a pronoun-like quality, although pronouns are optionally, not obligatorily, free. Chomsky continues that in (154) PRO is pronoun-like in that it can be bound or free.

(154) i. we told them that John is too stubborn [PRO to bother β about]

ii. I thought you said that John is too stubborn [PRO to bother β about]

β in (154i) can be *ourselves*, *themselves* or *oneself*, and β in (154ii) can be *ourselves*, *myself*, *yourself* or *oneself*. Thus PRO is pronoun-like in that it can either be free or bound by *we*, *them*, *you*, or *I*, and in that it can take a split antecedent. However, PRO is anaphor-like in that, when it is bound, PRO must be in the domain of the antecedent. That is, in (154i), *we* cannot be replaced by *our friends* when $\beta=ourselves$.

(155) * our friends told them that John is too stubborn [PRO to bother ourselves about]

5.7.4 CONTROL IN SUBJECT SENTENCES

In subject sentences, if there is a potential binder, PRO must be bound, as in (156), and an arbitrary reading, a reading with *oneself* is excluded, as in (157).

(156) [PRO to have to feed himself] would be a nuisance for John

(157) * [PRO to have to feed oneself] would be a nuisance

(Note that this contradicts Manzini's claims about these constructions.)
Chomsky continues that when no potential binder is present, as in (158), then PRO gets the arbitrary interpretation.

(158) [PRO to have to feed oneself] is a nuisance

When PRO is in a subject sentence, it is not in the domain of a binder.

The binder is, in an intuitive sense, the most prominent element within the complement that makes sense as a binder... (Chomsky 1986a:129)

Thus *John's friends* makes sense as a binder in (159i) and *John* in (159ii). In addition, the binder can be in a PP complement, as in (159iii). However, the antecedent cannot be too deeply embedded, as in (160).

- (159) i. [PRO to have to feed themselves] would annoy John's friends
 ii. [PRO to have to feed himself] would assist John's development
 iii. [PRO to have to feed himself] would be a nuisance for John
- (160) * [PRO to be able to feed himself] would cause us to conclude that John is competent

5.8 C-COMMAND

Contrary to Chomsky's assertion, PRO is not always in the domain of its controller in declarative sentences. In the discussion outlined above, Chomsky maintains that when PRO is bound, it has to be in the domain of its antecedent. However, PRO can also be bound by an antecedent that does not have PRO in its domain, as in (161).

(161) John signaled [to Mary][PRO to leave]

In (161), PRO is pronoun-like in that it is not in the domain of its antecedent. However, it is anaphorlike in that *Mary* is obligatorily its antecedent. The existence of verbs like (161) which have control

from the PP complement suggest that the structural description for control in (137) should be revised to indicate that an NP or PP can appear in the object position of α .

The following examples indicate that when control is local, PRO need not be in the domain of its antecedent. In the first place, if the controller must c-command PRO, then (162i)-(165i) are control structures while (162ii)-(165ii) are not:

- (162) i. John motioned Mary to sit down.
- ii. John motioned to/at Mary to sit down.

- (163) i. The general signaled his officers to begin the attack.
- ii. The general signaled to his officers to begin the attack.

- (164) i. John beckoned the child to come closer.
- ii. John beckoned to the child to come closer.

- (165) i. John cabled/telephoned/wrote Mary to come home.
- ii. John cabled/telephoned/wrote to Mary to come home.

If the controller must c-command PRO then (162i)-(165i) are control structures. However, in order to handle (162ii)-(165ii) a device is needed to assign an antecedent for PRO which resembles control in every way except that it does not require c-command.

In addition, if a possible controller must c-command, then all of the examples in (166) should be controlled by the subject position. However, control is from the PP object.

- (166) i. John signaled to Mary to sit down.
- ii. John called on/upon Mary to answer the question.
- iii. John depended on/upon Mary to carry out the garbage.
- iv. John motioned to/at Mary to sit down.
- v. John relied on/upon Mary to do the laundry.

As with *count on* discussed in Williams (1980:214) these verb-preposition pairs have not been reanalyzed as complex verbs with NPs as direct ob-

jects. By Williams' test, since all of the examples in (166) can be pied piped, reanalysis cannot have occurred.

- (167) i. To whom did John signal to sit down?
 ii. Upon whom did John call to answer the question?
 iii. Upon whom did John depend to carry out the garbage?
 iv. At whom did John motion to sit down?
 v. On whom did John rely to do the laundry?

Thus the list of anaphorlike properties of PRO should be revised to indicate that PRO is pronominal in that it does not have to be in the domain of its controller.

5.9 CONTROL IN PASSIVE CONSTRUCTIONS

There is a considerable literature on the control properties of passive and double passive constructions. The main facts are summarized below.

Generally, verbs with subject control do not passivize. Object control verbs do passivize. Thus one gets the distinction between (168ii) and (169ii).

- (168) i. John promised Mary to leave.
 ii. * Mary was promised to leave.

- (169) i. John persuaded Mary to leave.
 ii. Mary was persuaded to leave.

However, this is not always true. Subject control verbs do passivize when they have indirect question complements, as in (171). Their passives differ from those of object control verbs in that PRO gets the arbitrary interpretation. Thus the difference in reference for PRO in (170ii) and (171ii).

- (170) i. Mary told John how to win.
 ii. John was told how to win.

- (171) i. John asked Mary how to win.
 ii. Mary was asked how to win.

In double passive constructions, the distribution of data is much more complex. As has already been noted, *promise*, which cannot passivize in a single passive construction, does appear in the double passive, as in (172).

(172) Mary was promised to be allowed to leave.

However, verbs by no means take these double passive constructions freely. For example, *persuade* and *ask* do not.

(173) * Mary was persuaded to be allowed to leave.

(174) * Mary was asked to be allowed to leave.

(175) * Mary was asked when to be allowed to leave.

None of the control theories covered attempt to predict the antecedent for PRO in the double passive constructions.

5.10 SUMMARY

The control theories I have covered have tried to bifurcate control phenomenon. Williams hypothesizes that control can be predicted based upon whether or not the infinitive is in a predicate position. Manzini tries to predict control based on whether or not the infinitive has a domain-governing category. Koster and H&L try to predict control relations based upon whether or not PRO is governed. As I have shown, these systems either predict control incorrectly, or are so underspecified that they only predict when something is controlled, not what the controller is.

In the various environments in which infinitive clauses can occur, PRO never exhibits properties which are exclusively anaphoric or exclusively pronominal. Thus it appears that Chomsky's (1981) conclusion

that PRO is simultaneously an anaphor and a pronoun, a pronominal anaphor, holds.

6.0 THE DISTRIBUTION AND INTERPRETATION OF PRO

There are three basic questions concerning PRO:

1. What is the distribution of PRO?
2. What is the interpretation of PRO?
3. If PRO is interpreted as having an antecedent or antecedents, which NP(s) in the string is/are the antecedent(s)?

In response to (1), I assume Chomsky's analysis that the distribution of PRO follows from the PRO theorem: PRO is [+anaphoric,+pronominal] and therefore can only occur in a structure if PRO is ungoverned. Since all instances of PRO carry the features [+a,+p], I do not include them in future examples. The presence of these features, however, is assumed.

6.1 THE DISTRIBUTION OF PRO

An anaphor has the following three properties:

1. An anaphor has a c-commanding antecedent in the domain of its governing category.
2. An anaphor must obligatorily have an antecedent, with the correct person, number and gender, in the sentence.
3. The antecedent for an anaphor must be a single node. Anaphors cannot have split antecedents.

Examples of anaphors include: *herself*, *myself*, *oneself*, and *each other*.

A pronoun has the following three properties:

1. A pronoun must be free in its governing category.

2. A pronoun does not require any antecedent.
3. If a pronoun has an antecedent, it may be (a) more than one node and (b) have referents outside of the sentence. Pronouns may take split antecedents.

Some examples of pronouns include: *her, she, us, one, and you.*

The diagnostics for determining whether an element is a pronoun or an anaphor are: whether it is bound or free in its governing category; whether it requires an antecedent; whether the antecedent must c-command it; and whether it can take split antecedents.

PRO, the null subject of infinitives and gerunds, sometimes exhibits properties of anaphors and sometimes exhibits properties of pronouns. In some instances, like an anaphor, PRO must be bound to an antecedent. Compare PRO in (1) with the anaphor *himself* in (2).

(1) John_{*i*} is eager [PRO_{*i*} to please]

(2) John_{*i*} likes himself_{*i*}

In other instances, like a pronoun, PRO is free when it lacks a possible antecedent. Thus PRO in (3) is free, as is the pronoun *him* in (4).

(3) It is time [PRO_{*arb*} to leave]

(4) It is time [for him_{*i*} to leave]

Also like a pronoun, there are some configurations in which PRO must be free even when there is a possible antecedent. Compare (5) where PRO must have an arbitrary interpretation with (6) where the pronoun *him* cannot be coindexed with *John*.

(5) John is easy [PRO_{*arb*} to please]

(6) John_{*i*} likes him_{*j*}

Again, like a pronoun, there are instances where PRO can either be bound or free. Compare (7) where PRO can either be bound by *John* or have an arbitrary interpretation with (8) where the pronoun *him* can either refer to *John* or some other unspecified person.

- (7) i. John_i said that it would be appropriate [PRO_i to leave]
 ii. John said that it would be appropriate [PRO_{arb} to leave]
- (8) i. John_i said that it would be appropriate [for him_i to leave]
 ii. John_i said that it would be appropriate [for him_j to leave]

Like an anaphor, PRO cannot take split antecedents when it is locally bound, as in (9). However, when PRO is not locally bound it, like a pronoun, can take a split antecedent, as in (10).

- (9) * John wanted Bill [PRO to go to the movies together]
- (10) John wanted Bill to accept the invitation [PRO to go to the movies together]

This kind of distribution of data led Chomsky to conclude that PRO has the features of [+anaphor] and [+pronominal].

We then take PRO to be a pronominal anaphor [+a, +p], sharing the properties of pronouns and anaphors.... (Chomsky 1986a:164)

What does it mean to say that PRO has the two features [+anaphor] and [+pronoun]? It could mean that in some contexts PRO is an anaphor and in other contexts PRO is a pronoun. Or it could mean that PRO is a summation of the two features and is both [+anaphor] and [+pronoun]. Or it could mean that PRO is a hybrid, neither an anaphor or pronoun, which is partially [+anaphor] and partially [+pronoun]. If we regard nodes and categories as feature complexes, then this last hybrid-PRO has some, but perhaps not all, of the features of anaphors and pronouns.

Chomsky (1981) assumes that PRO has the features of [+anaphor] and [+pronoun] simultaneously. NPs which are anaphors are subject to

Binding Principle A and NPs which are pronominal are subject to Binding Principle B. Thus PRO has to obey the constraints imposed by Principles A and B:

- (11) i. Principle A: An anaphor is bound in a local domain.
 ii. Principle B: A pronominal is free in a local domain.

This leaves us with the logical contradiction that a single element is both bound and free (not bound) in a single domain. To avoid the contradiction, Chomsky concludes that PRO cannot have a local domain.

Chomsky defines *local domain* as follows:

... the local domain for an anaphor or pronominal α [is] the minimal *governing category* of α , where a governing category is a maximal projection containing both a subject and a lexical category governing α (hence, containing α). (Chomsky 1986a:169)

In order for PRO not to violate either binding constraint, it is required not to have a governing category. Hence PRO must always be ungoverned.

The position that PRO is never governed, which is known as the *PRO Theorem*, continues to be maintained by Chomsky and has been adopted by Williams (1980). However, the Pro Theorem is not universally adopted within the Principles-and-Parameters (or GB) framework.

While everyone agrees with Chomsky's insight that PRO has features [+anaphor] and [+pronoun], it is not agreed that PRO has both features at the same time. For example, Hornstein and Lightfoot view PRO as behaving sometimes like an anaphor and at other times like a pronoun. Their research focuses on solving the Control problem by determining under what conditions PRO is [+anaphor] and under what conditions PRO is [+pronoun].

Manzini (1983) argues that PRO is always an anaphor, but that in certain structural configurations, anaphors do not require antecedents.

Koster (1984, 1987) and Hornstein & Lightfoot (1987) contend that PRO is not simultaneously [+anaphor] and [+pronoun]. They argue, further, that PRO can be governed. According to Koster, when PRO is governed, it is an anaphor and subject to Binding Principle A, and when it is ungoverned it is subject to Control Theory. According to Hornstein and Lightfoot, when PRO is governed, it is an anaphor and subject to Binding Principle A, and when PRO is ungoverned it is pronominal. Ungoverned PRO "will receive the arbitrary interpretation".

In all structural contexts where infinitives appear -- whether as complements, adjuncts, or clausal subjects -- the subject PRO exhibits properties of both anaphors and pronouns. This leads us to conclude, as does Chomsky, that PRO has as its inherent feature complex [+pronominal, +anaphor]. The question remains of how it is determined which features of pronouns and which features of anaphors each instantiation of PRO exhibits.

I propose that two extrinsic features are assigned to PRO by the parser. These features, taken in combination, determine which anaphoric and pronominal properties PRO exhibits. The features are called [\pm generic] and [\pm definite]. The definite and generic features are assigned based on (1) lexical structures and (2) structural configurations.

A fundamental part of this analysis is that PRO is sometimes interpreted as the empty pronoun *one*. Before presenting an analysis of the [generic] and [definite] features, I show that PRO often receives the same interpretation as the overt pronoun *one*, and how PRO may have emerged from *one* diachronically.

6.2 THE PRONOMINAL ANAPHOR 'ONE'

An anaphor requires an antecedent or has to be bound, and a pronoun optionally has an antecedent or can be free. Hence, in sentences (12)-(14), where no antecedent is possible, PRO is being pronoun-like.

(12) It is unclear [how [PRO to wash out that stain]]

(13) [PRO to behave oneself in public] is a social requisite

(14) It is a social requisite [PRO to behave oneself in public]

I will argue that when PRO is the antecedent of *oneself* it is not arbitrary in reference. Instead it has many of the referential properties of the lexical indefinite pronoun *one*. When PRO has no antecedent, it exhibits the same referential properties as it does when it is the antecedent of *oneself*. I will call this non-coindexed element PRO_{one}.

The assumption that what has been known as arbitrary PRO is really the empty pronoun *one* helps to explain the phenomenon of split antecedents (excluding lexically assigned split antecedents).¹⁶ The relevance of split antecedents for any of the theories discussed in "Control Facts and Control Theories" on page 141 remains unclear since split antecedents are not assigned by any of their mechanisms. Nevertheless, as we have seen, split antecedents are commonly used as a diagnostic for non-obligatory control.

The similarities between arbitrary PRO and the pronoun *one* has not gone unnoticed. Chomsky writes:

¹⁶ I distinguish between a split antecedent in long-distance control, as in *John told Mary that it would be appropriate PRO to leave* and split antecedents which are both within the argument structure of the verb, as in *John proposed to Mary PRO to go to the movies*.

Either PRO is bound, or it is free with (typically) an arbitrary interpretation similar to *one*, as in "one's beliefs often prove false." (Chomsky 1986a:164)

I explore the precise relation between the pronoun *one* and arbitrary PRO below. The term arbitrary PRO or PRO_{arb} is a misnomer, since its interpretation is not, in fact, arbitrary.

6.2.1 THE PERSON, NUMBER, AND GENDER OF 'ONE' AND PROONE

Concerning the distribution of the pronoun *one*: *One* is syntactically and morphologically a third person singular pronoun. Concerning the interpretation of the pronoun *one*: Semantically, according to the *OED*, *it* refers to the speaker (first person), the hearer (second person), and anyone else (third person). Consider these passages from the *OED* (emphasis added):

V. *one*. *Indefinite pronoun* (with genitive *one's*).

20. A person or being whose identity is left undefined; some one, a certain one, an individual, a person (L. *quidam*). A following pronoun referring to *one* is in the 3rd. pers. sing., as 'One showed himself to his townsmen, who derided him'. (*OED*, VII:121)

21. Any one of everybody; any one whatever; including (and in later language often specially meaning) the speaker himself; 'you or I, or any one'; a person, a man; we, you, people, they (=OE. *man*, ME. *me*, G. *man*, F. *on*). Poss. *one's*, obj. *one*, reflexively ONESELF (formerly *one's self*); but for these the third person pronouns *his*, *him*, *himself* were formerly usual, and are still sometimes used; thus, 'If one showed oneself (himself) to one's (his) townsmen, they would know one.' The pl. pronouns *their*, *them*, *themselves* were formerly in general use on account of their indefiniteness of gender, but this is now considered ungrammatical.) In this sense *one* is quite toneless (/w n/), proclitic or enclitic. (*OED*, VII:121-2)

Oneself. pron. Also ones, one's self. [orig. one's self (see ONE 21 and SELF), after *my self*, etc.; afterwards assimilated to himself, itself.] An emphatic or distinctive equivalent of the indefinite pronoun ONE, used chiefly in the objective (after vb. or prep) or (in sense 1) as a nominative in apposition. The corresponding possessive is *one's own*: 'occupied with oneself and one's own affairs.'

1. Emphatic use: A person's self; himself or herself (meaning or including the speaker or writer).

2. Reflexive use: objective case of ONE 21, as 'One is obliged to keep oneself by oneself.'
(*OED*, VII:125)

The above quotes are summarized in (15).

- (15) *one*, along with the possessive *one's* and the reflexive *oneself*, can mean "anyone whatever" including (a) "the speaker himself"; (b) "you or I, or anyone"; (c) "a person's self, himself or herself (meaning or including the speaker or writer)".

One does not have all interpretations in all contexts. Since lexical *one* is syntactically and morphologically a third person singular pronoun, it has to agree in person and number with another pronoun in the string. Thus *one* licenses all and only third person singular forms, both definite and indefinite, with the exception of *its* and *itself*.

Since *PRO_{one}* is an empty category, it lacks morphological constraints, and so has a less restricted distribution than lexical *one*. The distribution of *PRO_{one}* more closely follows the semantics of the definition of *one* quoted from the *OED* above, and summarized in (15). Like overt *one*, *PRO_{one}* does not co-occur with second and third person plural forms which exclude the speaker from their scope. Also like *one*, *PRO_{one}* does co-occur with *one*, *one's* and *oneself*, and does not license *its* and *itself*. However, unlike lexical *one*, *PRO_{one}* licenses first and second person pronouns which include the speaker and hearer: *me*, *my* and *myself*; *you*, *your* and *yourself*; and *us*, *our* and *ourselves*. Unlike *one*, the only third person forms that *PRO_{one}* takes is *one*, *one's* and *oneself*. Other third person forms with *PRO_{one}* is interpreted as excluding the speaker, which violates the semantics of *one*. When the third person form *one* occurs with *himself*, *herself* or *oneself*, the in-

terpretation includes the speaker. I illustrate these facts about the distribution and interpretation of overt *one* and PRO_{one} below:

Reflexive Pronouns

The overt pronoun *one* licenses all and only third person singular animate forms: *herself*, *himself*, and *oneself*, as illustrated in (16):

- (16) i. * It is unclear [how [one can keep myself occupied here]]
 ii. * It is unclear [how [one can keep yourself occupied here]]
 iii. It is unclear [how [one can keep himself occupied here]]
 iv. It is unclear [how [one can keep herself occupied here]]
 v. It is unclear [how [one can keep oneself occupied here]]
 vi. * It is unclear [how [one can keep itself occupied here]]
 vii. * It is unclear [how [one can keep ourselves occupied here]]
 viii. * It is unclear [how [one can keep yourselves occupied here]]
 ix. * It is unclear [how [one can keep themselves occupied here]]

PRO_{one} can be the antecedent of the reflexive pronouns *myself*, *yourself*, *oneself*, and *ourselves*, but not of *himself*, *herself*, *itself*, *yourselves*, or *themselves*, as illustrated in (17):

- (17) i. Dear Warden, it is unclear [how [PRO_{one} to keep myself occupied here]]
 ii. Dear Warden, it is unclear [how [PRO_{one} to keep yourself occupied here]]
 iii. * Dear Warden, it is unclear [how [PRO_{one} to keep himself/herself occupied here]]
 iv. Dear Warden, it is unclear [how [PRO_{one} to keep oneself occupied here]]
 v. Dear Warden, it is unclear [how [PRO_{one} to keep ourselves occupied here]]
 vi. * Dear Warden, it is unclear [how [PRO_{one} to keep yourselves occupied here]]
 vii. * Dear Warden, it is unclear [how [PRO_{one} to keep themselves occupied here]]

The reflexive pronouns licensed by *one*, as in (16), and PRO_{one}, as in (17), are summarized in Figure 21 on page 213.

	one		PRO _{one}	
	SINGULAR	PLURAL	SINGULAR	PLURAL
FIRST	*myself	*ourselves	myself	ourselves
SECOND	*yourself	*yourselves	yourself	*yourselves
THIRD	himself herself *itself oneself	*themselves	*himself *herself *itself oneself	*themselves

Figure 21. Reflexive Pronouns Licensed by *one* and PRO_{one}

Possessive Pronouns The overt pronoun *one* licenses all and only third person singular animate forms, both definite: *her*, *him*, and indefinite: *one's*, as illustrated in (18):

- (18) i. * It is unclear [how [one can keep my/your self-respect here]]
 ii. It is unclear [how [one can keep his/her/one's self-respect here]]
 iii. * It is unclear [how [one can keep our/your/their self-respect here]]
 iv. * It is unclear [how [one can keep its self-respect here]]

PRO_{one} can be the antecedent of *my*, *your (sg)*, *one's* and *our*, but not of *his*, *her*, *its*, *your (pl)* or *their*, as illustrated in (19):

- (19) i. Dear Warden, it is unclear [how [PRO_{one} to keep my self-respect here]]
 ii. Dear Warden, it is unclear [how [PRO_{one} to keep your self-respect here]]
 iii. * Dear Warden, it is unclear [how [PRO_{one} to keep his/her/its self-respect here]]
 iv. Dear Warden, it is unclear [how [PRO_{one} to keep one's self-respect here]]
 v. Dear Warden, it is unclear [how [PRO_{one} to keep our self-respect here]]
 vi. * Dear Warden, it is unclear [how [PRO_{one} to keep your self-respect here]]
 vii. * Dear Warden, it is unclear [how [PRO_{one} to keep their self-respect here]]

	one		PRO _{one}	
	SINGULAR	PLURAL	SINGULAR	PLURAL
FIRST	*my	*our	my	our
SECOND	*your	*your	your	*your
THIRD	his her *its one's	*their	*his *her *its one's	*their

Figure 22. Possessive Pronouns Licensed by *one* and PRO_{one}

The possessive pronouns licensed by *one*, as in (18), and PRO_{one}, as in (19), are summarized in Figure 22 on page 214.

Objective Personal Pronouns

The objective form of overt *one* is *one*, *him*, or *her*, as illustrated in (20):

- (20) i. * For one to carry money on me is not a good idea.
 ii. * For one to carry money on you is not a good idea.
 iii. For one to carry money on him/her/one is not a good idea.
 iv. * For one to carry money on us/you/them is not a good idea.

The objective form of PRO_{one} is *me*, *your (sg)*, *one* and *us*, but not *him*, *her*, *it*, *you (pl)* or *them*, as illustrated in (21):

- (21) i. PRO_{one} to think that Mary loves me is a silly idea.
 ii. PRO_{one} to think that Mary loves you is a silly idea.
 iii. PRO_{one} to think that Mary loves one is a silly idea.
 iv. * PRO_{one} to think that Mary loves him is a silly idea.
 PRO_{one} = him
 v. PRO_{one} to think that Mary loves us is a silly idea.
 vi. * PRO_{one} to think that Mary loves you is a silly idea.
 PRO_{one} = you
 vii. * PRO_{one} to think that Mary loves them is a silly idea.
 PRO_{one} = them

	one		PRO _{one}	
	SINGULAR	PLURAL	SINGULAR	PLURAL
FIRST	*me	*us	me	us
SECOND	*you	*you	you	*you
THIRD	him her one	*them	*him *her one	*them

Figure 23. Objective Personal Pronouns Licensed by *one* and PRO_{one}

The objective personal pronouns licensed by *one* as in (20), and by PRO_{one}, as in (21), are summarized in Figure 23 on page 215.

There is a considerable literature concerning the history of *one* and its relation to the definite and indefinite pronouns. According to Curme, *one* can be the antecedent of *oneself*, *one's self*, *one's own self* and *himself*:

Indefinite *one* is also used as an absolute indefinite pronoun, i.e., without reference to a noun that has been previously mentioned, usually with a genitive in *-s*: 'He died in 1859, leaving his property to *one* Ann Duncan' (i.e., to *a certain one*, namely, Ann Duncan, the name standing in apposition with *one*). 'He looked like *one* [who was] dead.' 'One must do *one's* duty.' The reflexive form is either *one's self* or *oneself*, the former, the older form, after the analogy of *a man's better self*, *one's own self*, *myself* (70D, last and next to last par.), the latter, the newer form, after the analogy of *himself* (47D, next to last par.), both forms in common use in America, while in England the newer form is probably more common than the older: 'One should not praise *one's self*' (or *oneself*). Both *one's self* and *oneself*, however, are comparatively recent formations, which do not occur in Shakespeare. In older English, *himself* was used here.

The nominative, genitive, dative, accusative corresponding to indefinite *one* in present-day English are , *one's*, *one*, *one*, but the older forms *he*, *his*, *him*, *him* still linger on: 'One never realizes *one's* blessings while *one* enjoys them.' 'One hates *one's* enemies and loves *one's* friends.' 'In life

one only notices what interests *one*' (but in Galsworthy's *Patrician*, p. 48, we find here *him* as in older English). 'Vulgar habit that is people have nowadays of asking *one*, after *one* has given them an idea, whether *one* is serious or not' (Oscar Wilde, *A Woman of No Importance*, Act I). But *he*, *his*, *him* correspond to the numeral *one*, to *one...another*, and to *no one*, *someone*, *everyone*, *anyone*: 'One of these men hates *his* enemies.' 'One hates *his* enemies and another forgives *his*.' 'If *someone* (or *anyone*) should lose *his* purse, *he* should apply to the Lost Property Office.' Similarly, the reflexive object *one's self*, or *oneself*, corresponds to the indefinite subject *one*, while *himself* corresponds to the numeral *one*, to *one... another*, and to *no one*, *someone*, *everyone*, *anyone*: 'One cannot interest *one's self* in everything,' but *One* of the boys fell and hurt *himself*' and *No one* can interest *himself* in everything.' We often, however, hear *himself* instead of *one's self* or *oneself*, as in older English: 'One might fall and hurt *himself*.' In careless expression we often find here even in good authors a plural form corresponding to *one*: 'As though *one* went to tea with a woman for the sake of talking about the very things *you* (= *one*) have been doing all day' (Mrs. Ward, *Sir George Tressady*, I, Ch. V). 'One must be on *their* (= *one's*) guard against bargains that are worthless' (Rev. E.J. Hardy, *How to Be Happy Though Married*, Ch. XII, 128) (Curme: 531-2)

So far I have covered the person, number, and gender of objective, possessive, and reflexive pronouns which can have *one* or PRO_{one} as antecedents. Let us now turn to the person, number, and gender of *one* and PRO_{one} independent of any other pronoun in the sentence.

PRO_{one} can be plural, as in (22), but *one* cannot be, as in (23).

- (22) i. It would be inappropriate [PRO_{one} to leave together]
 ii. It would be inappropriate [PRO_{one} to meet in New York]
 iii. It would be inappropriate [PRO_{one} to kiss each other on the cheek]
 iv. It is necessary PRO_{one} to look alike [PRO_{one} to be cast in the play]]
- (23) i. * It would be inappropriate [for one to leave together]
 ii. * It would be inappropriate [for one to kiss each other on the cheek]
 iii. It is usual [for one to send flowers]

6.2.2 THE REFERENCE PROPERTIES OF OVERT 'ONE' AND EMPTY 'ONE'

One is like *you* in that there is only a single referent for the pronouns *you* and *one* in the sentences in (24), but in (25) each occurrence of *he* could be some new referent, as in (26)

- (24) i. If one commits a crime, and one tries to escape, one cannot escape one's conscience.
- ii. If you commit a crime, and you try to escape, you cannot escape your conscience.
- (25) If he commits a crime, and if he tries to escape, he cannot escape his conscience.
- (26) If Tom commits a crime, and if Dick tries to escape, Harry cannot escape his conscience.

In sentences where every subject is PRO_{one}, there is understood to be only a single referent:

- (27) To eat, to drink, and to be merry is a good life.
- (28) PRO_{one} to eat, PRO_{one} to drink, and PRO_{one} to be merry is a good life.

There is no reason to indicate the coreference of *one* and PRO_{one} with subscripts in (24i) and (28) since they are like *I*, *you* and *we* ... and not in the class of definite pronouns *he*, *she*, and *they*.

I will assume that PRO_{one} is like *one*, as described in (15), and that PRO_{one} can be first person singular and plural, second person singular, or third person singular. This will describe the facts concerning agreement of pronouns cited above. However, there seem to be some constraints on semantic interpretation.

Putting aside questions of well-formedness,¹⁷ sentences like the following have a single understood referent, the addressee (*you*) in (29) and the speaker (*I*) in (30).

- (29) PRO_{one} to eat yourself silly and then PRO_{one} to drink one's self into a coma is fun.
- (30) PRO_{one} to indulge myself on earth and then PRO_{one} to have one's sins forgiven after death sounds like a fair deal.

Sentences like (31) should have two referents understood, a first person and a second person. Since (31) seems much worse than (29) and (30), it is probably the case that PRO_{one} can be either the speaker or the addressee, but not both in the same sentence.

- (31) PRO_{one} to eat yourself silly and PRO_{one} to drink myself into a coma is fun.

One last problem concerns the reference of PRO_{one} when it occurs with second person singular pronouns. In the sentences discussed earlier, (17ii) and (19ii), although the pronouns are second person singular, and *the warden* is the person being addressed, the second person pronouns are not understood to be *the warden*.

- (17) ii. Dear Warden, it is unclear [how [PRO_{one} to keep yourself occupied here]]
- (19) ii. Dear Warden, it is unclear [how [PRO_{one} to keep your self-respect here]]

PRO_{one} seems to be like the indefinite *you* in the following sentences:

¹⁷ Sentences with *one* and PRO_{one} can be very long, and in these long sentences, person changes are not infrequent. Consider the example cited above by Curme:

As though *one* went to tea with a woman for the sake of talking about the very things *you* (=one) have been doing all day. (Mrs Ward, *Sir George Tressady*, I, Ch. V) (Curme: 532)

More examples can be found in Curme, Jespersen, and the *OED*.

(32) It is difficult to look farther ahead than you can see.

(33) You never know who is your friend until you need one.

6.2.3 THE FEATURE [HUMAN]

Jespersen (II.10.14) claims that *one* when used coreferentially, it may be animate or inanimate, but when used independently, *one* must be animate. Jespersen uses the word *anaphoric* where I use *coreferential*.

10.14. It will be necessary to have terms for a distinction that is important in this as well as in some of the following chapters. I propose to apply the word *anaphoric* to *one* (or any other word) if it refers to some word already mentioned, while I say *independent* if there is no such contextual reference. *The little one* is used anaphorically if it means 'the little flower' or whatever it is that has just been mentioned, but independently if it means 'the child' without some such substantive being already named. The independent *one* always denotes a person, while the anaphoric *one* may refer to a thing.

There is a play on the anaphoric and independent employments in the wish addressed to newly married people: May all your future troubles be *little ones*.
(Jespersen, II:247-8)

According to Jespersen, *one* can be understood to be referring to a previously mentioned NP, or *one* can be independent. When coreferential with another NP, *one* can be [+human]. When *one* is independent, *one* must be [+human].

(34) An old rusty one is better than a new one.

(35) I see one on the floor.

(36) One fell and broke.

One in (34)-(36) refers to a noun previously mentioned in the discourse that is understood by the speaker and hearer. Suppose we were looking over a yard full of chickens and said (37)-(38). According to Jespersen, *one* in (37) has *the chickens* as an antecedent or referent but *one* in (38) has no antecedent or outside referent and would refer to a person.

- (37) One tastes as good as the next.
- (38) One is better off dead than scratching, pecking, and eating worms.

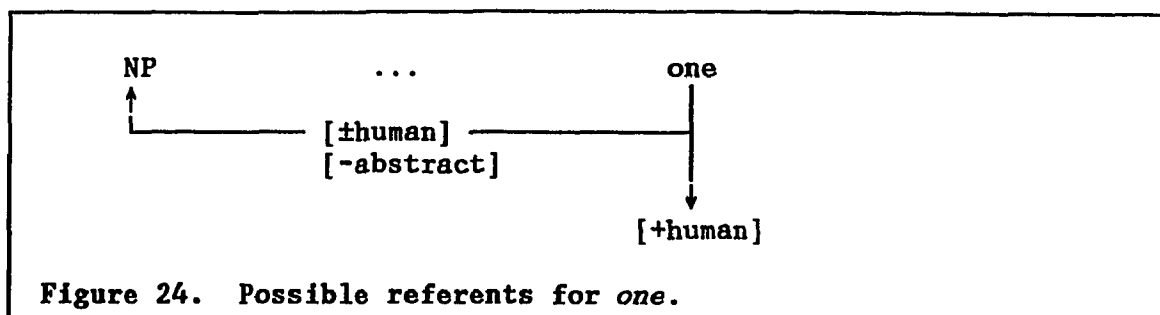
According to Curme, an indefinite *one* cannot refer back to an abstract noun or any other noun that does not denote an individual, person, or thing, and sometimes *one* alternates with *such*. Curme states:

In the substantive relation, indefinite reference is in the singular expressed by *one* when it is desired to refer back to a noun that has just been mentioned, but *such* is sometimes still as in older usage employed here: 'He is a friend, and I treat him as *one*.' 'Two or three low broad steps led to a platform in front of the altar, or what resembled *such*' (Scott, *Aunt Margaret's Mirror*, II), now more commonly *one*. An indefinite *one* cannot refer back to an abstract noun or any other noun that does not denote an individual, person, or thing, it is here usually replaced by *it*, *this* or *such*: 'I should like to find other examples if they are to be had.' 'To call for more facts and experiments, if *such* are possible' (Geike in *Nature*, Sept. 19, 1889). (Curme: 530)

Under *abstract nouns* Curme would include gerunds, such as *John's singing*. Hence, (39ii) is ill-formed because *one* cannot refer back to an abstract noun.

- (39) i. John's singing is so bad that hearing it drives me crazy.
 ii. * John's singing is so bad that hearing one drives me crazy.
- (40) i. John's song is so bad that hearing it drives me crazy.
 ii. John's songs are so bad that hearing one drives me crazy.

The precise properties of *one* is not within the scope of this thesis. Figure 24 on page 221 summarizes Jespersen's and Curme's analyses:



Unlike *one*, *oneself* and *one's* are always human. They are understood to be a participant in the discourse: the speaker, the addressee, or an indefinite human.

- (41) i. It is a psychological necessity PRO_{one} to relieve myself of my tensions.
 ii. It is a psychological necessity PRO_{one} to relieve yourself of your tensions.
 iii. It is a psychological necessity PRO_{one} to relieve oneself of one's tensions.
 iv. It is a psychological necessity PRO_{one} to relieve ourselves of our tensions.
- (42) To relieve oneself of one's tensions is a physiological necessity.
- (43) i. * It is a geological necessity PRO_{one} to relieve oneself of one's pressure.
 ii. * PRO_{one} to relieve oneself of pumice and magma is a volcanic necessity.

In (44), it sounds as though people must wear horseshoes. Even in the context of a discussion on horses and horseshoes, *one's* in (45) would not normally be understood as referring to a horse.

- (44) Shoes must be worn when walking on stones PRO_{one} to avoid cutting one's feet.
- (45) Horseshoes must be worn when walking on stones PRO_{one} to avoid cutting one's feet.

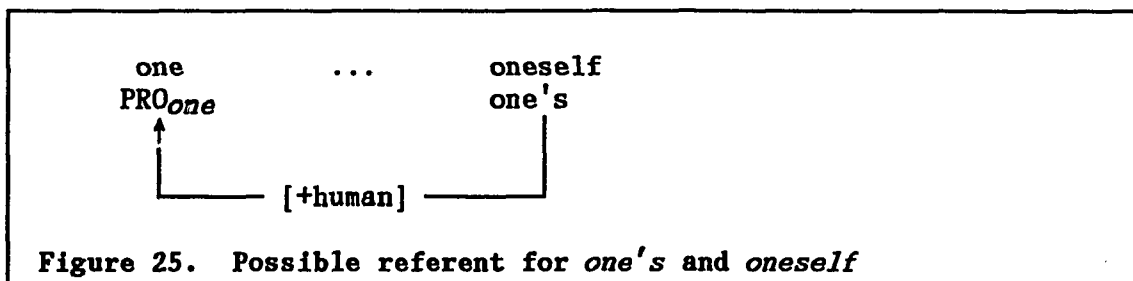
Notice that sentences which would link *oneself* or *one's* with a non-human, as in (43), (47), (49) and (50), antecedent are deviant.

- (46) People should wear shoes when walking on stones to avoid cutting one's feet.
- (47) * Horses should be careful when walking on stones to avoid breaking one's leg.
- (48) Horses should be careful when walking on stones to avoid slipping.
- (49) * To rid oneself of pumice and magma, Mt. St. Helens erupts regularly.
- (50) * It is a fact that to eject one's dust and fumes is the main function of a volcano.

The sentences are much better if *oneself* and *one's* are replaced by *itself* and *its* respectively.

- (51) i. To rid itself of pumice and magma, Mt. St. Helens erupts regularly.
- ii. It is necessary for Mt. St. Helens to erupt to get rid of its gas.
- iii. It is a fact that to eject dust and fumes is the main function of a volcano.

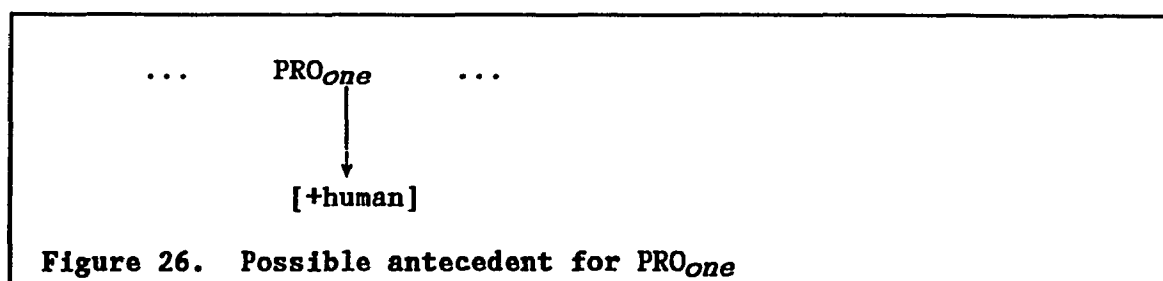
The only possible antecedents for *one's* and *oneself* are *one* and PRO_{one} , and further, the antecedent must be [+human]. This is summarized in Figure 25 on page 222.



Consider next, sentences like (52)-(54), which contain no possible antecedent for PRO_{one} .

- (52) [PRO_{one} to float in turbulent water] is difficult.
- (53) It is hard to imagine [how [PRO_{one} to get out of this situation]]
- (54) [Whether [PRO_{one} to eat it or not]] is a difficult question.

In sentences which contain no possible antecedent for PRO_{one}, PRO_{one} is always understood as [+human]. This is summarized in Figure 26 on page 223.



6.2.4 THE HISTORY OF THE OVERT PRONOUN 'ONE' AND EMPTY 'ONE'

A historical analysis of the overt indefinite pronoun *one* might indicate that the empty indefinite pronoun, *eone*, was a part of English syntax before the form *one* came into use. *One* might best be considered an expletive element like *it* or *there*. Jespersen gives the following analysis for the development of *one* in English:

10.12. The decisive innovation was the rise of combinations like *a good one*. This, I think, is to some extent like the use of *it* in *it rains* or in *I think it necessary to wait*, and like the use of *there* in *there was peace* or *let there be peace*: in all these cases a word becomes necessary because speakers are accustomed to have some word in that particular place: *it* and *there* take the place of a subject or of an object, and similarly *one* takes the place usually occupied by a substantive. People were so accustomed to say *a good man*, *a great house*, etc. that they felt *a good* and *a great* as rather bald expressions: the ordinary balance was restored by saying *a good one*, *a great one*. Or, we might perhaps rather say: the balance that was felt wanting in both ME constructions *an (one, an, a) great* and

great an (one), was restored by a blending of both. Corresponding developments are found in Danish (*en god een, et godt et; sadan een; hvilken*, originally *hvilik + een* 'which one; now also *hvilken een*; even *sikken en een < se hvilik en en een*, literally 'see which an a one') and in Cape-Dutch (*das 'n mooie ene* 'that's a nice one', *di ander een* 'the other one', H. Meyer, *Die Sprache der Buren* 1901, p. 40). But when once the use of *one* as a prop-word (to use Sweet's term) had been established, English went further than Danish at any rate, though some dialects (in Jutland) go further than standard Danish and admit, for instance, *den ækle en* 'that abominable one' and *din grimme en* 'you (literally your) hideous one' (both found in Bregendahl, *Dodsnat* 95). The chief difference between English and Danish in this respect is the use of the plural *ones*, which to some extent resembles Spanish *unos*, colloquial Danish has a repetition of *nogle* or *nogen*: *nogen gode nogen* 'some good some', and correspondingly in the neuter singular *noget godt noget* 'something good'.
(Jespersen, II:246-247)

The diachronic development of *one* in a construction like (55) is interesting since both the indefinite determiner *a/an* and the nominal *one* derive from the numeral *one*.

(55) [NP a big one]

Thus *a big one* has as a diachronic precursor: *one big one*. Constructions with *one* as the subject of an infinitive seem relatively recent. According to Jespersen, since the 19th century, the element *one* can occur as the subject of an infinitive:

10.3(appendix) We may perhaps give the following tentative chronology, the dates of course to be taken as approximative:

1300 *a good one*
 1400 *the good one*
 1500 *never a one, such a one, good ones, the good ones*
 1600 *one good one*
 18th c. *that one, a silver one*
 19th c. *the one (we) preferred, those ones, a one to keep company, the ones that..., my one*

(Jespersen, II:502)

In modern non-idiomatic usage, *one* as an infinitive subject is almost always animate with no reference in the sentence. Jespersen (18.25)

does give an example of inanimate *one* with a reference in the sentence (56):

(56) Consider the matter as one to be arranged between you and me.

However, the vast majority of sentences have human subjects. The following sentences are shortened examples of 19th century usage drawn from Jespersen, Curme, and the *OED*.

(57) Always one to suspect the worst, John was considered paranoid.

(58) When a marriage dissolves suddenly, it leaves one or both to weep.

(59) He entered the room with the air of one about to communicate a novel idea.

(60) John is not one to originate new ideas.

(61) John is not one to choose to play King Lear/to cast as King Lear.

(62) John is not the one to choose to play King Lear/to cast as King Lear.

(63) We are the ones to do the job.

My analysis is compatible with the view that the empty category *eone* first occurred in constructions like *a big eone*. As Jespersen says, *one* was introduced like the use of *it* in *it rains* and like the use of *there* in *let there be peace*. In the 19th century, *one* was used as the subject of an infinitive in sentence like (56)-(62). Since the 19th century, *eone* has been used as the subject of an infinitive. One can expect dialect variation in a new construction, as the use of *one* and *eone* with the infinitive seems to be.

6.3 THE INTERPRETATION OF PRO

The range of possible interpretations for PRO cannot be predicted by any of the Control theories discussed in the last chapter. In this

section, I propose two extrinsic features ([±definite] and [±generic]) which are assigned by the parser. Both lexical and structural factors contribute to licensing the parser to assign the extrinsic features.

Since I propose two extrinsically assigned features, the theory predicts that PRO has four distinct interpretations, each with its own predictable properties. The literature reviewed in the last chapter allows for two interpretations of PRO. The first is where PRO is unambiguously coindexed with a c-commanding NP. This has been called obligatory control PRO or governed PRO. The second is all other interpretations. This has variously been called PRO_{arb}, non-obligatory control PRO, or ungoverned PRO. One of the feature bundles roughly corresponds to obligatory control or governed PRO. I say "roughly corresponds" because my analysis diverges from the standard analysis which requires the controller to be a c-commanding NP. As I have argued, the controller can be in an object PP. The remaining three feature bundles are traditionally lumped together as PRO_{arb}. The PRO_{arb} classification is impoverished in that by merging all instances of PRO_{arb}, interpretations which are predictable appear to be unpredictable.

6.3.1 THE INTERPRETATION OF THE FEATURES: GENERIC AND DEFINITE

I call the extrinsic features [generic] and [definite]. The generic feature indicates that reference is outside of the adjacent sentence. The definite feature indicates that PRO requires an antecedent which is a single node in the phrase marker. All instantiations of PRO are assigned values for both features [generic] and [definite]. However, I will first introduce the features and their interpretations individ-

ually, before presenting the anaphoric and pronominal properties that their combinations induce.

In describing the effects of these features, I refer to PRO "including an antecedent but not being identical to it." Williams discusses this property with respect to *want*:

We find a similar phenomenon with *want* (Debbie Nanni, personal communication): the antecedent need not be exactly the subject--it can be some group which merely includes the subject. The verb *meet* (as opposed to *meet with*) normally takes a plural subject, but in the following example, the antecedent of the PRO subject of *want* is singular:

<68> I want to meet at 6.

(Williams 1980:218)

In <68>, PRO includes the antecedent *I* as well as some other person or people.

PRO with the feature [+generic] is interpreted generically in that it is like the indefinite *you* or *one*. Although [+generic] PRO may have an antecedent NP in the sentence, it is understood to include the antecedent but not necessarily to be identical to it. A [+generic] PRO can license (i.e. be the antecedent of) *oneself* even if the PRO has an antecedent that is not the pronoun *one*.

PRO with the feature [-generic] is interpreted as having at least one specific NP antecedent in the sentence. The antecedent may be split. However, [-generic] PRO is not interpreted as including reference to anyone or anything not explicitly represented as an NP in the sentence. A [-generic] PRO cannot be the antecedent of *oneself* unless the antecedent of PRO is an overt *one*.

PRO with the feature [+definite] must have one explicitly mentioned NP of an appropriate person, number, and gender in the sentence as an antecedent. A [+definite] PRO may be identical to the antecedent NP or may simply include the antecedent NP.

PRO with the feature [-definite] does not have one single explicit NP as its antecedent. A [-definite] PRO is indefinite in that there may be no antecedent NP in the sentence or the antecedent may be split in reference between two NPs.

6.3.2 THE INTERPRETATION OF THE FEATURE BUNDLES

Below, I illustrate the various possible combinations of [\pm generic, \pm definite] by appealing to examples. These proposed features are not an exhaustive classification of features in the lexical entry for any particular verb or verb class. They are, however, a complete characterization of the range of interpretations of PRO. See Figure 27 on page 229 for a chart summarizing the possible feature combinations. Any of the verbs used to illustrate the various interpretations of PRO may also license other infinitive complements beyond those listed.

PRO [-generic,+definite] or Exclusive-PRO

This is the feature complex of indicative complements of verbs. PRO[-g,+d] must have an antecedent NP in the sentence. This antecedent must be a single NP. *Persuade* is lexically marked to require a PRO[-g,+d] in the complement clause:

(64) John persuaded Mary [PRO[-g,+d] to go]

Further examples of verbs which require PRO[-g,+d] in their complements are given in (65)-(67):

	+generic	-generic
+definite	Inclusive-PRO <i>wonder</i>	Exclusive-PRO <i>persuade</i> <i>promise</i>
-definite	PRO <i>one</i> <i>say</i>	Split-PRO <i>propose</i>

Figure 27. Extrinsic Features assigned to PRO

- (65) Mary promised John [PRO to go]
 (66) Mary swore to John [PRO to stop drinking]
 (67) John shouted to Mary [PRO to stop drinking]

In each of these sentences, PRO is understood to be unambiguously coreferential with the single NP node *Mary*.

PRO_[-g,+d] is the most anaphoric of the four possibilities for PRO. This feature bundle roughly corresponds to what has been variously called obligatory control PRO and governed PRO. Like an anaphor, it has an obligatory antecedent which is a single node. PRO_[-g,+d] is pronominal, however, in that it takes a non-c-commanding antecedent when PRO is in a complement clause, as in (67) and (68):

- (68) i. John motioned to/at Mary to sit down.
 ii. The general signaled to his officers to begin the attack.
 iii. John beckoned to the child to come closer.
 iv. John cabled/telephoned/wrote to Mary to come home.

Since PRO_[-g,+d] is interpreted as being identical to its antecedent, and excludes anyone or anything else from its scope, I call it **exclusive-PRO**.

PRO [+generic,+definite] or Inclusive-PRO

This is typically the feature complex of *wh*-, or indirect question, complements of verbs. Being definite, PRO_[+g,+d] must have at least

one specific NP antecedent in the sentence. Being generic, PRO_[+g,+d] may be understood to include the antecedent but not be identical to it. Since it is generic, PRO_[+g,+d] can license *oneself*, even if the pronoun *one* is not the antecedent. Since PRO_[+g,+d] can be interpreted as including other people or things as well as the antecedent in its scope, I will call it inclusive-PRO.

Verbs like *wonder* license inclusive-PRO. When the adjective *unclear* has a PP complement, it assigns the features of inclusive-PRO

- (69) i. John_i wondered [how [PRO_[+g,+d] to behave himself_i]]
 ii. John_i wondered [how [PRO_[+g,+d] to behave oneself_i]]
- (70) i. It is unclear to John [when [PRO to leave]]
 ii. It is unclear to John [how [PRO to behave oneself in public]]

In each of these cases, PRO is understood to be coreferential with the single NP node *John*. PRO does not have to be identical with the NP node, but must include it. In sentences (69ii) and (70ii), although the reflexive *oneself* is licensed by the feature [+generic] on PRO, PRO is understood to include *John* in its reference.

Inclusive-PRO is anaphoric in that it must be bound by a single node of the appropriate person, number, and gender. Compare (72i) and (72ii):

- (72) i. John wondered how to behave himself.
 ii. * John wondered how to behave herself.

Inclusive-PRO is pronominal in that the antecedent need not be c-commanding when PRO is in a complement clause, as in (70), and the antecedent may be arbitrarily far away. I will discuss PRO_[+g,+d] and long-distance control in "Long-Distance Control" on page 249.

PRO [-generic,-definite] or Split-PRO

PRO[-g,-d] is the feature complex of indicative complements of verbs which have split antecedents. Being [-generic], PRO[-g,-d] must have at least one antecedent NP explicitly marked in the sentence. Being [-definite], the antecedent is not a single NP. Verbs that are lexically marked to permit split antecedents assign PRO[-g,-d] in the infinitive complement clause. Since this feature bundle licenses split antecedents, I will call PRO with the features [-g,-d] split-PRO. Examples of verbs which license split-PRO are given in (73)-(74).

(73) John_i proposed to Mary_j [PRO[-g,-d]_{i,j} to go]

(74) John_i agreed with Mary_j [PRO_{i,j} to go]

(75) John_i planned with Mary_j [PRO_{i,j} to go in one car]

In these cases, PRO is understood to include two NPs in its reference. Not uncommonly, the matrix sentence includes a *with-* commitative prepositional clause.

Split PRO is anaphoric in that it requires an antecedent. Split PRO is pronominal in that the antecedent is split. However, the antecedents must both be in the matrix sentence.

PRO[+generic,-definite] or PRO_{one}

This is the feature bundle that corresponds to PRO_{one} discussed in "The Pronominal Anaphor 'One'" on page 209. PRO[+g,-d] is generic in that it does not refer to any antecedent NP in the sentence, but it may include reference to the speaker (first person) or the hearer (second person). The complement of the intransitive *say* licenses PRO[+g,-d]. Being generic, the PRO[+g,-d], or PRO_{one} can license the pronoun *oneself*.

- (76) i. John said [PRO[+g,-d] to behave]
 ii. John said [PRO[+g,-d] to behave myself]
 iii. John said [PRO[+g,-d] to behave yourself]
 iv. * John said [PRO[+g,-d]_i to behave himself_i]
 v. John said [PRO[+g,-d] to behave oneself]

In addition, adjectives which do not assign an external θ -role license PRO_{one}.

- (77) It is interesting [PRO[+g,-d] to read old books]

PRO[+g,-d], or PRO_{one}, is always [+animate] and appears to be always [+human]. Being [+generic] it can license *oneself*.

PRO_{one} is the most pronoun-like of the four possibilities for PRO. Like a pronoun, it does not require an antecedent. However, it does act like an anaphor when there is more than a single PRO_{one} in the string. As was discussed in "The Reference Properties of Overt 'One' and Empty 'One'" on page 217, in strings like (27), repeated below, there is a single referent for PRO.

- (27) To eat, to drink, and to be merry is a good life.

The above discussion is intended as a heuristic to illustrate the full range of interpretations of PRO -- as opposed to the traditional bifurcation of PRO into PRO_{index} and PRO_{arb}. The actual content of the features of PRO can only be determined by the role the features play in the theory.

7.0 LICENSING MECHANISMS FOR INFINITIVE CONSTRUCTIONS

The distribution of PRO follows from principles of Universal Grammar. More specifically, the distribution of PRO is a result of binding theory. By the PRO Theorem, since PRO is [+pronominal] and [+anaphoric], it can only occur in ungoverned locations. This is a non-parameterized principle of UG. English particular, or parameterized, principles of UG also play a role in determining the distribution of PRO. The fact that each sentence of English requires a subject guarantees that PRO will be the subject of an infinitive.

The features assigned to PRO, [\pm generic, \pm definite] and, if [+definite], which NP is the antecedent of PRO, are determined by three factors:

- (a) Extrinsic features and the referent are assigned to PRO by lexical formatives. Lexical formatives are open class lexical items like verbs, adjectives and nouns.
- (b) Extrinsic features and the referent are assigned to PRO by grammatical formatives. Grammatical formatives are closed class lexical items like *wh*- elements, comparative adverbials (like *too* and *enough*), and the passive marker *be+en*.
- (c) Extrinsic features and the referent are assigned to PRO in specific structural configurations, such as to adjuncts.

I propose a parsing strategy for infinitive constructions in (1).

- (1) General strategy for the grammar of infinitives:
 - a. Tree Building Mechanisms: First, the grammar constructs the tree containing the infinitive(s). The tree building

mechanisms which assign a phrase marker to a string utilize the lexical entry to build a phrase marker, but they do not utilize any information about the features or reference properties of PRO. A phrase marker is assigned to the string, but there are no extrinsic features on PRO and PRO is not coindexed. The tree building mechanism makes use of X-bar theory, binding theory, case theory, theta-theory, and the extended projection principle. The phrase marker which is the output of the tree building mechanism includes binding relations to indicate the reference of anaphors and non-reference of pronouns.

- b. **Control Module:** The control module indicates the reference for control of PRO. The PRO interpretation mechanisms which assign extrinsic features to PRO and which indicate control by coindexation operate on the fully constructed tree.

In this and the following chapters, I motivate four basic PRO interpretation/reference mechanisms. The PRO interpretation mechanisms (1) indicate any coindexing with PRO, and (2) assign extrinsic features to PRO. These mechanisms operate on the fully constructed tree produced by the tree building mechanisms. There are four PRO interpretation and reference mechanisms and they apply in order. The properties of PRO in a string are the features and reference assigned by the last mechanism which applies. Later mechanisms can overwrite or change features of earlier mechanisms. The four mechanisms for interpreting PRO are listed below:

1. The *wh*-PRO_{one} licensing mechanism, which is discussed in "Feature Assignment by Grammatical Formatives" on page 254.
2. The lexical entry of the matrix verb assigns features and indicates reference, as discussed in "Lexically Specified Anaphora" on page 237. The V1-V2 lexical mechanism for double-passive constructions is motivated in "The Domain of Licensing PRO" on page 338.
3. The long-distance PRO mechanism, which can rewrite features, is discussed in "Long-Distance Control" on page 249.
4. The adjunct mechanism is presented in "The Constituent Structure of Adjuncts" on page 261.

The assumptions that the mechanisms apply in order and that later mechanisms can overwrite the features and reference of earlier mechanisms is not fundamental to these proposals. Rather, this ordering relation is imposed by, and reflects, the limitation on theoretical machinery to implement the processes required for the grammar of PRO.

I assume that the processes apply in sequence only to make the proposals more intelligible. One can imagine a production line in which a string is loaded onto the conveyor belt at one end and a completed phrase marker with the features and reference of PRO is unloaded at the other. Along the line stand various workers, each ready to assign some structure to the string. As the string moves down the line, each worker is presented with the structure that has been added by previous workers who can no longer have any effect. This suggests that the assignment of structure to a string is a basically a set of sequential operations, each of which changes the state of the structural description of the string to advance it from what it was. A grammar in which one process

must apply before another, and once having applied can never apply again, is a *sequential processing grammar*. It is entirely possible that this model of sequential grammatical operations applied to a string is incorrect.

Quite possibly we should think of a string as lying on a table and surrounded by workers. Each worker recognizes certain structural configurations and patterns, and in response to a recognized pattern, performs some alteration on the structure assigned to the string. Each worker can add to or change any aspect of the tree upon recognizing some structural configuration. Depending on the string being analyzed, the workers may work simultaneously or in any sequence in order to complete the analysis of the string. A grammar in which any process can apply at any time, and any two processes can apply in any order, will be called a *parallel processing grammar*.

This proposal is neutral between a sequential processing grammar and a parallel processing grammar. Even though the basic grammatical processes of control are presented as though the mechanisms were implemented in a sequential processing grammar, it is quite likely that the grammatical processes would more naturally be implemented as mechanisms in a parallel processing grammar.

7.1 LEXICAL SPECIFICATION OF FEATURES AND ANTECEDENT

It is uncontroversial that for any particular verb, V, its lexical entry indicates (a) structure licensed by that lexical item: e.g., transitive, intransitive, etc.; and (b) the type of complement licensed by V: e.g., [$\pm wh$ -] complement, infinitive, subjunctive, etc. In addition, I argue that the lexicon specifies (c) the referent of PRO: e.g., ex-

ternal argument or internal argument; and (d) the features of PRO: e.g., [\pm generic] and [\pm definite].

7.1.1 LEXICALLY SPECIFIED ANAPHORA

The lexical entry for a word not only contains information about possible complement structures of the word, it also contains information about obligatory anaphoric relations among the elements of the complements.

Examples of intrasentential anaphors are given in (2) and (3).

- (2) i. John perjured himself.
 ii. * John perjured herself/oneself/yourself/ourselves.
- (3) i. John bet his bottom dollar.
 ii. * John bet her/one's/your bottom dollar.

In (2) and (3), the lexically specified anaphora takes the subject of the predicate as its antecedent. Other verbs which lexically specify for anaphor and take subject antecedents include the following: NP *behave* (pro-self), NP *avail* pro-self of NP, NP *content* pro-self with NP, NP *pride* pro-self on NP, NP *make* NP pro-self at home, NP *enjoy* pro-self, NP *pull* pro-self together.

Transitive verbs which lexically specify for anaphors sometimes require the subject to be the antecedent, as in (4), and sometimes assign the object as antecedent, as in (5) and (6):

- (4) i. John gave Mary his word.
 ii. * John gave Mary her/one's/your word.
- (5) i. John took Mary at her word.
 ii. * John took Mary at his/one's/your word.
- (6) i. John sent Mary on her way.
 ii. * John sent Mary on his/one's/your way.

This same process occurs intersententially. Examples of lexically specified anaphora which occur across sentence boundaries are in (7)

and (8), where the subject of the complement must be referential with the noun phrase in the matrix subject.

- (7) i. John confessed to Mary [that [he did it]]
 ii. * John confessed to Mary [that [she/you did it]]
- (8) i. John professed to Mary [that [he believes in God]]
 ii. * John professed to Mary [that [she/you believes in God]]

The case of intersentential lexically specified anaphora is known as Control. When a verb lexically specifies for anaphora in its complement, and the complement is an infinitive clause, then the verb specifies the antecedent of PRO, as with *promise* and *persuade* in (9) and (10).

- (9) i. Mary promised John [PRO to perjure herself]
 ii. * Mary promised John [PRO to perjure himself/yourself/oneself]
- (10) i. Mary persuaded John [PRO to perjure himself]
 ii. * Mary persuaded John [PRO to perjure herself/yourself/oneself]

When control is lexically assigned, it is obligatory. The obligatoriness of control in complement clauses is illustrated below where the sentence is ill-formed when PRO and the subject are disjoint in reference, as in (11ii) and (12ii).

- (11) i. John_i wanted [PRO_i to leave]
 ii. * John_i wanted [PRO_j to leave]
- (12) i. John_i asked [what [PRO_i to wear]]
 ii. * John_i asked [what [PRO_j to wear]]

The antecedent cannot freely be in either subject or object position. For any given lexical entry, the antecedent either must be subject or must be object or must be split. For example, *ask* assigns object control when it has a propositional complement, as in (13), and subject

control when it has a *wh*- complement, as in (15); *tell* always has object control, as illustrated by (14) and (16).

- (13) i. John_i asked Mary_j [PRO_j to leave]
 ii. * John_i asked Mary_j [PRO_i to leave]
- (14) i. John_i told Mary_j [PRO_j to leave]
 ii. * John_i told Mary_j [PRO_i to leave]
- (15) i. * John_i asked Mary_j [what [PRO_j to wear]]
 ii. John_i asked Mary_j [what [PRO_i to wear]]
- (16) i. John_i told Mary_j [what [PRO_j to wear]]
 ii. * John_i told Mary_j [what [PRO_i to wear]]

Some verbs, such as *propose*, allow split antecedents in that they assign control from both subject and object. However, when infinitives in complement positions do allow split antecedents, the antecedents must both be within the argument structure of the verb, as in (17). Antecedents cannot be in another sentence, as in (18).

(17) John proposed to Mary to go to the movies together.

(18) * John thought that Mary wanted to go to the movies together.

When control is lexically specified, the controller need not c-command PRO. All of the analyses I have discussed assume that a possible controller must c-command PRO. I have illustrated in "C-command" on page 199, that this is not the case. The antecedent of PRO can be in an object PP. An additional claim of Chomsky's that I take issue with below is that PRO in indirect questions may be bound or free.

Chomsky's (1986a) argument that control in *wh*- complements is subject to c-command and can have an arbitrary interpretation is based on (19).

- (19) i. John's mother asked me [how [PRO to feed herself]]
 ii. * John's mother asked me [how [PRO to feed himself]]
 iii. John's mother asked me [how [PRO to feed oneself]]

When the lexicon assigns control to the external argument, as it does in *ask* when it takes a *wh*- complement, the subject is the controller. In (19), the subject, and therefore the controller, is *John's mother* and not *John*. Thus (19ii) is bad not because *John* does not c-command PRO but because *John* is not the external argument.

All of the theories I have discussed assume that if PRO can be indexed with *oneself*, then it can optionally have an arbitrary interpretation. This reflexive test for arbitrary interpretation is inadequate. As we saw in "The Person, Number, and Gender of 'One' and PROone" on page 210, the interpretation for the indefinite pronoun *oneself* typically includes the speaker or writer. PRO and *oneself* in (19iii) are coindexed with *John's mother*. That is, PRO in (19iii) is not free.

Chomsky further claims that in (20) PRO "may be bound or free":

(20) They asked how to rig the boat.

For those who can get an arbitrary interpretation for (20), this interpretation results from the indefiniteness of *they*. If *they* in (20) is replaced by *Mary*, as in (21), *Mary* must be included in the reference for PRO.

(21) Mary asked how to rig the boat.

It is true that, unlike declarative complements, interrogative complements license the inclusion of *oneself*. Thus we get (19iii) and (22), but not (23).

(22) John asked Mary [how [PRO to behave oneself]]

(23) * John promised Mary [PRO to behave oneself]

However, Chomsky, Manzini and others assume that the inclusion of *one-self* in (19iii) and (22) indicates that PRO is arbitrary in reference.

This is not so. Consider the following sentences:

- (24) i. * John asked Mary how to behave myself.
 ii. * John asked Mary how to behave yourself.
 iii. John asked Mary how to behave oneself.
 iv. John asked Mary how to behave himself.
 v. * John asked Mary how to behave herself.
 vi. * John asked Mary how to behave ourselves.
 vii. * John asked Mary how to behave yourselves.
 viii. * John asked Mary how to behave themselves.

The only reflexives allowed in this environment are *himself* and *oneself*. That is, either the gender, person, and number features of the reflexive must agree with the subject of *ask*, *John* and *himself* in (24iv) or they must agree in person and number and have a non-clashing gender feature, *John* and *oneself* in (24iii). Interrogative complements have a generic interpretation when they contain the indefinite reflexive pronoun *oneself*.

Since control in complements is lexically specified, an NP is a potential controller only when it is licensed by the lexical structure. Thus controllers must have θ -roles. If a verb or adjective does not assign an external θ -role, then the subject cannot be a controller. Thus only predicate adjectives which assign an external θ -role can assign control to a complement clause using this strategy.

To summarize, PRO in all clausal complements is pronoun-like in that it does not have to be c-commanded by its antecedent. It is anaphor-like in that its antecedent cannot freely be subject or object; if a controller is assigned, PRO cannot be free; its antecedent is within the argument structure of the verb. As far as split antecedents are concerned, PRO is pronoun-like with a restricted range of verbs in that

it can be co-indexed with more than one NP, however it is anaphor-like in that the antecedents must exist within the argument structure of the verb.

In the structure (25), the controller is indicated in the verb's lexical entry. It is impossible to determine the antecedent of PRO in the matrix sentence on the basis of structural considerations alone.

(25) [CP NP V (P) NP [CP (wh-) [IP PRO ...]]]

In the configuration (25), the lexical entry for V can select for one of three kinds of control, illustrated in (26)-(28).

(26) Subject control:
[NP_i V (P) NP [(wh-) [PRO_i ...]]]

(27) Object control:
[NP V (P) NP_i [(wh-) [PRO_i ...]]]

(28) Split control:
[NP_i V (P) NP_j [(wh-) [PRO_{i,j} ...]]]

A fourth option is predicted by the theory, that of no control or PRO_{One}. However, there appear to be no transitive verbs in English which selects PRO_{One}. Thus the structure (29) is ill-formed in English:

(29) No control:
* [NP_i V (P) NP_j [(wh-) [PRO_k ...]]]

The distribution of PRO_{One} is rather restricted. It never appears on a transitive verb. Since adjectives are never transitive, PRO_{One} is most commonly licensed by adjectives which assign no external θ -role. However, PRO_{One} can be assigned when a verb is intransitive, as in the structure (30).

(30) [CP NP V [CP (wh-) [IP PRO ...]]]

In the structure (30), V can select for one of two kinds of control: subject control, illustrated in (31), and no control, illustrated in (32).

(31) Subject control:
 [NP_{*i*} V [(wh-) [PRO_{*i*} ...]]]

(32) No control:
 [NP V [(wh-) [PRO_{*one*} ...]]]

7.1.2 LEXICALLY SPECIFIED ANAPHORA ON VERBS

Let us consider verbs which exemplify the range of control structures permitted by the model. The following is intended to illustrate the range of control structures. Just as verbs often have more than one subcategorization structure, they can likewise have more than one lexical control structure.

Since this model has two extrinsically assigned features, it predicts four classes of verbs: (1) those which license exclusive-PRO or PRO[-*g*,+*d*]; (2) those which license split-PRO or PRO[-*g*, -*d*]; (3) those which license PRO_{*one*} or [+*g*, -*d*]; and (4) those which license inclusive-PRO or PRO[+*g*, +*d*]:

1. Verbs which lexically license exclusive-PRO or PRO[-*gen*, +*def*] include:

promise: NP_{*i*} V NP [[PRO[-*g*, +*d*]_{*i*} to ...]]
John promised *Mary* PRO to leave.

persuade: NP V NP_{*i*} [[PRO[-*g*, +*d*]_{*i*} to ...]]
John persuaded *Mary* PRO to leave.

swear to: NP_{*i*} V [PP to NP] [[PRO[-*g*, +*d*]_{*i*} to ...]]
John swore to *Mary* PRO to stop smoking.

shout to: NP V [PP to NP_{*i*}] [[PRO[-*g*, +*d*]_{*i*} to ...]]
John shouted to *Mary* PRO to stop smoking.

2. Verbs which lexically license split-PRO, or PRO[-*gen*, -*def*] include:

propose to: $NP_i V [pp \text{ to } NP_j] [[PRO[-g,-d]i,j \text{ to } \dots]]$
John proposed to Mary to go to the movies.

3. Verbs which lexically license PRO_{one} or $PRO[+gen,-def]$ include:

say: $NP_i V [[PRO[+g,-d]j \text{ to } \dots]]$
John said PRO_{one} to leave.

4. Verbs which lexically license inclusive- PRO or $PRO[+gen,+def]$ include:

wonder: $NP_i V [wh- [PRO[+g,+d]i \text{ to } \dots]]$
John wondered what to wear.

ask: $NP_i V NP [wh- [PRO[+g,+d]i \text{ to } \dots]]$
John asked Mary PRO where to go.

tell: $NP V NP NP_i [wh- [PRO[+g,+d]i \text{ to } \dots]]$
John told Mary where PRO to go.

learn from: $NP_i V [pp \text{ from } NP] [wh- [PRO[+g,+d]i \text{ to } \dots]]$
John learned from Mary PRO how to knit.

explain to: $NP V [pp \text{ to } NP_i] [wh- [PRO[+g,+d]i \text{ to } \dots]]$
John explained to Mary how PRO to knit.

When a verb supports more than one subcategorization frame, it can specify for more than one kind of control. There is only one control specification per subcategorization frame.

7.1.3 LEXICALLY SPECIFIED ANAPHORA ON ADJECTIVES

1. Adjectives which lexically license exclusive- PRO , or $PRO[-gen,+def]$, include:

be afraid: $NP_i V [[PRO[-g,+d]i \dots]]$
John is afraid [PRO] to go out in the dark.

be painful: be ADJ $[pp \text{ to } NP_i] [[PRO[-g,+d]i \text{ to } \dots]]$
It is painful to John PRO to say good-bye.

2. In order to have an adjective which licenses $PRO[-g,-d]$, i.e. has a split antecedent, it is necessary to have an adjective which assigns a θ -role to a subject, a θ -role to a PP complement, and an

infinitival complement clause. Such adjectives appear not to exist in English.

3. Adjectives which assign no θ -roles license PRO_{one} , or $PRO[+gen,-def]$. These include:

be appropriate: be ADJ [[$PRO[+g,-d]$ to ...]]
It is appropriate PRO_{one} to say good-bye.

be easy: be ADJ [[$PRO[+g,-d]$ to ...]]
It is easy PRO_{one} to say good-bye.

be painful: be ADJ [[$PRO[+g,-d]$ to ...]]
It is painful PRO_{one} to say good-bye.

be unclear: be ADJ [$wh-$ [$PRO[+g,-d]$ to ...]]
It is unclear when PRO_{one} to say good-bye.

PRO_{one} is only assigned to infinitive complements of adjectives that do not assign an external θ -role and do not take a PP complement.

4. Adjectives which lexically license inclusive- PRO , or $PRO[+gen,+def]$, include:

be unclear: be ADJ [pp to N_I] [$wh-$ [$PRO[+g,+d]_i$ to ...]]
It is unclear to *John* when *PRO* to leave.

7.2 LONG-DISTANCE CONTROL AND BINDING

Long-distance control is a phenomenon which occurs with predicate adjective constructions. Much of the literature on long-distance control centers around the predicates *be appropriate* and *be unclear*. With these, as well as all other instances of long-distance control, the adjective assigns a clausal complement and does not assign an external θ -role. Since the adjective has no θ -roles to assign, it cannot lexically specify control of PRO in the infinitive complement. Therefore, my analysis predicts that the subject of the infinitive is PRO_{one} .

However, in cases of long-distance control, PRO is coreferential with one or more NPs in a higher clause. By my analysis of the interpretation of PRO, PRO_{one} cannot be coindexed with a higher NP because PRO_{one} has the feature [-definite].

Due to the principles of UG which account for the distribution of PRO, PRO cannot be assigned an antecedent by Binding Theory. However, nothing prevents binding theory from assigning PRO as an antecedent for a reflexive pronoun. I show that when PRO is the antecedent of a reflexive pronoun, it inherits, among other features, the feature [+definite] from the reflexive definite pronoun.

In "The reference of the Indefinite Empty Pronoun", I review the referential properties of PRO_{one} in predicate adjective constructions, and then show how they differ when embedded in another sentence. Then in "Long-Distance Control" on page 249, I propose a control strategy which operates on binding structures.

7.2.1 THE REFERENCE OF THE INDEFINITE EMPTY PRONOUN

When there is no possible NP antecedent in the sentence, the interpretation of PRO must be PRO_{one} (PRO[+gen,-def]). The reference of PRO_{one} is discussed in detail in "The Pronominal Anaphor 'One'" on page 209. PRO_{one} is generic in that it does not have an antecedent NP in the sentence, but it may include reference to the speaker (first person) or the hearer (second person). Being [+generic], PRO_{one} can license the pronoun *oneself*.

The lexical entry for the adjective *appropriate* licenses the infinitive complement and indicates that the features on PRO are [+gen,-def]. In addition, *appropriate* does not assign an external

θ-role. Thus the entry for *appropriate* licenses the structure shown in (33).

(33) be appropriate: [NP it] _____ [CP [IP PRO[+gen,-def] to VP]]

The sentences in (34) are licensed by *appropriate*, when PRO is interpreted as PRO_{one}:

- (34) i. It would be appropriate [[PRO[+g,-d] to pay for Bill]]
 ii. It would be appropriate [[PRO[+g,-d] to leave]]
 iii. It would be appropriate [[PRO[+g,-d] to pay for oneself]]
 iv. * It would be appropriate [[PRO_i to pay for himself_i]]
 v. It would be appropriate [[PRO[+g,-d] to pay for each other]]
 vi. It would be appropriate [[PRO[+g,-d] to leave together]]

There are two lexical entries for the adjective *unclear*, both of which take a *wh*- infinitive complement. When *unclear* occurs with a PP complement, PRO is coreferential with the NP object of the preposition, as in (35). When *unclear* has no PP complement, PRO is interpreted as PRO_{one}, as in (36). Neither *unclear*₁ nor *unclear*₂ assigns an external θ-role.

(35) be unclear₁: [NP it] _____ [PP to NP_i] [CP *wh*- [IP PRO[+g,+d]_i to VP]]

(36) be unclear₂: [NP it] _____ [CP *wh*- [IP PRO[+g,-d] to VP]]

The following structures are licensed by *unclear*₁ as in (35).

- (37) i. It is unclear to John_i [when [PRO[+g,+d]_i to pay the bill]]
 ii. It is unclear to John_i [when [PRO[+g,+d]_i to leave]]
 iii. It is unclear to John_i [when [PRO[+g,+d]_i to pay for oneself]]
 iv. It is unclear to John_i [when [PRO[+g,+d]_i to pay for himself_i]]
 v. * It is unclear to John [when [PRO_i to pay for [each other]_i]]
 vi. * It is unclear to John [when [PRO to leave together]]

The lexical entry for *unclear*₁ licenses (a) specific features assigned to PRO and (b) the coreference of PRO and *John*. The feature [+gen] on PRO licenses *oneself* in (37iii). In (37iv), *himself* is bound to PRO_i

which is lexically specified to be coreferential with *John_i*; and *John*, being third person masculine singular, licenses *himself*. In (37v), *each other*, requiring a plural, is bound to PRO_{*i*} which is lexically indicated as coreferential with *John_i*, a singular, hence, the sentence is ill-formed. Similarly, in (37vi) the predicate *leave together* requires a plural subject, but PRO_{*i*} is lexically linked to the singular *John_i*.

The following structures are licensed by the lexical entry for *unclear₂*, as in (36).

- (38) i. It is unclear [when [PRO[+g,-d] to pay the bill]]
 ii. It is unclear [when [PRO[+g,-d] to leave]]
 iii. It is unclear [when [PRO[+g,-d] to pay for oneself]]
 iv. * It is unclear [when [PRO_{*i*} to pay for himself_{*i*}]]
 v. It is unclear [when [PRO[+g,-d] to pay for each other]]
 vi. It is unclear [when [PRO[+g,-d] to leave together]]
 vii. It is unclear [when [PRO[+g,-d] to pay for myself]]
 viii. It is unclear [when [PRO[+g,-d] to pay for yourself]]
 ix. It is unclear [when [PRO[+g,-d] to pay for ourselves]]
 x. * It is unclear [when [PRO_{*i*} to pay for themselves_{*i*}]]

PRO[+gen,-def] is PRO_{*one*}. PRO_{*one*} is a possible antecedent for a plural (*each other, ourselves*) and can be the subject of a plural predicate (*leave together*). PRO_{*one*} cannot be the antecedent of a definite reflexive (*himself, herself, or themselves*).

The sentences (38i-iii) and (38v-vi) can be embedded in other sentences without affecting the interpretation of PRO:

- (39) i. Mary told Bill that it is unclear when PRO[+g,-d] to pay the bill.
 ii. Mary told Bill that it is unclear when PRO[+g,-d] to leave.
 iii. Mary told Bill that it is unclear when PRO[+g,-d] to pay for oneself.
 iv. Mary told Bill that it is unclear when PRO[+g,-d] to pay for each other.
 v. Mary told Bill that it is unclear when PRO[+g,-d] to leave together.

PRO_{one} in the sentences in (38) is understood to be anyone. In sentences (39), PRO_{one} can also be understood to be anyone, including anyone mentioned in the higher sentences. However, when (38iv) and (38vii-x) are embedded in a higher sentence, the subjects of the infinitive are no longer interpreted as PRO_{one}. For example, if the infinitive subject in (40) were PRO_{one}, then (40) would be predicted to be ill-formed.

- (40) John told Mary [that it is unclear [when [PRO to pay for themselves]]]

Not only is (40) well-formed, but PRO is being interpreted as [+gen,+def]. The [-def] feature is being rewritten as [+def]. This and other examples of this phenomenon, which is called long-distance control, are discussed in the next section. Notice that in the sentences in (41), PRO_[+g,+d] is lexically specified via (35) to be coreferential with the object of the PP, *Bill*.

- (41) i. Mary told John that it is unclear to Bill_{*f*} when PRO_[+g,+d]_{*i*} to pay.
 ii. Mary told John that it is unclear to Bill_{*f*} when PRO_[+g,+d]_{*i*} to pay for oneself.
 iii. * Mary told John that it is unclear to Bill_{*f*} when PRO_{*f*} to pay for each other.

7.2.2 LONG-DISTANCE CONTROL

Long-distance control is the name for the type of PRO-antecedent reference found in sentences like (42) and (43), where PRO is controlled over two CP nodes. In the sentences in (42)-(43), PRO is [+generic] in that it may include the antecedent but not be identical to it. It is [+definite] in that it must have at least one antecedent. Thus PRO in long-distance control environments has the features [+gen,+def].

- (42) i. Mary told Bill it would be appropriate [[PRO_[+g,+d] to pay for herself]]

- ii. Mary told Bill it would be appropriate [[PRO[+g,+d] to pay for himself]]
 - iii. Mary told Bill it would be appropriate [[PRO[+g,+d] to pay for themselves]]
- (43) i. Mary thinks it would be appropriate [[PRO[+g,+d] to pay for herself]]
- ii. * Mary thinks it would be appropriate [[PRO to pay for himself]]
 - iii. Mary thinks it would be appropriate [[PRO[+g,+d] to pay for themselves]]

In summary, PRO[+gen,+def] in long-distance binding behaves like a pronoun with respect to possible antecedents, but like an anaphor in that it must have a possible antecedent NP in the sentence. The question remains: How does PRO acquire these features?

There are four salient properties of long-distance control:

1. In cases of long-distance control, PRO is always linked by binding theory to a third person reflexive with gender and number (*herself*, *himself*, and *themselves*).
2. A sentence with long-distance control is always well-formed if the definite reflexive is replaced by an NP which either takes PRO_{one} as an antecedent (*oneself*) or requires no antecedent (a pronoun or an r-expression). If the definite reflexive in (42)-(43) is replaced by *oneself*, as in (44), the sentence remains well-formed, with perhaps a slightly different interpretation.

- (44) i. Mary told Bill it would be appropriate [[PRO[+g,-d] to pay for oneself]]
- ii. Mary thinks it would be appropriate [[PRO[+g,-d] to pay for oneself]]

When the reflexive is *oneself*, the features on PRO are PRO[+gen,-def], since PRO needs no specific reference to any NP in the sentence. A sentence with long-distance control is always

well-formed if the definite reflexive bound to the PRO is replaced by an r-expression, e.g. *Sue* in (45).

- (45) i. Mary told Bill it would be appropriate [[PRO[+g,-d] to pay for Sue]]
 ii. Mary thinks it would be appropriate [[PRO[+g,-d] to pay for Sue]]

The structure and feature assignment to the infinitive clauses in sentences (44)-(45) are licensed by the lexical structure for *appropriate*, as in (33).

3. PRO in long-distance control structures has the features [+gen,+def] and must have an antecedent of the correct person, number, and gender in the sentence; however, PRO is understood to include the antecedent, not to be necessarily coreferential with it.
4. When PRO exhibits long-distance control, it is always in an adjective construction with the structure (46). The adjective has no PP complement and assigns no external θ -role.

- (46) ... it be ADJ [*CP* [*IP* PRO to [*VP* V ... [*NP* himself/herself/themselves]]]]

I propose the following analysis for *long-distance control*. I assume the parsing strategy proposed in (1). First the grammar builds a phrase marker using information from the lexicon. The sentence containing PRO is licensed as the infinitive complement of an adjective. Since the adjective has no θ -roles to which control can be assigned, PRO is assigned the features [+g,-d], as in (47).

- (47) Mary told Bill it would be appropriate [[PRO[+g,-d] to pay for herself]]

Second, the parser binds anaphors. The definite reflexive *herself* must be A-bound by an antecedent in a local domain. When a definite reflexive

is bound to PRO, its features are also bound to the PRO. PRO now has, among other features, the features PRO[+gen,+def], as shown in (48).

- (48) Mary told Bill it would be appropriate [[PRO[+g,+d]_i to pay for herself_j]]

Finally, in determining the reference for control of PRO, we are confronted with two issues: (a) What is the interpretation of the PRO[+gen,+def]? (b) Since a [+def] PRO must have an antecedent NP, what mechanisms indicate which NP is the antecedent?

The feature [+def] means that PRO must have an antecedent. PRO in long-distance control is an anaphor in that it must have an antecedent. In choosing an antecedent, it behaves exactly like a pronoun of the same person, number, and gender of the reflexive from which it gets its features. When it is the antecedent of a singular reflexive pronoun (*herself* and *himself*), the antecedent of PRO must agree in person, number and gender. The antecedent of PRO in (49i) could be *Mary* or *Sue*, although it must be one of them.

- (49) i. Mary told Sue it would be appropriate [[PRO[+g,+d]_i to pay for herself_j]
 ii. Mary told Bill it would be appropriate [[PRO[+g,+d]_i to pay for herself_j]
 iii. Bill told Sue it would be appropriate [[PRO[+g,+d]_i to pay for herself_j]
 iv. * Bill told Tom it would be appropriate [[PRO[+g,+d]_i to pay for herself_j]

When the reflexive is plural (*themselves*), then the antecedent can be satisfied by anything that would satisfy the antecedent of a plural pronoun. In the latter case, PRO includes, but is not identical to, the antecedent.

- (50) i. Mary said that it would be appropriate [[PRO[+g,+d] to pay for themselves]]

- ii. Mary told Bill that it would be appropriate [[PRO[+g,+d] to pay for themselves]]

The antecedent NP of the PRO cannot be lexically specified in the entry of the adjective since the antecedent can be arbitrarily far from the adjective, and it may be any one of several possibilities, as we saw in (49).

The basic principles for determining the antecedent of PRO[+gen,+def] in the case of long-distance control are these:

1. Since PRO is [+anaphor,+pronoun], the antecedent of PRO cannot be determined by binding theory.
2. When binding assigns PRO as an antecedent to a definite reflexive pronoun, PRO inherits the features [+def, α -feminine, β -plural]. This results in the [-def] feature on PRO[+gen,-def] being rewritten as [+def]. The possibilities are shown below:

PRO[+gen,-def]_i...himself_i ==> PRO[+gen,+def,-fem,-plu]

PRO[+gen,-def]_i...herself_i ==> PRO[+gen,+def,+fem,-plu]

PRO[+gen,-def]_i...themselves_i ==> PRO[+gen,+def,+plu]

3. The feature [+def] means that PRO must have an antecedent.
4. The feature [+gen] means that PRO need not be identical with the antecedent, but only include the antecedent.
5. The features [α -feminine, β -plural] indicate what is a possible antecedent.

Whatever mechanisms determine the controller of PRO are the same as those which determine possible antecedents for overt pronouns across two sentence boundaries. Notice that binding theory has nothing to say about either problem, other than that the pronoun is free in the local domain.

Since PRO is in subject position, there is no A-position to which it could possibly be bound in its local domain.

7.3 FEATURE ASSIGNMENT BY GRAMMATICAL FORMATIVES

Grammatical formatives are closed class lexical items. Like lexical formatives (open class lexical items), grammatical formatives give the parser instructions about how to build a tree. However, whereas lexical formatives license predictable patterns, grammatical formatives alert the parser to expect something out of the ordinary. For example, the grammatical formative *be+en* alerts the parser that the order of elements it will encounter in the input string does not correlate with the order of elements it would expect given the lexical entries of the other words in the string. In this sense, grammatical formatives can be viewed as triggering subroutines in the parser which are required to map the information in the lexicon onto the input string. An example of a grammatical formative that we have encountered is the adverbial *too* which, as a specifier, licenses a complement.

In this section, I discuss how the parser builds structure for passive clauses and assigns control to infinitive complements in passives. I show how the set of grammatical formatives also participate in the feature assignment to PRO in passives.

7.3.1 THE ASSIGNMENT OF STRUCTURE AND CONTROL TO PASSIVES

The lexical entries for the transitive verbs *persuade*, *promise*, and *ask* license the structures in (51)-(54). Notice that two structures are proposed for *ask*: *ask*₁ takes a propositional clause and *ask*₂ takes an indirect question.

- (51) *persuade*: NP _____ NP_{*i*} [CP [IP PRO[-gen,+def]_{*i*} to VP]]
- (52) *promise*: NP_{*i*} _____ NP [CP [IP PRO[-gen,+def]_{*i*} to VP]]
- (53) *ask*₁: NP _____ NP_{*i*} [CP [IP PRO[-gen,+def]_{*i*} to VP]]
- (54) *ask*₂ NP_{*i*} _____ NP [CP *wh*- [IP PRO[+gen,+def]_{*i*} to VP]]

The lexical structures will license the tree building mechanisms to construct a tree like (55) for the strings (56)-(59).

- (55) [CP [IP NP₁ [VP V NP₂ [CP [IP PRO to VP]]]]]
- (56) John persuaded Mary_{*i*} [[PRO[-g,+d]_{*i*} to leave]]
- (57) John_{*i*} promised Mary [[PRO[-g,+d]_{*i*} to leave]]
- (58) John asked Mary_{*i*} [[PRO[-g,+d]_{*i*} to leave]]
- (59) John_{*i*} asked Mary [when [PRO[+g,+d]_{*i*} to leave]]

The passive morphology, *be*+en, instructs the parser to manipulate lexical information differently than for the indicative strings (55)-(59). For example, the parser encounters passive morphology upon parsing (60).

- (60) John was arrested.

The lexical entry for *arrest* indicates that *arrest* is a transitive verb and that it assigns an external θ -role. The passive morphology informs the parser that (1) the external θ -role is suppressed and (2) the verb's objective case is suppressed. *John* in subject position is therefore not in a θ -position. Since the object of *arrest* is missing, an empty category is inserted in the object position. As this object position is a governed, uncased position, the empty category is determined to be an anaphor. As such, the anaphor must be bound in its local domain. *John* is the only available antecedent. In addition, if *John* is not bound to a θ -position, the string will have a Theta Criterion violation. The

parser binds the empty category to *John*, as in (61), thereby satisfying both requirements.

(61) *John*_{*i*} was arrested *e*_{*i*}

Consider now the passive forms of *persuade*, *promise*, and *ask*. The abstractions from the lexical entries shown in (51)-(54), in conjunction with information supplied by the passive marker, license the tree building mechanisms to assign a tree like (55) to each of the following.

(62) *Mary*_{*i*} was persuaded *e*_{*i*} [[PRO[-*g*,+*d*]_{*i*} to leave]]

(63) * *Mary*_{*j*} was promised *e*_{*j*} [[PRO[-*g*,+*d*]_{*i*} to leave]]

(64) *Mary*_{*i*} was asked *e*_{*i*} [[PRO[-*g*,+*d*]_{*i*} to leave]]

(65) * *Mary*_{*j*} was asked *e*_{*j*} [when [PRO[+*g*,+*d*]_{*i*} to leave]]

Before we discuss the mechanisms that license (62) and (64), and block (63) and (65), I want to point out that (65), when assigned with the features of inclusive-PRO, is ill-formed. When PRO has the features [+definite], it requires an antecedent. In (65), PRO cannot include *Mary* in its reference. However, (66), where PRO is interpreted as PRO_{one}, is well-formed. I will return to this directly.

(66) *Mary*_{*j*} was asked *e*_{*j*} [when [PRO[+*g*, -*d*]] to leave]]

Consider the mechanisms which license the structure, the feature assignment of PRO, and the interpretation of PRO in (62). The lexical entry for *persuade*, as reflected in (51), licenses the assignment of tree structure (55) to string (62). It then indicates that PRO must be bound to the NP₂ position.

(51) *persuade*: NP _____ NP₁ [CP [IP PRO[-*gen*,+*def*]_{*i*} to VP]]

(55) [CP [IP NP₁ [VP V NP₂ [CP [IP PRO to VP]]]]]

(62) *Mary*_{*i*} was persuaded *e*_{*i*} [[PRO[-*g*,+*d*]_{*i*} to leave]]

The NP₂ position is present in the structure and the lexical marking of *persuade* licenses the coindexing of NP₂ with PRO. *Mary* in NP₁ position is indexed with the empty category in NP₂ position. Since *Mary* is coindexed with *e*, and *e* is coindexed with PRO, the sentence is well-formed with the interpretation that it is *Mary* who leaves.

In (63), the lexical entry reflected in (52) licenses the assignment of structure (55) to string (63).

(52) *promise*: NP₁ _____ NP [CP [IP PRO[-gen,+def]_i to VP]]

(55) [CP [IP NP₁ [VP V NP₂ [CP [IP PRO to VP]]]]]

(63) * *Mary_j* was promised *e_j* [[PRO[-g,+d]_i to leave]]

The lexical entry for *promise* indicates that PRO is obligatorily coindexed with the NP bearing the external θ -role assigned by *promise*. However, the NP bearing the external θ -role of *promise* is not present in the sentence. Since the string (63) cannot satisfy the lexical markings for the features and coreference of PRO of *promise*, (63) is ill-formed.

If the external θ -role of *promise* is present in the sentence as the NP object of the preposition *by*, PRO may not be linked with the NP *John*, as in (67).

(67) * *Mary_j* was promised *e_j* by *John* [[PRO[-g,+d]_i to leave]]

Sentence (64) is licensed by the lexical structure reflected in (53) of *ask*₁.

(53) *ask*₁: NP _____ NP₁ [CP [IP PRO[-gen,+def]_i to VP]]

(64) *Mary_i* was asked *e_i* [[PRO[-g,+d]_i to leave]]

When *ask* takes a propositional complement, PRO is coindexed with the internal argument of *ask*, NP₂. *Mary*, in NP₁ position, is coindexed with

NP₂, and NP₂ is coindexed with PRO. The presence of a *by* prepositional phrase does not alter the analysis, and (68) has the same properties of PRO reference as (64).

(68) Mary_{*i*} was asked *e_i* by John [[PRO[-*g*,+*d*]_{*i*} to leave]]

However, if these mechanisms are applied to *ask*₂ in a passive construction, they assign a bad structure, producing the ill-formed structure in (65). That is, the lexical entry for *ask*₂, (54), indicates that when *ask* has a *wh*- complement, PRO is coindexed with the external argument of *ask*, as in (59).

(54) *ask*₂: NP_{*i*} _____ NP [*CP wh*- [*IP* PRO[+*gen*,+*def*]_{*i*} to VP]]

(59) John_{*i*} asked Mary [when [PRO[+*g*,+*d*]_{*i*} to leave]]

(65) * Mary_{*j*} was asked *e_j* [when [PRO[+*g*,+*d*]_{*i*} to leave]]

The lexical entry for *ask*₂ licenses the structure (65), but since there is no external argument to be coindexed with PRO_{*i*}, the sentence is ill-formed.

According to this analysis, (69) is falsely predicted to be ill-formed for the same reason that (70) is ill-formed. However, sentence (69) is well-formed.

(69) Mary was asked when to leave.

(70) * Mary was promised to leave.

Given that (69) is well-formed despite the fact that it cannot be generated from the lexical entry (54), the question arises: For sentence (69), what principles and mechanisms license (a) the phrase marker, (b) the features on PRO, and (c) the reference properties of PRO?

The features on PRO in (69) are PRO_[+gen,-def], i.e., PRO_{one}. PRO does not refer to any NP in the sentence, in particular, not *Mary* in (71), and neither *Mary* nor *John* in (72).

(71) *Mary*_{*i*} was asked *e*_{*i*} [when [PRO_[+g,-d] to leave]]

(72) *Mary*_{*i*} was asked *e*_{*i*} by *John*_{*j*} [when [PRO_[+g,-d] to leave]]

Sentence (69) can be licensed by the grammar if we make the following assumption: There is a *wh*-PRO_{one} licensing mechanism, which has the effect of licensing PRO_{one} in an indirect question when there is no antecedent.

(73) *wh*-PRO_{one} licensing mechanism: If there is a structure like (73i) in which PRO has no features, then the presence of a *wh*-element in COMP licenses the assignment of the features [+gen,-def] to PRO, as in (73ii).

- i. [CP *wh*- [IP PRO to VP]]
- ii. [CP *wh*- [IP PRO_[+gen,-def] to VP]]

This mechanism is also postulated to account for control in adjuncts with empty operators.

Given the assumption that structure is built and binding operates before control does, the lexical entry (54) would license the grammar to assign structure (55) to (69) giving a structure like (74) with no features or reference indicated for PRO.

(54) *ask*_{*2*} NP_{*i*} _____ NP [CP *wh*- [IP PRO_[+gen,+def]_{*i*} to VP]]

(55) [CP [IP NP_{*1*} [VP V NP_{*2*} [CP [IP PRO to VP]]]]]]

(69) *Mary* was asked when to leave.

(74) [*Mary*_{*i*} was asked *e*_{*i*} [when [PRO to leave]]]]

The *wh*-PRO_{one} licensing mechanism assigns the features [+gen,-def] to (74), yielding (75).

(75) [*Mary*_{*i*} was asked *e*_{*i*} [when [PRO_[+g,-d] to leave]]]]

When the PRO interpretation mechanisms refer to the lexical entry for *ask*₂ for lexically specified feature assignment, the feature assignment is not licensed because the lexical entry for *ask*₂ specifies control by the external θ -role, and the external θ -role is suppressed in the passive.

The passive of *promise* is blocked by the grammar because the *wh*-PRO_{one} licensing mechanism does not apply to it. The *wh*-PRO_{one} licensing mechanism only applies to indirect question complements.

By these mechanisms, the derivation for (76) is as follows:

(76) John asked Mary when to leave.

The tree structure (77) is licensed by general principles and the lexical entry of *ask*:

(77) John asked Mary [when [PRO to leave]]

The *wh*-PRO_{one} licensing mechanism assigns the feature bundle [+gen,-def] to PRO in interrogative complements, as in (78).

(78) John asked Mary [when [PRO[+g,-d] to leave]]

Finally, the lexical entry for *ask*₂ licenses the features [+gen,+def], along with control by the external argument. Since the external argument is present, PRO is assigned the feature [+definite], and PRO is coindexed with the external argument, *John*, as in (79).

(79) John_i asked Mary [when [PRO[+g,+d]_i to leave]]

The *wh*-PRO_{one} licensing mechanism also licenses feature assignment to the strings in (80)-(82). In (80) output of the *wh*-PRO_{one} licensing mechanism in (80ii) gets rewritten as (80iii) by lexical feature specification. In (81) and (82), since there is no possible antecedent, the final interpretation remains PRO_{one}.

- (80) i. John wonders how to brush his teeth.
 ii. John wonders [*CP* how [*IP* PRO[+g,-d] to brush his teeth]]
 iii. John wonders [*CP* how [*IP* PRO[+g,+d] to brush his teeth]]
- (81) i. It is unclear when to leave.
 ii. It is unclear [*CP* when [*IP* PRO[+g,-d] to leave]]
- (82) i. What to wear is always a problem.
 ii. [*CP* what [*IP* PRO[+g,-d] to wear]] is always a problem.

7.4 THE CONSTITUENT STRUCTURE OF ADJUNCTS

Infinitival adjunct constructions are discussed by different authors using different names. Terms used to classify the constructions include causatives, purpose clauses, relative clauses, adjective modifiers or qualifiers, and degree phrases. I will classify adjunct types according to the category of the maximal projection. The class of infinitival adjunct constructions include these sentences:

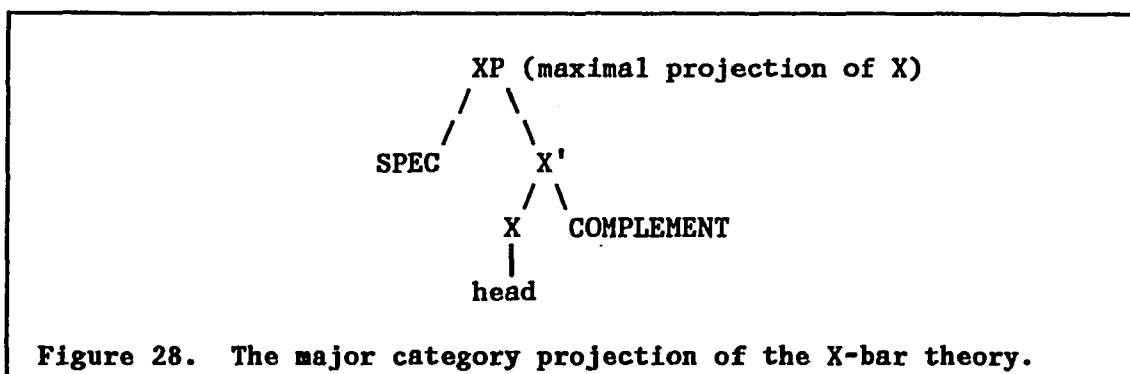
- (83) John sneezed [(in order) [PRO to get attention]]
- (84) i. John shot Bill [(in order) [PRO to keep himself busy]]
 ii. Bill was shot [(in order) [PRO to keep him quiet]]
- (85) i. Mary knows the candidate [PRO to beat]
 ii. Mary knows the candidate [PRO to beat Bill]
- (86) i. John is too stubborn [PRO to talk to]
 ii. John is too stubborn [PRO to talk to Bill]
- (87) i. John ran slowly enough [PRO to catch]
 ii. John ran quickly enough [PRO to catch Bill]
- (88) i. Mary is pretty [PRO to look at]
 ii. Perfume is sweet [PRO to smell]

In "Types of Adjuncts" on page 262, I show what phrase marker is assigned to adjunct constructions in general, and infinitival adjuncts in particular. In "The Tree Building Mechanisms for Infinitival Adjunct Clauses" on page 267, I discuss the tree building mechanisms which license the construction of the embedded clause containing PRO and how

it is attached to the matrix clause. Then, in the sections that follow, I discuss the mechanisms of the control module which indicate any coindexing of PRO and assign PRO its features [\pm gen, \pm def] are proposed.

7.4.1 TYPES OF ADJUNCTS

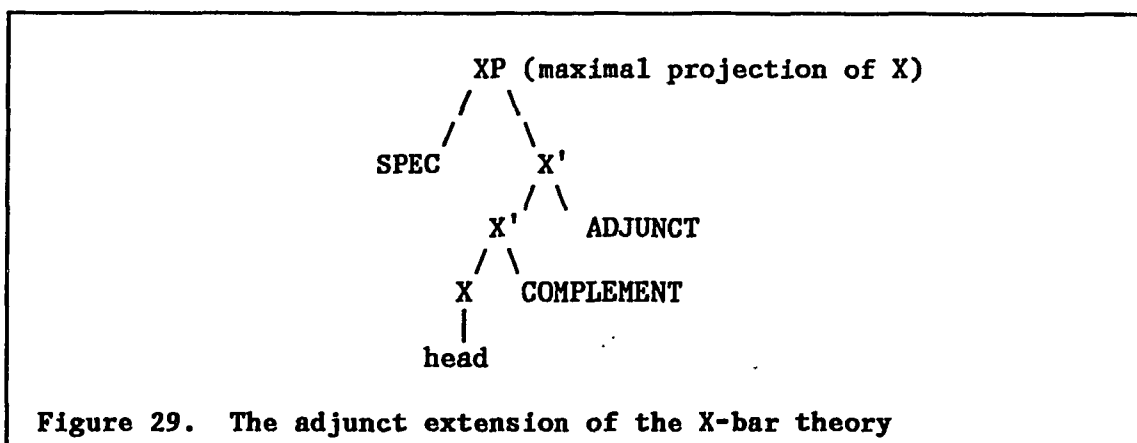
I assume, in accordance with X-bar theory, that each major category is a projection of a lexical category. The internal structure of each major category is given by the structure in Figure 28 on page 262.



The elements which are present in the structure Figure 28 on page 262 are licensed by the lexical entry of the head. All elements in the X-bar structure which are licensed by the lexical entry of a head are called the complements of the head. By definition, complements are licensed by the lexical entry of the head.

I will call the structure in Figure 29 on page 263, the adjunct extension of the X-bar notation. Major categories are projections of lexical categories, but there is an X' node introduced by the adjunct mechanism, and this X' node dominates the X' node assigned by the X-bar convention in Figure 28 on page 262. The X' adjunct node dominates

adjuncts. By definition, adjuncts are not licensed by the lexical entry of the head, but by the general type of category X and by the conditions imposed by various modules. The attachment of an adjunct to the adjunct extension X' is licensed, subject to distributional requirements discussed below, by the category type, X, and not by the lexical entry of the head.



I assume that the four lexical categories are defined by the features [$\pm V$] and [$\pm N$]. Each of these lexical categories can license one or more adjuncts. The categories can license the following types of constructions. This list is illustrative, not exhaustive.

(89) Verb: [$+V, -N$]

Mary and John sneezed <i>together</i> .	manner adverbial
(<i>how</i>)	
John sneezed <i>so that he could breath</i> .	causative adverbial
(<i>why</i>)	
John sneezed (<i>in order</i>) <i>to breath</i> .	causative adverbial
(<i>why</i>)	

(90) Noun: [$-V, +N$]

The fact <i>that Bill put in front of us</i> .	relative clause (<i>which</i>)
The candidate <i>who we must beat</i> .	relative clause (<i>which</i>)
The candidate <i>to beat</i> .	relative clause (<i>which</i>)

(91) ADVerb: [-V,-N]

John ran quickly <i>enough</i> .	degree adverbial (<i>how</i>)
John ran quickly <i>enough to catch Bill</i> .	degree adverbial (<i>how</i>)
John ran quickly <i>enough to catch</i> .	degree adverbial (<i>how</i>)
John is <i>too stubborn to talk to Bill</i> .	degree adverbial (<i>how</i>)

(92) ADJective: [+V,+N]

Mary is pretty <i>enough</i> .	degree adverbial (<i>how</i>)
Mary is <i>very</i> pretty.	degree adverbial (<i>how</i>)

The adjunct constructions are labeled *manner adverbial*, etc. because of the type of *wh*- question they answer. The precise labels or interpretation of the adjuncts is peripheral to our main concern which is to license infinitive adjunct clauses. For reference, I assume the following labels for constructions.

- (93) Manner adverbials answer the questions: *how*, or *in what manner*.

How did Mary and John sneeze?

- (94) Causative adverbials answer the questions: *why*, *how come*, or *for what reason*.

Why did John sneeze?

- (95) Degree adverbials answer the questions: *how*, *how much*, or *to what extent*.

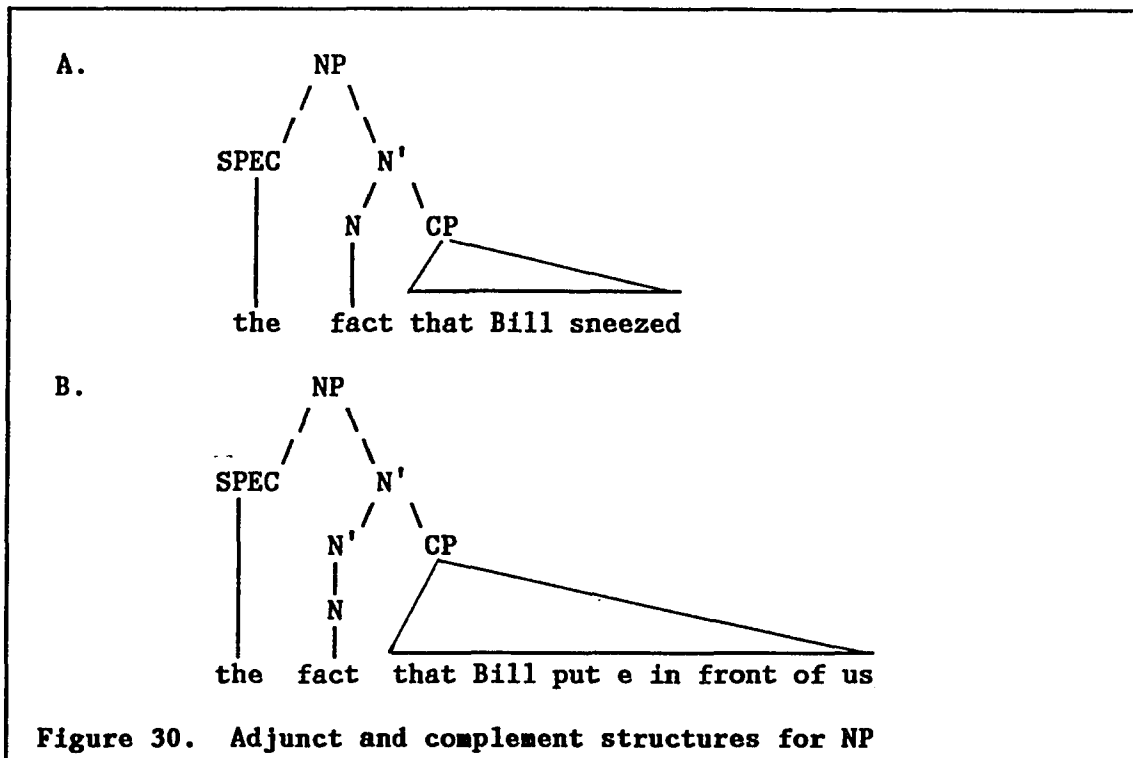
How quickly did John run?
How pretty is Mary?

- (96) Relative clauses answer the question: *which*, *what*, or *which one*.

Which fact did Bill put in front of us?
Which candidate is the most important one?

A relative clause is an adjunct on an NP. Some nouns, like *fact*, license a CP complement in their lexical entry. The CP in (97) is a complement of *fact* with the complement structure (A) in Figure 30 on

page 265. The CP in (98) is a relative with the adjunct structure (B) in Figure 30 on page 265.¹⁸



¹⁸ There is some literature on the difference between complements and adjuncts for nouns like *fact*, *scandal*, etc. Often there are determiner/complement restrictions:

- (a) It was a/*the scandal that John sneezed on the Queen.
 (b) Mary remembered the/*a fact that John sneezed on the Queen.

The complement and the adjunct can occur together, and, as follows from the structure in Figure 29 on page 263, the complement must be adjacent to the N and the adjunct to the right of the complement:

- (c) Mary remembered the fact that John sneezed on the Queen which the rest of us had forgotten.
 (d) Mary remembered the fact which the rest of us had forgotten that John sneeze on the queen.
 (e) Columbus' promise to Isabella to find gold, which was put in a contract, led to the discovery of America.
 (f) The advice of the lawyers, which was never followed, was to sue.

- (97) i. The fact that Bill sneezed annoyed everyone.
 ii. * The fact which Bill sneezed annoyed everyone.
- (98) i. The fact that Bill put in front of us annoyed everyone.
 ii. The fact which Bill put in front of us annoyed everyone.

The names assigned to the adjunct infinitive clauses are in the chart in Figure 31 on page 266.

CATEGORY	FEATURES	ADJUNCT INFINITIVE CLAUSE
Verb:	[+V,-N]	in order to clause
Noun:	[-V,+N]	infinitival relative clause
ADV/PREP:	[-V,-N]	degree adverbial clause
ADJ:	[+V,+N]	copula clause

Figure 31. Adjunct infinitive clause types

Below are examples of each of the four types of adjunct infinitive clauses:

1. VP adjuncts, or *in order to* clauses:

- (99) i. John sneezed in order to get attention.
 ii. John_i [_{VP} [_{V'} [_{V'} [_V sneezed]]] [_{CP} in order [_{PRO}[-g,+d]_i to get attention]]]]
- (100) i. John shot Bill in order to keep him quiet.
 ii. John_i [_{VP} [_{V'} [_{V'} [_V shot] Bill]] [_{CP} in order [_{PRO}[-g,+d]_i to keep him quiet]]]]
- (101) i. Bill was shot in order to keep him quiet.
 ii. Bill_i was [_{VP} [_{V'} [_{V'} [_V shot] e_i]] [_{CP} in order [_{PRO}_{one} to keep him quiet]]]]

2. NP adjuncts or infinitival relative clauses:

- (102) i. Mary knows the candidate to beat.

- ii. Mary knows [NP [SPEC the] [N' [N' [N candidate_i]]]
[CP O_i [PRO_{one} to beat e_i]]]]

(103) i. Mary knows the candidate to beat Bill.

- ii. Mary knows [NP [SPEC the] [N' [N' [N candidate_i]]] [CP [
PRO[-g,+d]_i to beat Bill]]]]

3. ADVP adjuncts or degree adverbial clauses:

(104) i. John ran slowly enough to catch.

- ii. John_i ran slowly [ADVP [ADV' [ADV' [ADV enough]]]
[CP O_i [PRO_{one} to catch e_i]]]]

(105) i. John ran quickly enough to catch Bill.

- ii. John_i ran quickly [ADVP [ADV' [ADV' [ADV enough]]]
[CP [PRO[-g,+d]_i to catch Bill]]]]

(106) i. John is too stubborn to talk to.

- ii. John_i is [AP too [A' [A' [A stubborn]]] [CP O_i
[PRO[+g,-d] to talk to e_i]]]]

(107) i. John is too stubborn to talk to Bill.

- ii. John is [AP too [A' [A' [A stubborn]]] [CP [PRO[-g,+d]_i to
talk to Bill]]]]

4. AP adjuncts or copula clauses:

(108) i. Mary is pretty to look at.

- ii. Mary_i is [AP [A' [A' [A pretty]]] [CP O_i [PRO_{one} to look at
e_i]]]]

(109) i. Perfume is sweet to smell.

- ii. perfume_i is [AP [A' [A' [A sweet]]] [CP O_i [PRO_{one} to smell
e_i]]]]

7.4.2 THE TREE BUILDING MECHANISMS FOR INFINITIVAL ADJUNCT CLAUSES

I will restrict my attention to adjunct infinitive clauses, and consider examples like (110)-(117).

(110) John sneezed in order to attract attention.

(111) John is stubborn enough to talk to.

(112) John is too stubborn to talk to (Bill).

(113) John sneezed to attract attention.

(114) Mary knows the candidate to beat (Bill).

(115) Mary is pretty to look at.

In sentences (110)-(112), the adjunct infinitive phrase is signaled by the presence of one or more grammatical formatives. The lexicon contains the following information:

(116) *in order*:

complementizer
VP adjunct
CP(Inf)

(117) *enough*:

[-N, -V]
ADV, ADJ adjunct
CP(Inf)

(118) *too*:

[-N, -V]
ADV, ADJ specifier
CP(Inf)

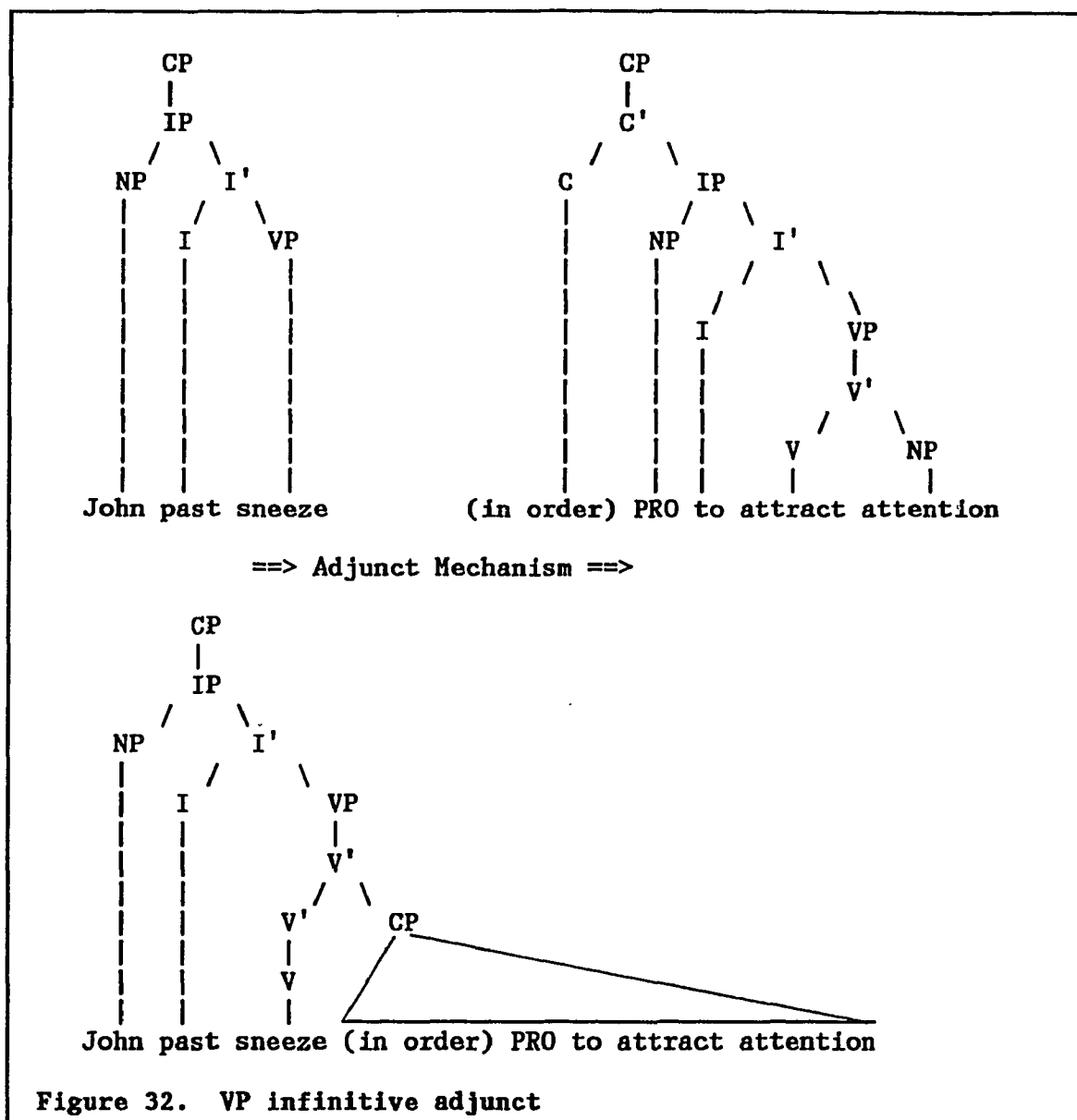
When the tree building mechanisms encounter one of these elements, the adjunct mechanism is invoked to create the structure to assign to the string. The adjunct mechanism adjoins the CP to the phrase marker without changing the bar level.

Suppose sentence (110) is presented to the tree building mechanisms.

(110) John sneezed in order to attract attention.

The tree building mechanism can assign the following structures in (110'), more fully specified in Figure 32 on page 269, to the string.

- (110') i. [CP John sneezed]
ii. [CP in order PRO to attract attention]



There are two separate structures: *John sneezed* is a CP, and *in order to attract attention* is a CP. Following the lexical entry for the complementizer *in order*, (110'ii) is an adverbial phrase that is an adjunct to VP.¹⁹ The adjunct mechanism joins the two trees in an adjunct

¹⁹ The node CP functionally is an adverb when the complementizer is *in order*. It answers the question *why?* I also assume that CP can

structure, shown in Figure 32. This structure is then operated on by the control module, which is discussed in "The Control Mechanisms for Adjuncts" on page 278.

Suppose sentence (111) is input to the tree building mechanisms.

(111) John is stubborn enough to talk to.

The lexical entry of *stubborn* does not license any complement. Further, no verb licenses an adverb in its lexical entry. An adverb, like *enough*, is an adjunct to an adjective or adverb. The tree building mechanisms assign the structures shown in Figure 33 on page 271 to the strings *John is stubborn* and *enough to talk to*. The assignment of structure to *John is stubborn* is straight-forward. Let us consider the steps by which the tree building mechanisms produce the phrase marker for *enough to talk to*.

The lexical entry for *enough* indicates it is an adverb with a CP complement. The entry for *talk* indicates it takes a PP internal argument. The tree building mechanisms must insert two empty categories since there are two missing NPs:

(111') [ADVP [ADV' [ADV enough] [CP PRO to talk to e]]]

Since the subject of *to* is ungoverned, it must be PRO. The empty category object of *talk to* can only be a variable bound to an operator in a non-A position since it has both case and a θ -role. Since there is no morphological operator, the *e* must be a variable bound to an empty

function as an NP when it is the subject, as in (c).

- (a) Why did John sneeze?
- (b) John sneezed in order to attract attention.
- (c) For Bill to make noise would annoy John.

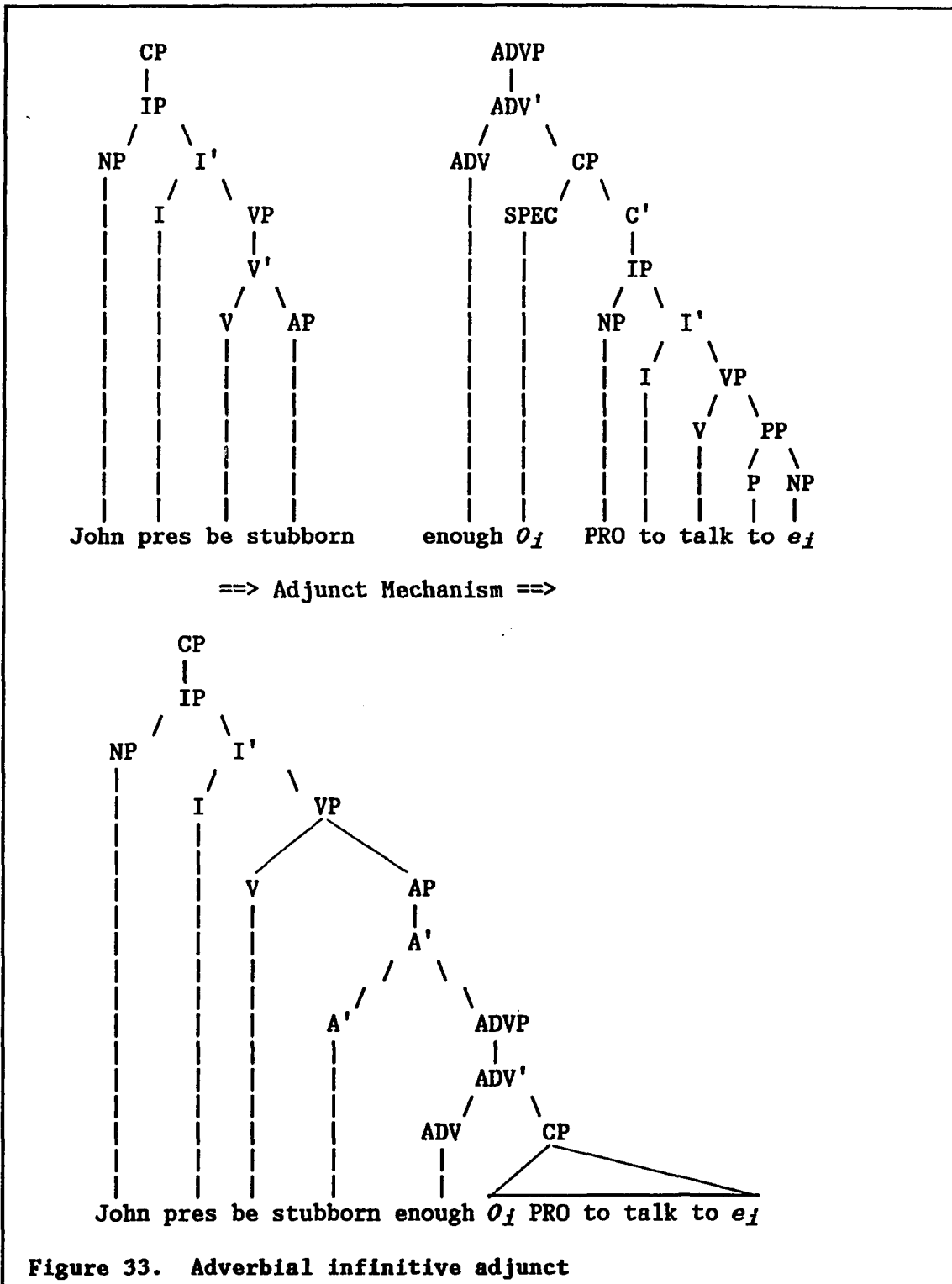


Figure 33. Adverbial infinitive adjunct

operator, O . The tree building mechanisms construct the phrase marker (111'') for the embedded clause.

(111'') [*ADVP* [*ADV'* [*ADV* enough] [*CP* O_i [*PRO* to talk to e_i]]]]]

The adjunct mechanism attaches the *enough* adverbial phrase to the adjunct position of the AP projection of *stubborn*. The output of the tree building mechanisms is shown in Figure 33. This output is then operated on by the control module.

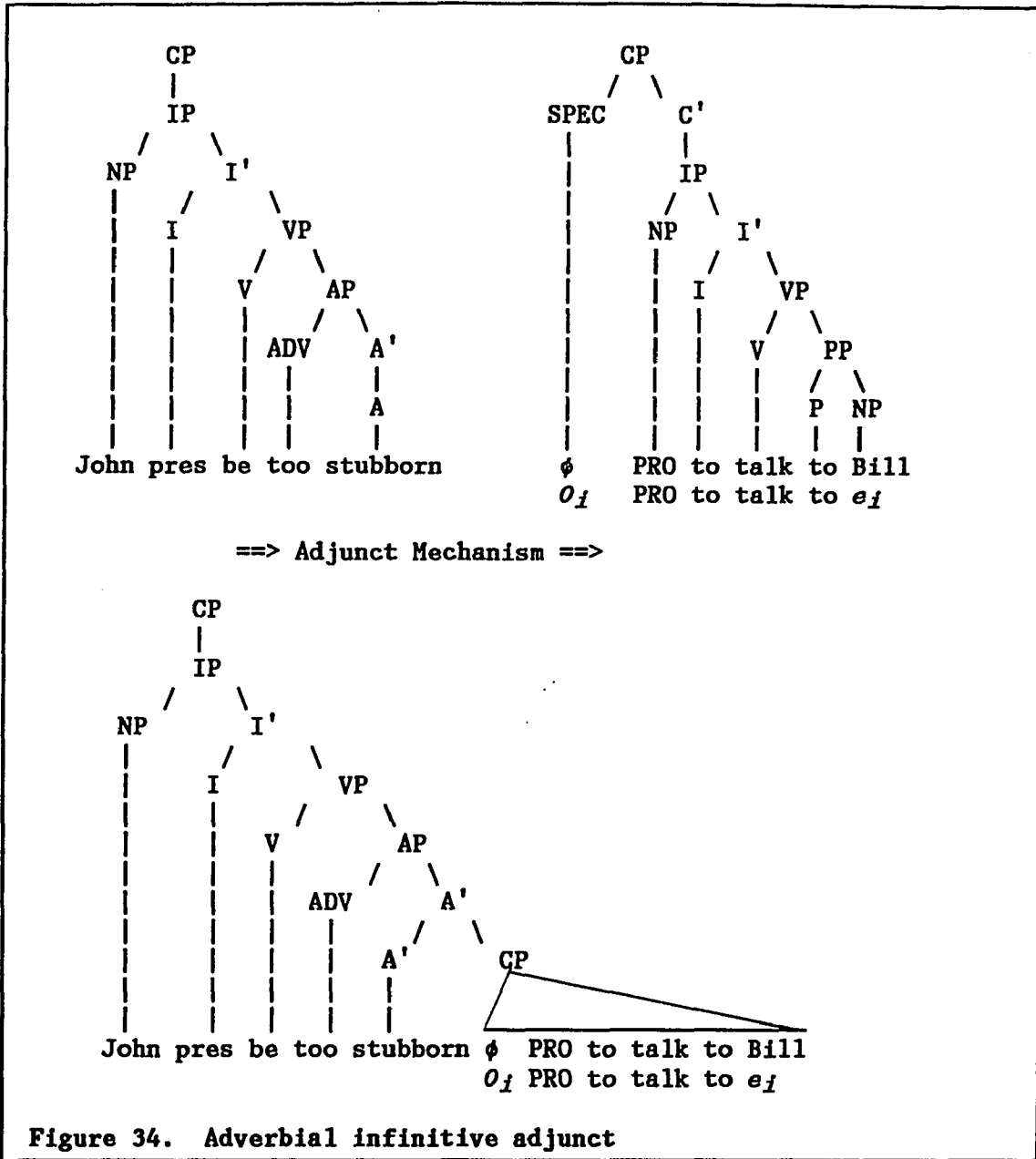
Suppose sentence (112) is input to the tree building mechanisms.

(112) John is too stubborn to talk to (Bill).

Since *stubborn* licenses no complements, the tree building mechanism will not analyze *to talk to Bill* as a complement of *stubborn*. The tree building mechanism, using the lexical entry for *too*, will construct the phrase markers in Figure 34 on page 273 for the string *John is too stubborn*, and for the strings *to talk to Bill* and *to talk to*.

The string *John is too stubborn* is analyzed as having the same structure as *John is very stubborn*, *John is quite tall*, etc., with *too* as an adverb which appears as the specifier of the adjective or adverb it modifies. The complement licensed by *too* appears in the adjunct position of the adjective or adverb modified by *too*.

In the string *to talk to Bill*, there is only one empty NP and it must be PRO. The string *to talk to* is assigned two empty categories. PRO is the specifier of IP. The internal argument of *talk to* is assigned case and a θ -role, hence, e must be a variable bound to an operator in COMP. Since there is no overt operator, the tree building mechanism inserts the empty operator, O , and coindexes O and e .



The lexical entry for *too* licenses the adjunction of the CP to the adjective or adverb modified by *too*. The adjunction mechanism combines the two structures and produces the phrase marker in Figure 34 on page 273. The control module then applies to this structure.

Suppose sentence (113) is input to the tree building mechanism.

(113) John sneezed to attract attention.

This is exactly the same as (110) except that the complementizer *in order* is not present in the infinitive phrase. The tree building mechanisms will assign two separate and unattached phrase markers to the strings: *John sneezed* and *to attract attention*. The adjunct mechanism places the infinitive clause as an adjunct on the V' expansion yielding the structure in Figure 32 on page 269.

Suppose the sentences represented in (114) are input to the tree building mechanism.

(114) Mary knows the candidate to beat (Bill).

The lexical entry for *candidate* does not license a complement. In addition, the lexical entry for *know* does not license NP and CP as internal arguments simultaneously, as is shown by the ill-formedness of (119iii), although either is fine on its own.

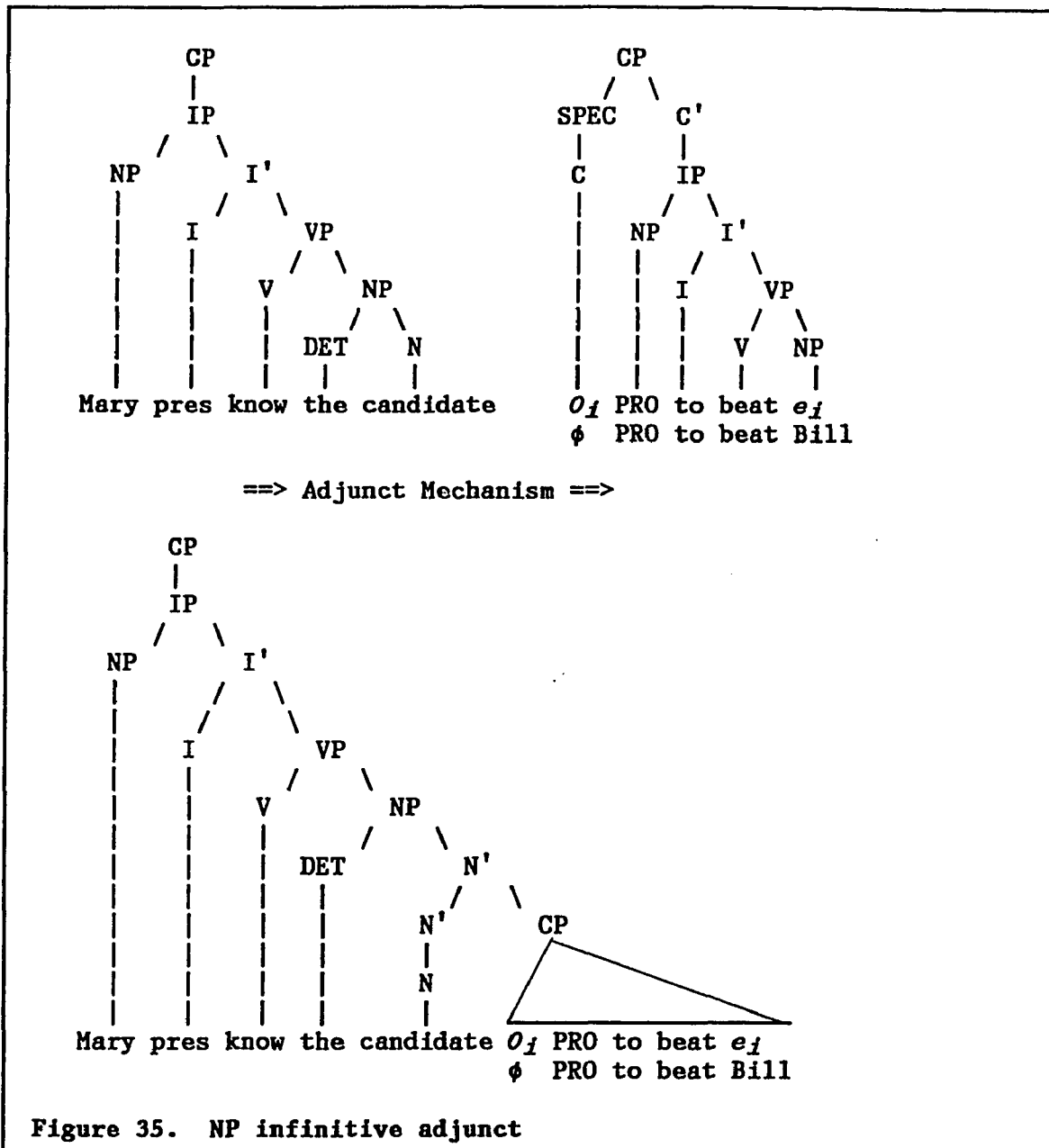
- (119) i. Mary knows John.
 ii. Mary knows to put out the cat.
 iii. * Mary knows John to put out the cat.

The tree building mechanisms assign the structures in (114') and (114'') to the strings: *Mary knows the candidate* and *to beat (Bill)*, illustrated in Figure 35 on page 275.

- (114') i. [CP Mary knows the candidate]
 ii. [CP PRO to beat Bill]

- (114'') i. [CP Mary knows the candidate]
 ii. [CP o_i PRO to beat e_j]

The adjunct mechanism can join the infinitive clause at two different adjunct positions. First, the adjunct mechanism could create a VP adjunct to *know* and add the infinitive clause as an *in order to* clause. I return to this possibility below. Second, the adjunct mechanism can



join the infinitive clause as an adjunct on the NP projection of *candidate*. This second possibility is illustrated in Figure 35 on page 275, where the adjunct mechanism creates an N' adjunct node, and attaches the CP infinitive clause to it.

In some cases, the adjunct mechanism will be able to attach the infinitive clause to either of two positions, perhaps as either a V adjunct or an N adjunct. In such cases, the adjunct mechanism will produce more than one parse. The tree building mechanisms will produce a different phrase marker for each of the possible ways of attaching an adjunct.

How will the tree building mechanism know which X-bar category to attach the infinitive clause to? The most natural assumption is that it attaches the clause everywhere it can, subject to certain limitations. For example, if *in order* is present in the COMP node, then the clause can only be a VP adjunct. However, the infinitive clause in (114) can be attached in more than one place, and the string may be ambiguous.

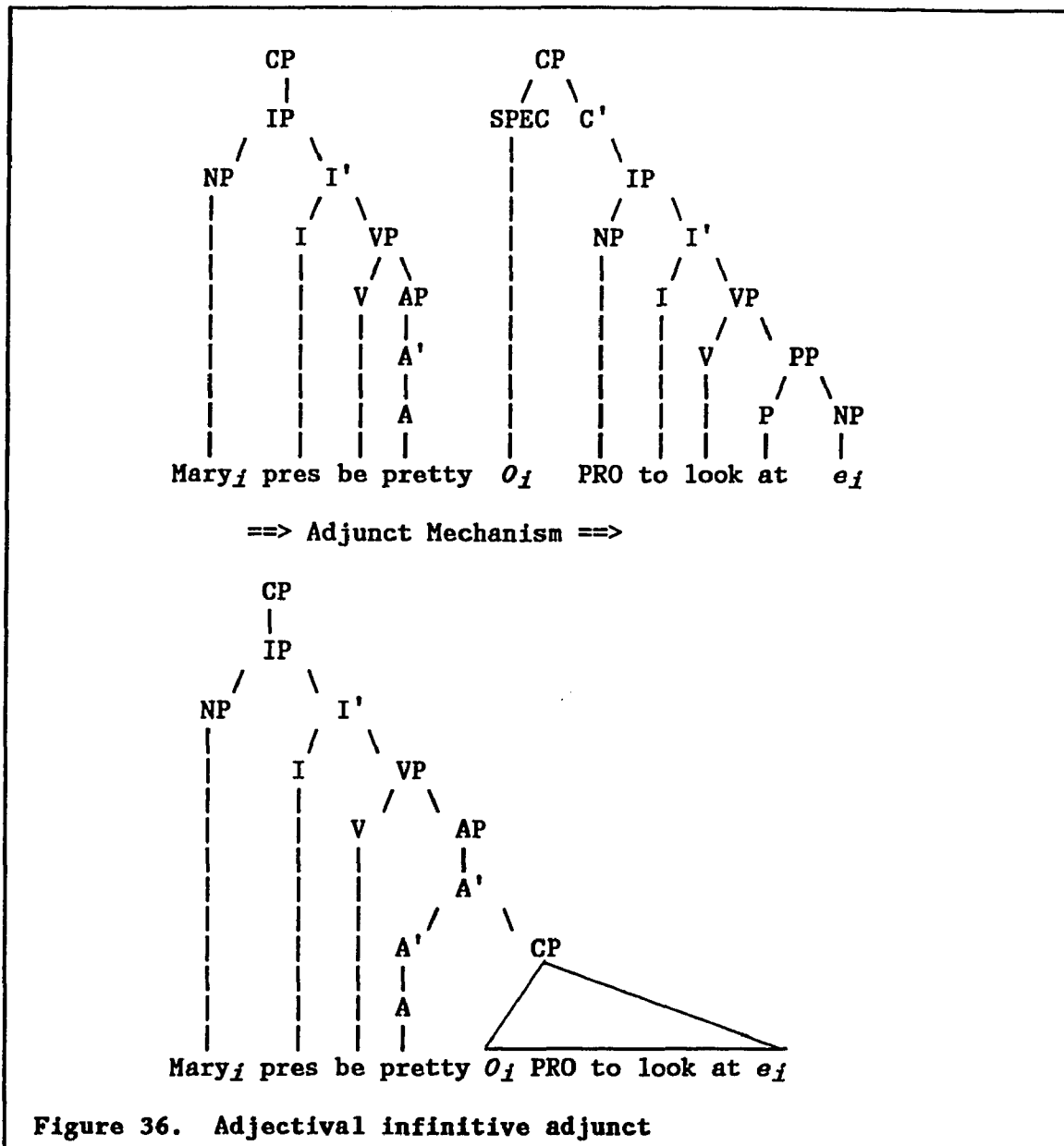
According to the control module, a string containing PRO is ill-formed if there are no features [\pm gen, \pm def] on PRO. If the infinitive sentences are attached where possible by the adjunct tree building mechanism, the control module will not be able to assign features to PRO in some cases, and these will be marked as uninterpretable by the control module. In other cases, PRO will be assigned features, and these will be interpretable.

Suppose the tree building mechanism inputs sentence (117).

(115) Mary is pretty to look at.

The lexical entry of the adjective *pretty* does not license a CP complement. Hence, the tree building mechanism will construct the two unattached phrase markers in (117') and in Figure 36 on page 277.

(115') i. [Mary is pretty]
 ii. [O_i PRO to look at e_j]



The X-bar projection of the copula *be* does not seem to license *in order to* clauses, as the examples in (120) show, so an *in order to* clause does not seem to be a possibility for the adjunct mechanism.

- (120) i. * Mary is pretty in order to get into the movies,
 ii. * John is sick in order to stay home from school.

The *why* question answer to a *be* sentence uses *because* not *in order to*:

Concerning distribution, PRO_{one} can occur in the adjunct infinitive clause on each category, as shown in (124).

- (124) i. V: John was shot (in order) PRO_{one} to keep him quiet.
 ii. N: John knows the candidate PRO_{one} to beat.
 iii. ADV: John is stubborn enough PRO_{one} to talk to.
 iv. ADV: John is too stubborn PRO_{one} to talk to.
 v. ADJ: Mary is pretty PRO_{one} to look at.


Exclusive-PRO can occur only in V, N, and ADV adjuncts. It does not occur on AP adjuncts. This is illustrated in (125).

- (125) i. V: John_i shot Bill (in order) PRO_i to keep him quiet.
 ii. N: John knows the candidate_i PRO_i to beat the incumbent.
 iii. ADV: John_i is stubborn enough PRO_i to talk to Bill.
 iv. ADV: John_i is too stubborn PRO_i to talk to Bill.
 v. ADJ: * Mary_i is pretty PRO_i to look at Sue.

The answers to questions (a)-(c) are simple in principle but complex in detail. Below I outline the basic principles underlying the assignment of control to PRO in adjuncts.

There are mechanisms in the grammar, unrelated to control, which isolate one NP from all the other NPs in the sentence. Let us consider two, which I call the N-Rel relation and the Agr-SR link.

The N-Rel relation is the relation between a relative clause and the nominal head it modifies, illustrated in (126). One might assume that N-Rel is an indexing relation that serves to indicate that the CP is related on a unique NP.

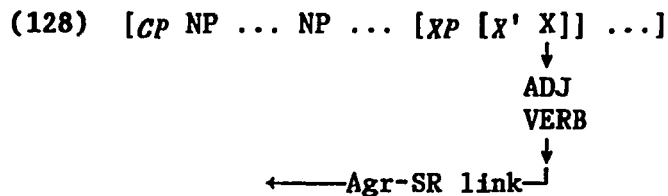
- (126) [NP [N' [N' N] [CP wh-RP_i [IP ... e_i ...]]]]

 N-Rel

In the sentences in (127), the N-Rel relation obtains between the noun head *woman* and the CP relative.

- (127) i. [NP the [N woman] [CP who beat the incumbent]]

- ii. [NP the [N woman] [CP who_i the incumbent beat e_i]]
- iii. [NP the [N woman] [CP PRO to beat the incumbent]]
- iv. [NP the [N woman] [CP O_i PRO_{one} to beat e_i]]

The Agr-SR link, illustrated in (128), refers to the relation between an adjective or a verb and the NP with which it must agree in person, number, and gender and/or with which it must satisfy any selection restrictions.



The following sentences are examples of the Agr-SR link. When the adjective does not agree with the designated NP, the sentence is ill-formed.

- (129) i. * Mary is alike.
 ii. Mary and Sue are alike.
 iii. Mary is like Sue.
- (130) i. Mary and Sue struck Bill as alike.
 ii. * Bill struck Mary and Sue as alike.
 iii. * Mary and Sue considered Bill alike.
 iv. Bill considered Mary and Sue alike.

Whatever mechanism determines that (129i) is ill-formed must utilize information about the relation, the Agr-SR link, between the noun *Mary* and the adjective *alike*. Whatever mechanism determines that (130ii) and (130iii) are ill-formed must utilize information about the Agr-SR link between the noun *Bill* and the adjective *alike* in (130ii) and (130iii). The Agr-SR link in (131) is between a noun and a verb.

- (131) i. * Mary collided.
 ii. Mary and Sue collided.
 iii. Mary collided with Sue.

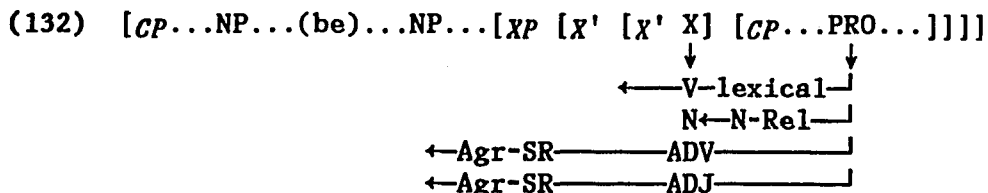
Whatever mechanism determines that (131i) is ill-formed must utilize information about the Agr-SR link between the noun *Mary* and the verb *collide*.

The Agr-SR link may be grammatically defined in terms of subject-predicate or it may be lexically defined. Some Agr-SR link exists for every adjective and every verb in a sentence, and the Agr-SR link on any particular adjective or verb points to a single NP in the sentence.

My description of control in adjuncts answers the questions (a)-(c) at the beginning of this section. In addressing these questions, I show when control in adjunct infinitive constructions can be reduced to either the N-Rel relation or the Agr-SR link. In these cases, the problem of assigning a referent to PRO is not one of control, it is the more general problem of determining what node APs and ADVPs are predicated of. Control in infinitival relative adjuncts is determined by the N-Rel mechanism that works for non-infinitival relatives. This is discussed in "Noun Infinitival Adjuncts: Infinitival Relatives" on page 306. Control in ADVP and AP infinitival adjuncts is determined by the Agr-SR link. This is discussed in "Adverbial Adjuncts: Degree Adverbial Phrases" on page 296 and "Adjectival Adjuncts: Copula Clauses" on page 310. Control in *in order to* infinitival VP adjuncts involves (a) information about the lexical entry of the V head, and (b) information about the presence or absence of *be+en* in the matrix sentence. I propose a VP adjunct control mechanism in "VP Adjuncts: 'in order to' clauses" on page 283.

To summarize, in the structure (122), there are three mechanisms which determine which noun will be the antecedent of PRO: (1) lexical

processes for *in order to* VP adjuncts, (2) the N-Rel relation for infinitival relative adjuncts, and (3) the Agr-SR link for ADVP and AP adjuncts. These processes are illustrated in (132).



When adjuncts are in the VP, in *in order to* clauses, control is lexical. Coindexing is a function of the lexical entry of the verb and the presence of the passive morphology (*be+en*) in INFL. When adjuncts are in the NP, as are infinitival relative adjuncts, control is determined by the N-Rel relation. Coindexing is always with the head of the NP containing the adjunct clause. When adjuncts are in ADV and ADJ degree and copula clauses, coindexing is a function of the Agr-SR link.

The fact that there are only two possible interpretations for PRO in adjuncts, PRO_{one} and PRO[-gen,+def], follows from the proposed adjunct control mechanisms. In adjuncts, when PRO is coindexed, it must be coindexed with a single node, and it does not have reference outside of the sentence. This follows from the properties of the mechanisms which assign the coindexing. The lexical mechanism coindexes PRO with the NP that is the external argument of the verb, and this is a single node. The N-Rel mechanism coindexes PRO with the noun which is the head of the NP in which the adjunct is attached, and this is a single node. The Agr-SR mechanism coindexes PRO in an AP or ADVP adjunct with the noun that meets selection and agreement requirements of the adjective or verb which the adverb modifies, and this is a single NP node.

7.6 VP ADJUNCTS: 'IN ORDER TO' CLAUSES

The bracketed structure (134) represents the general configuration for infinitive VP adjuncts.

- (134) [_{CP} NP (be) [_{VP} [_{V'} [_{V'} V(+en) (P (NP)) (CP)] [_{CP} (in order) PRO to VP]]]]

When the configuration represented in (134) is input to the control module, there are three possible outputs: (I) exclusive-PRO, (II) PRO_{one}, and (III) no extrinsic feature assignment (where PRO is uninterpretable).

Case I: PRO is assigned the features of exclusive-PRO, PRO_[-gen,+def], and coindexed with the external argument of the matrix verb (V_m), as in the sentences in (135)-(137).

- (135) John_i sneezed [(in order) PRO_i to attract attention]
 (136) John_i shot Bill [(in order) PRO_i to keep himself busy]
 (137) John_i operated on Bill [(in order) PRO_i to save his life]

Case II: PRO is assigned the features of PRO_{one} ([+gen,-def]), as in examples (138)-(141).

- (138) Bill was shot by John [(in order) PRO_{one} to keep him quiet]
 (139) Bill was shot [(in order) PRO_{one} to keep him quiet]
 (140) Bill was operated on [(in order) PRO_{one} to save his life]
 (141) Bill was operated on by John [(in order) PRO_{one} to save his life]

Case III: PRO is not assigned any features by the control module. In this case PRO is uninterpretable and the sentence is ill-formed.

- (142) * For John to smoke would annoy Bill in order to get on his nerves.
 (143) * For you to be rude to everyone stinks in order to lose friends.

The VP-adjunct mechanism for assigning control in structure (134) is (144).

(144) VP-adjunct mechanism for control of PRO:

WHEN $[+V]_m$ licenses an external argument in its lexical entry,

IF CP_m contains an external argument (does not contain *be+en*),

THEN Assign the features $[-gen,+def]$ to PRO and coindex $PRO[-g,+d]$ with the external argument. (CASE I)

OTHERWISE Assign PRO_{one} . (CASE II)

OTHERWISE Assign no coindexing or features to PRO, which will result in the PRO being uninterpreted by the control module. (CASE III)

In the fragments of lexical entries below, two verbs license an external θ -role (*shoot* and *operate*), and two verbs do not license an external θ -role (*annoy* and *stink*). *Shoot* and *operate* do not license CP complements, and so have no specification for control or extrinsic feature assignment. *Annoy* and *stink* do license CP complements, and so they specify extrinsic feature assignment. Since *stink* assigns no θ -role, the controller cannot be specified in its lexical entry. Only *annoy* specifies the controller because it alone assigns both a θ -role and a complement clause. All of these verbs can be assigned an *in order to* VP adjunct by the tree building mechanisms.

(145) *shoot*:

$[+V,-N]$
external θ -role
internal arguments: NP

(146) *operate*:

$[+V,-N]$
external θ -role
internal arguments: PP_{on}

(147) *annoy*:
 [+V,-N]
 internal: NP, CP
 control: internal, PRO[-g,+d]

(148) *stink*:
 [+V,-N]
 internal: CP
 control: [+g,-d]

Consider the parses of sentences (135), (136), (138), and (142).

(135) John sneezed (in order) PRO to attract attention.

(136) John shot Bill (in order) PRO to keep himself busy.

(138) Bill was shot (in order) PRO to keep him quiet.

(142) * For John to smoke would annoy Bill (in order) PRO to get on his nerves.

String (135) has the structure (135').

(135') [John [_{VP} [_{V'} [_{V'} sneezed] [_{CP} (in order) PRO to attract attention]]]]

V_m is *sneeze*, which licenses an external argument in its lexical entry. CP_m does not contain *be+en*, so (135') is Case I of the VP-adjunct mechanism, and PRO is assigned the features [-gen,+def] and coindexed with the external argument of *sneeze*, which is *John*.

(135a) [John_i [_{VP} [_{V'} [_{V'} sneezed] [_{CP} (in order) PRO[-g,+d]_i to attract attention]]]]

The tree building mechanism, which includes the binding module, assigns string (136) the structure in (136').

(136') [John [_{VP} [_{V'} [_{V'} shot Bill] [_{V'}(in order) PRO_i to keep himself_i busy]]]]

V_m is *shoot*, and *shoot* licenses an external argument. CP_m does not contain *be+en*. This is Case I of the VP-adjunct mechanism, so PRO is assigned the features of exclusive-PRO and coindexed with *John*. Since

PRO already has an index from binding, *John* is assigned the index of PRO yielding:

(136a) [$John_i$ [VP [V' [V' shot Bill] [V' (in order) PRO_{&lrbk.-g,+d}] i to keep himself $_i$ busy]]]

The tree building mechanism assigns string (138) the structure in (138').

(138') [$Bill_i$ was [VP [V' [V' shot e_i] [CP (in order) PRO to keep him quiet]]]

V_m is *shoot*, which licenses an external argument. CP_m does not contain an external argument, CP_m contains *be+en*. This is Case II of the VP-adjunct mechanism. The features of PRO_{one}, ([+gen,-def]), are assigned to the subject of the infinitive, as in (138a).

(138a) [$Bill_i$ was [VP [V' [V' shot e_i] [CP (in order) PRO_{one} to keep him quiet]]]

The tree building mechanism assigns string (142) the structure in (142').

(142') [[for John to smoke] would [VP [V' [V' annoy Bill] [CP (in order) PRO to get on his nerves]]]]

V_m is *annoy*, and *annoy* does not license an external θ -role. Hence, this is case III of the VP-adjunct mechanism, and PRO is uninterpretable. The structure is marked as ill-formed by the control module.

7.6.1 THE DOMAIN OF THE VP-ADJUNCT MECHANISM

In sentence (149), PRO can be either *the doctor* or *the patient*, depending upon whether the *in order to* adjunct is attached to the VP headed by *wonder* or the VP headed by *sign*. In (150)-(151), *or not* indicates the end of the embedded clause, as in *whether ... or not*. Thus, in (150), the adjunct is attached to the VP headed by *wonder*, and so PRO can only

be *the doctor*. In (151), the adjunct is attached to the embedded clause, and PRO is *the patient*.

- (149) The doctor wondered whether or not the patient had to sign the form (in order) to collect the insurance money.
- (150) The doctor wondered whether the patient had to sign the form or not (in order) to collect the insurance money.
- (151) The doctor wondered whether the patient had to sign the form (in order) to collect the insurance money or not.

In sentence (152), PRO can be only *the doctor*, but in (153), PRO can be *the doctor or the patient*.

- (152) In order to collect the insurance money, the doctor wondered if the patient had to sign the form.
- (153) The doctor wondered if the patient had to sign the form in order to collect the insurance money.

Control in (153), where the *in order to* adverbial is in topic position, follows if the mechanisms of the control module incorporate the assumption that the *in order to* phrase can be anywhere in the matrix clause (CP_m). It does not have to be in the VP adjunct position adjacent, to the right of, the head of the VP.

7.6.2 ADJUNCTS AND COMPLEMENTS IN THE VERB PHRASE

Some verbs, like *count on*, optionally have a CP complement. With such verbs, an infinitive clause is ambiguously either a complement or an *in order to* VP adjunct.

(154) *count on*

external θ -role
 internal arguments: PP_{ON}, (CP)
 control: internal argument, [-g,+d]

(155) John counted on Bill to arrive on time.

(156) John counted on Mary to get herself there on time.

(157) John counted on Mary (in order) to get himself here on time.

- (158) John counted on Mary to get herself there on time (in order) to get himself here on time.

The tree building mechanisms can assign sentence (155) two different structures: (159) and (160).

- (159) [John [_{VP} [_{V'} counted on Bill [_{CP} PRO to arrive on time]]]]

- (160) [John [_{VP} [_{V'} [_{V'} counted on Bill] PRO to arrive on time]]]]

In (159), the CP is licensed by the lexical entry and PRO must be coindexed with the object. The complement does not have an *in order* to interpretation. The control module, using the lexical information for *count on*, coindexes PRO and the internal argument, adding the features [-g,+d], to yield (159').

- (159') [John [_{VP} [_{V'} counted on Bill]_i [_{CP} PRO[-g,+d]_i to arrive on time]]]]

Since the lexical entry of *count on* indicates that the embedded CP is optional, the tree building mechanisms can construct (160), which puts the CP as a VP adjunct, and not in the complement position. The bracketed structure in (160) meets the structural description (134) for a VP-adjunct. V_m is *count on*, and this verb lexically licenses an external θ -role. CP_m does not contain *be+en* so this is Case I of the VP-adjunct mechanism, and PRO is coindexed with the external argument of *count on*, *John*, as in (160').

- (160') [John]_i [_{VP} [_{V'} [_{V'} counted on Bill] [PRO[-g,+d]_i to arrive on time]]]]

Any verb which optionally licenses a CP as an internal argument permits an ambiguity with an infinitive phrase. One structure has the CP as a verbal complement, and the other has the CP as a VP adjunct. In some situations, agreement phenomena can rule out one structure, and

in these cases, the infinitive construction is unambiguous. This happens in the following strings:

- (161) i. They asked me (in order) to amuse themselves.
ii. They asked me to amuse them. they=them
- (162) i. Mary shouted to Bill (in order) to make herself heard.
ii. Mary shouted to Bill to make himself heard.
- (163) i. Mary shouted to Bill (in order) to make him hear her.
ii. Mary shouted to Bill to make her hear him.
- (164) i. Mary appealed to Bill (in order) to get herself closer to him.
ii. Mary appealed to Bill to get himself closer to her.
- (165) i. Mary invited Bill (in order) to ingratiate herself to him.
ii. Mary invited Bill to ingratiate himself to her.

7.6.3 THE THEORY OF IMPLICIT CONTROLLERS

Chomsky, among others, has attempted to motivate control by an implicit argument or a missing argument in certain structures. This argument must presumably either be represented syntactically or represented only in the lexical structure of the verb. In either case, it is not clear how this kind of control would work. In this section, I question the Chomsky (1986a) motivation for implicit arguments.

Chomsky (1986a) discusses subjects in the following terms:

Note that, in principle, there are three ways in which a subject may fail to appear overtly in some clausal structure: (i) it may be syntactically realized as an empty category, (ii) it may be realized as a constituent of the V head of the VP predicate, which assigns it a θ -role; (iii) it may be missing in both syntactic and lexical representations. In fact, all three conditions are realized and they have distinctively different properties.

The three possible cases are illustrated in <122>, where in (ii), *e* is the trace of *the boat*:

- <122> (i) I decided [PRO to sink the boat]
- (ii) the boat was sunk *e*
- (iii) the boat sank

Take these to be both S-structure and LF representations.
(Chomsky 1986a:118)

Chomsky proceeds to analyze the difference between the constructions in terms of adverb and complement constructions. He concludes:

Thus, the three possible cases exist and are distinguished from one another. In particular, there is evidence for the syntactic presence of PRO and the lexical presence of a 'missing argument'. (Chomsky 1986a:119)

I agree that there is evidence for the syntactic presence of PRO, however, contrary to Chomsky, I argue that there is no evidence for the lexical presence of a missing argument. The evidence which he offers that is relevant to this discussion is that relating to control by the missing argument. Chomsky writes:

Consider control *by* the understood subject:

- <126> (i) it is time [PRO to sink the boat [PRO to collect the insurance]]
 (ii) the boat was sunk [PRO to collect the insurance]
 (iii) *the boat sank [PRO to collect the insurance]

An element present at the syntactic or lexical level can serve as a controller. In (iii) the sentence can be given only the senseless interpretation with *the boat* understood as the subject of *collect*. (Chomsky 1986:119)

This distribution of data does not provide motivation for lexically present missing arguments. Instead, these facts follow from the VP adjunct mechanisms. It should be noted that, in the cases where Chomsky posits an implicit controller, the implicit controller is always animate and indefinite. That is, it corresponds exactly to PRO_{one}.

The lexical entry for the noun *time* indicates that it licenses a CP complement with PRO_{one} as subject; licensing sentences like (166) and (167).

(166) [*NP* the time [PRO_{one} to sink the boat]] is near.

(167) It is [*NP* time [*PRO*_[+g,-d] to sink the boat]]

If there is an *in order to* adjunct clause in the string, the tree building mechanisms can assign the following structures:

(168) the time [*PRO*_[+g,-d] to sink the boat [(in order) *PRO* to collect the insurance] is near

(169) it is time [*PRO*_[+g,-d] to sink the boat [(in order) *PRO* to collect the insurance]]

In each case, V_m is *sink*, which licenses an external θ -role. CP_m does not contain *be+en*. This is Case I of the VP-adjunct mechanism, and *PRO* in the adjunct clause is coindexed with *PRO*_{one} of the matrix clause, as in (170)-(171).

(170) the time [*PRO*_[+g,-d]_{*i*} to sink the boat [(in order) *PRO*_[-g,+d]_{*i*} to collect the insurance]] is near

(171) it is time [*PRO*_[+g,-d]_{*i*} to sink the boat [(in order) *PRO*_[-g,+d]_{*i*} to collect the insurance]]

Let us now look at (172) where Chomsky posits a lexically present implicit subject.

(172) The boat was sunk (in order) *PRO* to collect the insurance.

The tree building mechanisms assign structure (172') to (172).

(172') [the boat_{*i*} was [*VP* [*V'* [*V'* sunk *e_i*] [*CP* (in order) *PRO* to collect the insurance]]]]

V_m is *sink*, and CP_m contains *be+en*. Hence, this is Case II, and *PRO* is marked *PRO*_{one} by the VP-adjunct mechanism, as in (173).

(173) [The boat_{*i*} was [*VP* [*V'* [*V'* sunk *e_i*] [*CP* (in order) *PRO*_[-g,+d] to collect the insurance]]]]

Let us consider the problem of "who sank the boat" and "who collects money" in more detail. In sentences (175i)-(178i), Case I of the VP-adjunct mechanism coindexes the matrix subject with the adjunct *PRO*. This coindexation implies (175ii)-(178ii). Thus it follows from Case

I of the VP-adjunct mechanism that if (i) is non-deviant then so is (ii). Likewise, if (ii) is deviant, then so is (i), and they are deviant in the same way.

- (174) i. [NP_i [VP V ... [(in order) PRO_i to VP]]]
 ii. [NP_i [IP pres VP]]
- (175) i. * Torpedoes sank the boat in order to collect the insurance.
 ii. * Torpedoes collect the insurance.
- (176) i. Frogmen sank the boat in order to collect the insurance.
 ii. Frogmen collect the insurance.
- (177) i. The second mate sank the boat in order to collect the insurance.
 ii. The second mate collects the insurance.
- (178) i. The owners sank the boat in order to collect the insurance.
 ii. The owners collect the insurance.

Since Case II of the VP-adjunct mechanism assigns PRO_{one}, which is not assigned an antecedent, in Case II of the VP-adjunct mechanism, if (i) is true, (ii) may be true, but it may not be. In the structure (179), (ii) is not implied by (i).

- (179) i. [... NP ... [(in order) PRO_{one} to VP]]
 ii. [... NP ... [IP pres VP]]

Therefore, in (180)-(182), which are all Case II, if (i) is true, (ii) is not implied, although it may be true. (iii) is not implied by (i), instead (iii) follows from general knowledge of insurance, boats, and their owners.

- (180) i. The boat was sunk by torpedoes in order to collect the insurance.
 ii. * Torpedoes collect the insurance.
 iii. The owners collect the insurance, thanks to the torpedoes.
- (181) i. The boat was sunk by frogmen in order to collect the insurance.
 ii. Frogmen collect the insurance.
 iii. The owners collect the insurance, thanks to the frogmen.

- (182) i. The boat was sunk by the second mate in order to collect the insurance.
 ii. The second mate collects the insurance.
 iii. The owners collect the insurance, thanks to the second mate.
- (183) i. The boat was sunk by the owners in order to collect the insurance.
 ii. The owners collect the insurance.
 iii. The owners, and no one else, collect the insurance.

The VP-adjunct mechanism assumes that PRO in (180)-(183) is the indefinite empty pronoun *one*, PRO_{one}. The above suggests that, in this configuration, PRO_{one} can be understood to refer to anyone in the world, including any persons referred to in the sentence. Common sense in language use, coupled with some knowledge about insurance, leads to inferred reference in (180iii)-(183iii). This is quite different from (175)-(178), where the person who collects the insurance is asserted.

Another example of Case I of the VP-adjunct mechanism is (184) (Chomsky's <126iii>).

- (184) The boat [_{VP}[_{V'}[_{V'} sank] [(in order) PRO to collect the insurance]]]

The verb *sink* assigns an external θ -role, and CP_m does not contain *be+en*. Therefore, this is Case I and PRO is coindexed with *boat* yielding (185i). Sentence (i) should be as acceptable as (ii).

- (185) i. * The boat_i [_{VP} [_{V'} [_{V'} sank] [(in order) PRO_i to collect the insurance]]]
 ii. * The boat collects the insurance.

Sentence (185i) is ill-formed for the same reason that the sentences in (186) are ill-formed.

- (186) i. * The boat collected the insurance.
 ii. * It was the boat that collected the insurance.
 iii. * The boat that collected the insurance sank.
 iv. * The insurance was collected by the boat.

The problem with (185i) is not that an agent is missing, it is rather that the sentence implies (185ii) and (186), that the boat collected the insurance. The fact that (185ii) is asserted to be true by (185i) follows from the assumptions of the theory of adjuncts.

Note, in addition, that sentences with VP-adjuncts are quite acceptable without any missing subject if the *in order to* phrase makes sense.

- (187) Aristotle said that a rock falls (in order) PRO to reach the center of the earth.
- (188) A rock falls (in order) PRO to reach the center of the earth.
- (189) A boat sinks (in order) PRO to reach Davy Jones' locker.
- (190) The sun sets (in order) PRO to allow the little birds PRO to sleep.

Chomsky also tries to motivate missing arguments by the distribution of agent-oriented adverbs. He states:

... Structures like (i) and (ii) are distinguished from (iii) by the possibility of adding "agent-oriented" adverbs such as "voluntarily":

- <124> (i) I decided [PRO to leave voluntarily]
- (ii) the boat was sunk voluntarily
- (iii) *the boat sank voluntarily

These adverbs require an agent, which may be expressed syntactically as in (i) or lexically as in (ii). Notice that in (i), *voluntarily* is associated with PRO, not its binder *I*, as it would be in "I decided [PRO to leave] voluntarily." In the latter, it is my decision that is voluntary, whereas in (i) it is my leaving. (Chomsky 1986a:118-9)

I want to note here that the principles which determine the antecedent of PRO in VP-adjuncts differs significantly from the mechanism which associates adverbs like *voluntarily* with a referent, perhaps an agent.

For example, in (191) PRO is *John*, but in (192) PRO is not *Bill*, it is PRO_{one}. PRO cannot be coindexed with *Bill* in either (191) or (192).

(191) John_i shot Bill [(in order) PRO_i to keep himself busy]

(192) Bill was shot [(in order) PRO_{one} to keep him quiet]

On the other hand, *voluntarily* in (193) refers to *John* and in (194) to *Bill*. In (195), *voluntarily* can refer to either *John* or *Bill*.

(193) John shot Bill voluntarily.

(194) Bill was shot voluntarily.

(195) Bill was shot by John voluntarily.

Insofar as one agrees with these intuitions, the principles which interpret PRO in VP-adjuncts are different than those which associate adverbs like *voluntarily* with an NP.

A final point concerns the interaction of *in order to* clauses with long-distance control, as in (196).

(196) John was given immunity in order to testify against himself.

This has the structure (196').

(196') John was [_{VP} [_{V'} [_V'given immunity] in order PRO to testify against himself]]]

V_m is *give*, which assigns an external θ -role. CP_m contains *be+en*. Hence, this is Case II of the VP-adjunct mechanism, and PRO[+g,-d] is assigned.

(196a) John was [_{VP} [_{V'} [_V'given immunity] in order PRO[+g,-d] to testify against himself]]]

However, binding requires that *himself* is coindexed with PRO[+g,-d], which converts [-def] to [+def] as discussed under long-distance binding in "Long-Distance Control" on page 249.

(196b) John was [_{VP} [_{V'} [_{V'}given immunity] in order PRO_[+g,+d]_i to testify against himself_j]]]

This is a case of long-distance control. PRO_[+gen,+def,-pl,-fem] is indexed with the only suitable NP in the sentence: *John*.

7.7 ADVERBIAL ADJUNCTS: DEGREE ADVERBIAL PHRASES

Degree adverbial infinitive phrases have the structures in (197) and (199). In (197) the infinitive phrase modifies an adjective, as in (198), and in (199), an adverb, as in (200).

(197) [_{CP} NP V (NP) ... (too) ADJ (enough) [_{CP} PRO to VP]]

(198) i. John is stubborn enough to talk to.

ii. John is too stubborn to talk to.

(199) [_{CP} NP V (NP) ... (too) ADV (enough) [_{CP} PRO to VP]]

(200) i. John ran quickly enough to catch Bill.

ii. John ran too slowly to catch Bill.

Since adjuncts are, by definition, optional, if sentences with structures (197) and (199) are well-formed, then the corresponding sentences with the structures (201) and (203) are also well-formed. Since (198) and (200) are well-formed, (202) and (204) must be well-formed. This fact will be significant in determining the referent of PRO in adjectival and adverbial adjuncts.

(201) [_{CP} NP V (NP) ... ADJ]

(202) John is stubborn.

(203) [_{CP} NP V (NP) ... ADV]

(204) John ran quickly/slowly.

There are three possible outputs for the assignment of features to PRO in adverbial infinitive adjuncts: Exclusive-PRO (PRO_[-g,+d]), PRO_{one} (PRO_[+g,-d]), and inclusive-PRO (PRO_[+g,+d]).

Case I: PRO can be assigned $PRO[-g,+d]$ (or exclusive-PRO) and be coindexed with an NP in the *mātrix* clause (CP_m), as in the bracketed structure (205), and in sentences (206)-(209).

- (205) [CP NP_{*i*} ... (too) ADJ/ADV (enough) [CP $PRO[-g,+d]_i$ to VP]]
 (206) John_{*i*} is stubborn enough [$PRO[-g,+d]_i$ to talk to Bill]
 (207) John_{*i*} is too stubborn [$PRO[-g,+d]_i$ to talk to Bill]
 (208) John_{*i*} runs quickly enough [$PRO[-g,+d]_i$ to catch Bill]
 (209) John_{*i*} runs too slowly [$PRO[-g,+d]_i$ to catch Bill]

Case II: PRO is assigned PRO_{one} , as in the bracketed structure (210), illustrated by sentences (211)-(214).

- (210) [CP ... (too) ADJ/ADV (enough) [CP^0 [PRO_{one} to VP]]]
 (211) John is stubborn enough [O_i [$PRO[+g,-d]$ to talk to e_i]]
 (212) John is too stubborn [O_i [$PRO[+g,-d]$ to talk to e_i]]
 (213) John runs slowly enough [O_i [$PRO[+g,-d]$ to catch e_i]]
 (214) John runs too quickly [O_i [$PRO[+g,-d]$ to catch e_i]]

Case III: In cases of long-distance control, PRO gets the features of Inclusive-PRO. PRO_{one} is given the features [+def, α -pl, β -fem] by virtue of being coindexed with a definite reflexive (*himself, herself, themselves*). PRO is $PRO[+g,+d]$, as in the bracketed structure (215), and illustrated by sentences (216)-(219).

- (215) [CP NP ... (too) ADJ/ADV (enough) [CP O_j [$PRO[+g,+d]_i$ to ... pro-self_{*i*} ... e_j]]]
 (216) Mary found Bill stubborn enough [O_j [$PRO[+g,+d]_i$ to bother herself_{*i*} with e_j]]
 (217) Mary found Bill too stubborn [O_j [$PRO[+g,+d]_i$ to bother herself_{*i*} with e_j]]
 (218) Mary found John spoke slowly enough [O_j [$PRO[+g,+d]_i$ to bother herself_{*i*} with e_j]]

(219) Mary found John spoke too quickly [0_j [PRO[+g,+d] $_i$ to bother herself $_i$ with e_j]]

For the purposes of this discussion, we must distinguish between direct and indirect control mechanisms. A direct control mechanism specifies or points directly to the NP that is to be coindexed with the PRO. Examples of direct control mechanisms are (1) the lexical assignment of control discussed in "Lexically Specified Anaphora" on page 237, where it is indicated in the lexical entry of a verb which NP is coindexed with PRO, and (2) the VP-adjunct mechanism of "VP Adjuncts: 'in order to' clauses" on page 283, that specifies which NP is to be coindexed with PRO. An indirect control mechanism instead specifies, or points to, a category which, in turn, specifies or points to the NP referent. As we will see indirect control mechanisms can also specify a category which, in turn, specifies another category which points to the referent.

7.7.1 THE MECHANISM OF INDIRECT CONTROL: CASE I

In structure (205), the NP that is coindexed with PRO is determined by indirect control mechanisms.

(205) [CP NP $_i$... (too) ADJ/ADV (enough) [CP PRO[-g,+d] $_i$ to VP]]

The two basic relations of the indirect control mechanism are the Agr-SR link which was discussed in "The Control Mechanisms for Adjuncts" on page 278, and the ADV-modifier link. Both of these relations are defined below.

The Agr-SR link is the relation between an adjective or verb and the NP with which it must agree in person, number, and gender and/or with which it must satisfy any selectional restrictions.

The ADV-modifier link is a relationship between an adverb (e.g. *too*, *enough*, *quickly*, *slowly*) and the adjective, adverb, or verb it modifies. In order to be licensed, every adverb must be linked to the adjective, adverb, or verb that it modifies.

The Agr-SR link obtains between *John* and *stubborn* in (220) and between *John* and *runs* in (221). The ADV-modifier link obtains between *runs* and *quickly* in (221).

(220) John is stubborn.
 └─Agr-SR─┘

(221) John runs quickly.
 └Agr-SR┘ └ADV-mod┘

The ADV-modifier link obtains between the adverbial *enough...[PRO to talk to Bill]* and the adjective *stubborn* in (222), and between the adverbial *enough...[PRO to catch Bill]* and the adverb *quickly* in (223).

(222) John is stubborn enough [PRO to talk to Bill]
 └─ADV-mod─┘

(223) John runs quickly enough [PRO to catch Bill]
 └─ADV-mod─┘

The mechanism for control in adverbial-adjuncts (ADV-adjuncts) is stated in (224).

(224) The indirect control mechanism of ADV-adjunct infinitival clauses.

- i. If the adverbial CP is in an AP, then coindex PRO with the NP pointed to by the Agr-SR link for that adjective.
- ii. If the adverbial CP is in an ADVP, and the head of the ADVP is linked to a verb via an ADV-modifier link, then coindex PRO with the NP pointed to by the Agr-SR link for that verb.

An example of (224i) is sentence (207), which is assigned the structure (207') by the tree building mechanisms.

7.7.2 THE INDIRECT REFERENT MECHANISM: CASE II

The tree building mechanisms can insert an empty operator to bind an empty category in a position which is assigned case and a θ -role. The tree building mechanisms assign a structure like (225) to sentences (211)-(214).

(225) [*CP* ... (too) ADJ/ADV (enough) [*CP* *O_i* [PRO to V ... *e_i*]]]

The *wh*-PRO mechanism of the control module, see "Feature Assignment by Grammatical Formatives" on page 254, assigns PRO the features of PRO_{one} when PRO follows a *wh*- element in COMP. The *wh*-PRO mechanism converts (225) to (226), hence, sentences (211)-(214) all contain PRO_{one}.

(226) [*CP* ... (too) ADJ/ADV (enough) [*CP* *O_i* PRO_{one} to V ... *e_i* ...]]]

In order to be licensed, the empty operator must be or strongly bound by an NP in the higher sentence. The question arises: What principles operate in coindexing an NP with the empty operator?

In adverbial adjunct infinitive clauses, the principles which determine the referent of the empty operator seem to be the same as those which determine the referent of PRO. The control mechanism (224) can be generalized to the referent mechanism (227).

(227) The indirect referent mechanism of ADV-adjunct infinitival clauses.

- i. If the adverbial CP is in an AP, then coindex PRO/*O* with the NP pointed to by the Agr-SR link for that adjective.
- ii. If the adverbial CP is in an ADVP, and the head of the ADVP is linked to a verb via an ADV-modifier link, then coindex PRO/*O* with the NP pointed to by the Agr-SR link for that verb.

7.7.3 INDIRECT REFERENCE AND LONG-DISTANCE CONTROL: CASE III

In cases where the adverbial complement contains both an empty operator and a reflexive pronoun bound to PRO, as in (216)-(219), the mechanisms of long-distance control will operate. The tree building mechanisms assign structure (228) to (216)-(219).

(228) [_{CP} ... (too) ADJ/ADV (enough) [_{CP} *O_i* [_{PRO_j} to V ... pro-self_j ... *e_i*]]]

The *wh*-PRO mechanism will assign the features of PRO_{one} to PRO yielding (228a).

(228a) [_{CP} ... (too) ADJ/ADV (enough) [_{CP} *O_i* [PRO[+g,-d]_j to V ... pro-self_j ... *e_i*]]]

Via the mechanisms of long-distance control, see "Long-Distance Control" on page 249, the features [+def,α-plural,β-feminine] are transferred onto PRO, yielding (228b).

(228b) [_{CP} ... (too) ADJ/ADV (enough) [_{CP} *O_i* [PRO[+g,+d]_j to V ... pro-self_j ... *e_i* ...]]]

If the reflexive pronoun is *herself*, as in (216)-(219), then PRO must have a singular, feminine NP as a referent in the sentence.

7.7.4 THE INTERNAL STRUCTURE OF THE ADV ADJUNCT INFINITIVE CLAUSE.

It is possible for an ADV adjunct infinitive clause to contain another ADV adjunct, as in (229)-(231).

(229) John is too hungry to be too tired to eat.

(230) John is too hungry to be considered too tired to eat.

(231) John struck Mary as too hungry to be considered by anyone to be too tired to eat.

(232) This airplane is worn out enough to be considered too risky to fly.

The tree building mechanisms assign structure (229') to (229).

(229') [John is [AP too hungry [CP PRO to be [AP too tired [CP PRO to eat]]]]]

Let us consider the higher ADV adjunct, *too hungry*. The Agr-SR relation indicates that *hungry* is linked to *John*. The ADV-mod link indicates that PRO is linked to *hungry*. The ADV-adjunct mechanism of indirect control indicates that PRO is coindexed with *John*, yielding (229a).

(229a) [John_i is [AP too hungry [CP PRO_i to be [AP too tired ...]]]]

Considering the lower ADV adjunct, *too tired*, the Agr-SR relation links PRO_i to the lower *tired*. The ADV-mod link relates PRO to *tired*. This indicates that the lower PRO is coindexed with the higher PRO, yielding (229b) as the output of the control module.

(229b) [John_i ... [CP PRO_i to be [AP too tired [CP PRO_i to eat]]]]

It is possible for the adverbial complement sentence to contain a passive, as in (233)-(239).

- (233) i. John is too sick to operate on.
ii. John is too sick to be operated on.

- (234) i. John is healthy enough to operate on.
ii. John is healthy enough to be operated on.

The following sentences contain a passive in the adverbial infinitive adjunct clause.

(235) Shakespeare's play is too sad (for anyone) to be amused by.

(236) Poverty is too tragic (for anyone) to be amused by.

(237) The acting was too vulgar (for anyone) to be entertained by.

(238) This music is too melodious (for anyone) to be annoyed by/at.

(239) The horror show is too stupid (for anyone) to be frightened by.

If the *by* phrase is removed from sentences (235)-(239), and PRO is coindexed with the matrix subject, the ill-formed sentences in (240)-(244) result.

- (240) * [Shakespeare's play]_{*i*} is too sad PRO_{*i*} to be amused.
 (241) * Poverty_{*i*} is too tragic PRO_{*i*} to be amused.
 (242) * [The acting]_{*i*} was too vulgar PRO_{*i*} to be entertained.
 (243) * [This music]_{*i*} is too melodious PRO_{*i*} to be annoyed.
 (244) * [The horror show]_{*i*} is too stupid PRO_{*i*} to be frightened.

The tree building mechanisms assign structure (235') to (235).

- (235') [Shakespeare's play is [AP too sad [CP O_{*i*} PRO_{*j*} to be amused
e_j by *e_i*]

In the control module, PRO is assigned PRO_{*one*} by the *wh*-PRO mechanism. The Agr-SR link obtains between *Shakespeare's play* and the adjective *sad*. The ADV-mod link obtains between the adjective *sad* and the adverbial CP. Hence, the ADV adjunct indirect referent mechanism indicates the referent of the empty operator is the NP *Shakespeare's play*. The output of the control module is (235a).

- (235a) [[Shakespeare's play]_{*i*} is [AP too sad [CP O_{*i*} PRO[+g,-d]_{*j*} to be amused *e_j* by *e_i*]

The tree building mechanisms assign structure (240') to (240).

- (240') Shakespeare's play is [AP too sad [CP PRO_{*i*} to be amused *e_i*]]

The Agr-SR mechanism links *Shakespeare's play* with the adjective *sad*. The ADV-mod mechanism links the adverbial infinitive complement with *sad*. Thus the ADV adjunct indirect mechanism indicates that PRO is coindexed with *Shakespeare's play*, yielding (240a) as the output of the control module.

- (240a) [[Shakespeare's play]_{*i*} is [AP too sad [CP PRO_{*i*} to be amused
e_i]]]

This sentence is deviant because *amuse* requires an animate object. Sentence (240a) has the same deviance as the sentences in (245).

- (245) i. * Shakespeare's play was amused.
 ii. * Someone amused Shakespeare's play.

Notice that sentence (246), which the tree building mechanisms assign structure (246'), is not deviant.

(246) Shakespeare is too sad to be amused.

(246') [Shakespeare is [AP too sad [CP PRO_i to be amused e_i]]]

The ADV-adjunct mechanism would coindex *Shakespeare* and PRO yielding (246a) as the output of control.

(246a) [Shakespeare_i is [AP too sad [CP PRO_i to be amused e_i]]]

When the animate noun *Shakespeare* is in the subject position, we get the sentences in (247).

- (247) i. Shakespeare is too sad to be amused.
 ii. Shakespeare is too sad to be amused by.

This grammar predicts that sentences (248i)-(250i) are well-formed with the analyses indicated in (248ii)-(250ii).

- (248) i. The old doctor is too incompetent to be operated on by.
 ii. [the old doctor_i is [AP too incompetent [O_i PRO_{one-j} to be operated on e_j by e_i]]]
- (249) i. The old doctor is too sick to be operated on.
 ii. [the old doctor_i is [AP too sick [PRO_i to be operated on e_i]]].
- (250) i. The old doctor is too sick to operate on.
 ii. [the old doctor_i is [AP too sick [O_i PRO_{one} to operate on e_i]]]

Lasnik and Fiengo (1974) offer an analysis of ADJ infinitival adjuncts, which they call OD sentences. They state:

We will now consider sentences such as <9c> and <10c>, which we will call OD sentences.

- [<9c> The mattress is too thin to sleep on. ...
 <10c> The football is soft enough to kick.]

In some respects these sentences are similar to the sentences we have just considered, while in other respects they differ. Consider first the following examples which illustrate some of the differences.

- <19> a. * Socrates_f is dull enough (for me) to be bored by ____f.
 b. * The policemen_f are too stupid (for the demonstrators) to be captured by ____f.
 c. * This music_f is too cacophonous (for Bill) to be put to sleep by ____f.
 d. * Mary_f is too incompetent for this job to be offered to ____f by I.B.M.

It is apparent that Passive cannot apply in the complement of OD sentences.
 (Lasnik and Fiengo 1974:538)

I disagree with their analysis. As I have just shown, passives do occur in these structures, but the sentences offered by L&F are sufficiently complicated as to make interpretation difficult.

7.8 NOUN INFINITIVAL ADJUNCTS: INFINITIVAL RELATIVES

Structure (251) represents the general configuration for noun infinitival adjuncts.

(251) [CP ... [NP [N' [N' N] [CP PRO to VP]]] ...]

When this configuration is input to the control module, there are three possible outputs: exclusive-PRO, PRO_{one}, and long-distance control PRO.

Case I: PRO is assigned the features of exclusive-PRO, and coindexed with the head noun of the NP that contains the adjunct. Sentences (253)-(255) are assigned the structure (252) by the tree building mechanisms.

- (252) [CP ... [NP [N' [N' N] [CP PRO to VP]]] ...]
- (253) Mary knows [NP the candidate [PRO to beat Bill]]
- (254) Mary knows [NP the plumber [PRO to fix the sink]]
- (255) [NP the person [PRO to fix the sink]] showed up late

The control mechanism for Case I is stated in (256).

- (256) The NP Adjunct Mechanism: In an NP infinitive complement, assign the features [-g,+d] and coindex PRO with the head noun of the NP to which the adjunct is attached, i.e. to structure (i) assign the features and indices (ii).
- i. [CP ... [NP [N' [N' N] [CP PRO to VP]]] ...]
- ii. [CP ... [NP [N' [N' N_i] [CP PRO[-g,+d]_i to VP]]] ...]

This mechanism will assign PRO in (253')-(255') the following referents:

- (253') Mary knows [NP the candidate_i [PRO[-g,+d]_i to beat Bill]]
- (254') Mary knows [NP the plumber_i [PRO[-g,+d]_i to fix the sink]]
- (255') [NP the person_i [PRO[-g,+d]_i to fix the sink]] showed up late

Case II: PRO is assigned the features of PRO_{one}. The tree building mechanisms assigns the structure in (257) to the sentences (258)-(260).

- (257) [CP ... [NP [N' [N' N] [CP O_i PRO ... e_i ...]]] ...]
- (258) Mary knows [NP the candidate [CP O_i PRO to beat e_i]]
- (259) Mary knows [NP the plumber [CP O_i PRO to hire e_i]]
- (260) [NP the person [CP O_i PRO to hire e_i]] will show up early

The *wh*-PRO mechanism will assign the features of PRO_{one}, yielding (258')-(260') as the output of the control module:

- (258') Mary knows [NP the candidate [CP O_i PRO[+g,-d] to beat e_i]]
- (259') Mary knows [NP the plumber [CP O_i PRO[+g,-d] to hire e_i]]
- (260') [NP the person [CP O_i PRO[+g,-d] to hire e_i]] will show up early

The empty operator is bound by the strong binding mechanism for adjuncts.

(261) The N Adjunct Strong Binding Mechanism: In an NP infinitive complement, coindex the empty operator, *O*, with the head noun of the NP to which the adjunct is attached, i.e. to structure (i) assign indices (ii).

- i. [CP ... [NP [N' [N' N] [CP *O* PRO to VP]]] ...]
- ii. [CP ... [NP [N' [N' N_{*i*}] [CP *O_i* PRO to VP]]] ...]

The mechanism (261) operates on (258')-(260') to yield the structures (258a)-(260a).

(258a) Mary knows [NP the candidate_{*i*} [CP *O_i* PRO[+g,-d] to beat *e_i*]]

(259a) Mary knows [NP the plumber_{*i*} [CP *O_i* PRO[+g,-d] to hire *e_i*]]

(260a) [NP The person_{*i*} [CP *O_i* PRO[+g,-d] to hire *e_i*]] will show up early

Case III: PRO is coindexed with a definite reflexive pronoun and is given a referent by long-distance control. Sentence (263) has the structure (262).

(262) [CP ... [NP [N' [N' N] [CP *O_i* PRO_{*j*} ... pro-self_{*j*} ... *e_i* ...]]] ...]

(263) Mary found a candidate [*O_i* PRO_{*j*} to get herself_{*j*} excited about *e_i*]

PRO, being coindexed with the definite reflexive pronoun *herself*, inherits the features [+def,+fem,-pl]. Mechanism (261) coindexes *O* with the noun head. *Mary* and PRO are coindexed by the long-distance control mechanism. The final structure of (263) is (263a).

(263a) Mary_{*j*} found a candidate_{*i*} [*O_i* PRO[+g,+d]_{*j*} to get herself_{*j*} excited about *e_i*]

Williams (1980) mentions some examples in which pragmatic factors play a role in determining the antecedent of PRO. He states:

...In some cases, pragmatic or semantic factors choose one controller over another (see Bach (1979)).

<62a> I bought a book [PRO to read]_{*arb*}

<62b> I bought Bill a book [PRO to read]*arb*

In <62a>, *I* is the controller; in <62b>, it is *Bill*. The controller is obviously neither thematically nor grammatically unique. (Williams 1980:217)

These examples are cases of NP adjunct clauses with PRO_{one} and the empty operator. Let us consider some examples containing reflexive pronouns.

(264) i. Mary bought some soap to wash the baby with.

ii. Mary bought [NP some [N' [N' soap] [O_i PRO_{one} to wash the baby with e_j]]]

(265) i. Mary bought some soap to wash her with.

ii. Mary bought [NP some [N' [N' soap] [O_i PRO_{one} to wash her with e_j]]]

(266) i. Mary bought some soap to wash herself with.

ii. Mary bought [NP some [N' [N' soap] [O_i PRO_[+g,+d] to wash herself_j with e_j]]]

(267) i. Mary bought Bill some soap to wash (oneself) with.

ii. Mary bought Bill [NP some [N' [N' soap] [O_i PRO_[+g,-d] to wash oneself_j with e_j]]]

Consider now the sentences in (268)-(271) which have essentially the structure (267ii).

(268) Mary bought Bill some soap PRO to wash himself/him with.

(269) Mary bought Bill some soap PRO to wash *herself/her with.

(270) Mary bought Bill some soap PRO to wash themselves with.

(271) Mary bought Bill some soap PRO to wash each other with.

In each case, the mechanisms of long-distance control transfer the features for [definite], [feminine], and [plural] onto the PRO_[+g,-d]. The principles which determine the possible referent(s) for PRO_[+g,+d,αfem,βplural] are those which determine the possible referent(s) of a pronoun (*he, she, they*) over two sentence boundaries.

The fact that (269) is bad when *herself* is coindexed with PRO is unusual, and follows from the meaning of the sentence. The implication is that *Bill* has *the soap*, but not that *Mary* has *the soap*. To make sense, the referent of PRO must have *the soap*. Consider who has *the soap* in the (272)-(274):

- (272) i. Mary bought for Bill some soap PRO to wash (oneself) with.
 ii. Mary bought from Bill some soap PRO to wash (oneself) with.
- (273) i. Mary bought for Bill some soap PRO to wash himself/*herself with.
 ii. Mary bought from Bill some soap PRO to wash *himself/herself with.
- (274) i. Mary obtained for Bill some soap PRO to wash himself/*herself with.
 ii. Mary obtained from Bill some soap PRO to wash *himself/herself with.

The (i) sentences imply that *Bill* has *the soap*, hence *Bill* is the most natural referent for PRO. The (ii) sentences imply that *Mary* has *the soap*, hence *Mary* is the most natural interpretation for the PRO. Notice in each case that [*PRO to wash themselves with*] would be well-formed with PRO including both *Bill* and *Mary*.

7.9 ADJECTIVAL ADJUNCTS: COPULA CLAUSES

Structure (275) characterizes the general structure for adjective (ADJ) infinitive adjunct clauses:

- (275) [_{CP} NP ... [_{AP} [_{A'} [_{A'} ADJ] [_{CP} PRO to VP]]]]

When this configuration is input to the control module, there is only one possible output: PRO_{one}. Examples of ADJ infinitive adjuncts are:

- (276) Mary is pretty to look at.
 (277) Pepper is hot to taste.

(278) Perfume is sweet to smell

ADJ infinitive adjuncts all have the empty operator assigned by the tree building mechanisms. They all have structure (279).

(279) [CP NP ... [AP [A' [A' ADJ] [CP O_i PRO ... e_i ...]]]]

The *wh*-PRO mechanism assigns the features of PRO_{one} to PRO following a *wh*- element in COMP, yielding (279a).

(279a) [CP NP ... [AP [A' [A' ADJ] [CP O_i PRO_{one} ... e_i ...]]]]

The empty operator is given its referent by the ADJ Adjunct Mechanism, given in (280).

(280) The Adjective Adjunct Mechanism.

The NP that is the referent of the empty operator is the NP pointed to by the Agr-SR link of the adjective head of the AP which contains the adjunct.

Consider sentence (276). The tree building mechanisms would assign structure (276') to (276).

(276') [CP [NP Mary] is [AP [A' [A' [ADJ pretty] [CP O_i PRO to look at e_i]]]]]

The *wh*-PRO mechanism of the control module would add the features of PRO_{one} to PRO yielding (276a).

(276a) [CP [NP Mary] is [AP [A' [A' [ADJ pretty] [CP O_i PRO_{one} to look at e_i]]]]]

The ADJ-Adjunct Mechanism indicates the referent of the empty operator is *Mary* since *Mary* is the NP pointed to by the Agr-SR link for *pretty*.

Sentences (281) and (282) would be assigned the structures (281a) and (282a) by the control module. In (281) the Agr-SR link for *pretty* is with *Mary*, hence, the empty operator has *Mary* as referent. In (282) the Agr-SR link for *pretty* is with *Sue*, hence, the empty operator points to *Sue*.

- (281) Mary struck Sue as pretty to look at.
- (282) Mary considered Sue as pretty to look at.
- (281a) [*CP* [*NP* Mary] struck [*NP* Sue] as [*AP* [*A'* [*A'* [*ADJ* pretty] [*CP* *O_i* PRO_{one} to look at *e_i*]]]]]
- (282a) [*CP* [*NP* Mary] considered [*NP* Sue] [*AP* [*A'* [*A'* [*ADJ* pretty] [*CP* *O_i* PRO_{one} to look at *e_i*]]]]]

While the above analysis provides a description of the data, it leaves many questions unanswered. First and foremost: Why is exclusive PRO not permitted in adjective infinitive constructions? In "The Constituent Structure of 'Mary is pretty to look at'" on page 125 I suggest that the external θ -role assigned by *pretty* prevents it from being a controller. However, this may be more a statement of the problem than it is a solution.

In the present version of this theory, ADJ adjunct infinitives can be described, but not explained. The correct interpretation can be assigned to the well-formed examples, but there is no explanation for why the construction is ill-formed with exclusive-PRO, as in (283).

- (283) * Mary is pretty to look at Sue.

7.10 A CLASSIFICATION OF CONTROL STRUCTURES ACCORDING TO FEATURES

PRO[-*gen*,+*def*] or exclusive-PRO has a single NP antecedent in the matrix CP that is interpreted as coreferential with PRO. The NP-PRO coreference is licensed by the lexical entry of the word which licenses the infinitive complement clause or by the mechanisms which license adjuncts.

- (284) John persuaded Bill [[PRO to go]]
- (285) John left [(in order) [PRO to get a paper]]
- (286) John is too stubborn [[PRO to talk to Mary]]

PRO[-gen,-def] or split-PRO requires split antecedents in the matrix sentence. The antecedent-PRO interpretation is licensed by the lexical item licensing the infinitive complement. Split-PRO appears only to be licensed in the complement of verbs, often verbs with the comitative *with*.

(287) John proposed to Mary [[PRO to go to the movies]]

(288) John chipped in with Mary [[PRO to buy a car]]

PRO[+gen,+def] or inclusive-PRO is interpreted to have a reference that includes, but is not necessarily identical to, a NP in the matrix sentence. Inclusive-PRO is the feature set of PRO in a *wh*- infinitive complement of a lexical item and of PRO in cases of long-distance control.

(289) i. Mary asked Bill [how [PRO to behave herself]]
 ii. Mary asked Bill [how [PRO to behave oneself]]
 iii. * Mary asked Bill [how [PRO to behave himself]]

(290) i. It is unclear to Bill [how [PRO to behave himself]]
 ii. It is unclear to Bill [how [PRO to behave oneself]]
 iii. * It is unclear to Bill [how [PRO to behave herself]]

(291) i. Mary told Bill it would be appropriate [[PRO to pay for herself]]
 ii. Mary told Bill it would be appropriate [[PRO to pay for himself]]
 iii. Mary told Bill it would be appropriate [[PRO to pay for oneself]]
 iv. Mary told Bill it would be appropriate [[PRO to pay for themselves]]

(292) i. Mary thinks it would be appropriate [[PRO to pay for herself]]
 ii. Mary thinks it would be appropriate [[PRO to pay for oneself]]
 iii. * Mary thinks it would be appropriate [[PRO to pay for himself]]
 iv. Mary thinks it would be appropriate [[PRO to pay for themselves]]

$PRO_{[+gen,-def]}$ is PRO_{one} . PRO_{one} can be interpreted as first (speaker) or second (hearer) person. PRO_{one} has no NP reference in the sentence. It is licensed as the PRO in the infinitive complement of some verbs and as PRO in an infinitive with a *wh*- element in the COMP node. PRO_{one} cannot have an antecedent NP in the sentence.

- (293) John said [[PRO to leave]]
- (294) [Where [PRO to sleep in NY]] is a problem.
- (295) John is too stubborn [\emptyset [PRO to talk to e]]
- (296) John is easy [\emptyset [PRO to talk to e]]
- (297) Mary is pretty [\emptyset [PRO to look at e]]
- (298) It would be appropriate [[PRO to leave early]]
- (299) It is impossible [[PRO to drink rubbing alcohol]]

8.0 LEXICAL STRUCTURES OF VERBS AND ADJECTIVES

In this chapter I will review the possible lexical structures of verbs and adjectives. The fragment of lexical structure with which we are concerned is that part which specifies the major category of the word, whether or not it assigns an external θ -role, whether or not it assigns an internal θ -role (an NP or PP), and whether it assigns a clausal complement. The only clausal complement which concerns us is an infinitival complement. Finally, the lexical structure indicates the extrinsic feature assignment to the subject of the infinitive, and which of the θ -roles it assigns is the controller of PRO. A model lexical entry is in (1).

- (1) *word*
- [$\pm V, \pm N$]
 - \pm external θ -role
 - internal argument(s)
 - Control: internal or external argument control,
 - Extrinsic feature assignment to PRO

Control is closely associated with θ -role assignment in that an antecedent of PRO is always in a θ -position. When a verb assigns more than one θ -role, an external θ -role (which is always an NP) and an internal θ -role (which is either an NP or a PP), the lexicon must specify which θ -role carrier is the controller. When only one θ -role is assigned, and the [$+\text{def}$] extrinsic feature is assigned to PRO, it is unnecessary to specify the controller, since it must be from the single θ -position.

8.1 VERBS WITH EXTERNAL THETA ROLE

Transitive verbs, such as *promise* and *persuade* assign both external and internal θ -roles, as well as a clausal complement, as in (2).

- (2) *promise, persuade* ...
 [+V,-N]
 external θ -role
 NP, CP

Their lexical structures differ with respect to assigning the controller. Before the control mechanisms assign extrinsic features to PRO and coindex PRO with an antecedent, *promise* and *persuade* have the same tree structure. However, when control information is included, the lexical structures diverge. *Promise* has the lexical structure (3) and *persuade* the lexical structure in (4).

- (3) *promise*
 [+V,-N]
 external θ -role
 NP, CP
 control: external, [-g,+d]

- (4) *persuade*
 [+V,-N]
 external θ -role
 NP, CP
 control: internal, [-g,+d]

The same holds for verbs which take PP complements. *Swear* and *shout* license similar trees, as in (5), only control differs, as in (6) and (7).

- (5) *swear, shout* ...
 [+V,-N]
 external θ -role
 PP_{to}, CP

- (6) *swear*
 [+V,-N]
 external θ -role
 PP_{to}, CP
 control: external, [-g,+d]

- (7) *shout*
- [+V,-N]
external θ -role
PP_{to}, CP
control: internal, [-g,+d]

The lexical structure of a verb which assigns an external θ -role but no internal θ -role, like *wonder*, is shown in (8).

- (8) *wonder* ...
- [+V,-N]
external θ -role
CP
Control: [+g,+d]

Wonder has only a single θ -role to assign and licenses an infinitive clause. In addition, *wonder* assigns the extrinsic feature [+definite] to PRO, so an antecedent is required. Since there is only one θ -role in the lexical structure, the controller is the NP which carries the θ -role.

8.2 VERBS WITH NO EXTERNAL THETA ROLES

A verb which assigns no external θ -role, licenses a CP, and either does or does not license an internal θ -role, will have a lexical entry like (9).

- (9) *verb*
- [+V,-N]
internal arguments: (PP|NP), CP
Control information: indexing and features of PRO

Postponing questions of control for the moment, I want to focus on the tree structures defined by such a lexical entry. The fragment of lexical structure represented by (9) licenses two phrase markers:

- (10) i. [CP [NP it] V (PP|NP) [CP PRO ...]]
ii. [CP [CP PRO ...] V (PP|NP)]

The verbs *annoy*, *escape*, and *stink* license a CP but no external θ -role and have entries like (9). *Annoy* and *escape* license internal θ -roles, while *stink* does not.

- (11) i. [It annoys John [PRO to eat soup]]
 ii. [[PRO to eat soup] annoys John]
- (12) i. [It escapes me [how PRO to tie my shoes]]
 ii. [[how PRO to tie my shoes] escapes me]
- (13) i. [It stinks [PRO to eat raw meat]]
 ii. [[PRO to eat raw meat] stinks]

Lexical entries specify the type of complements licensed by the verb but give no order to the elements. The order of elements follows from general principles of UG. For instance, (9) cannot result in the order of constituents in (14) because the NP must be adjacent to the verb in order to be assigned (abstract) Case.

- (14) i. * [it V CP NP]
 ii. * [CP PRO to eat soup] John

Entry (9) cannot result in the ill-formed order of constituents in (15) because the NP requires a θ -role, and verbs like (9) do not assign an external θ -role.

- (15) i. * [NP V CP]
 ii. * [CP PRO to eat soup]

There are four possible extrinsic feature combinations of PRO for verbs which assign no external θ -role:

1. Exclusive-PRO or [-gen,+def]:

- (16) *annoy*²⁰
 [+V,-N]
 internal: NP, CP

²⁰ More verbs in the [-gen,+def] class are: *confuse*, *bug*, *please*, *frighten*, *scare*, *amuse*, *goad*, *inspire*, *surprise*, *gall*, *pain*, *grieve*.

control: internal, PRO[-g,+d]

(17) *appeal*

[+V,-N]

internal: PP_{to}, CP

control: internal, PRO[-g,+d]

2. Split-PRO or [-gen,-def]:

This feature complex for PRO is, in principle, impossible in structures where no external θ -role is assigned. The features of Split-PRO require PRO to have split antecedents and occur in a structure like (18).

(18) * [NP *it*]_j V (P) NP_i [CP PRO[-g,-d]_{i,j} to VP]

Such a structure is impossible because the expletive *it* which occupies the subject position of verbs which assign no θ -role has no referential properties. Expletive *it* cannot be the antecedent of anything.

3. Inclusive-PRO or [+gen,+def]:

(19) *escape, elude* ...

[+V,-N]

internal: NP, CP

control: internal, PRO[+g,+d]

4. PRO_{one} or [+gen,-def]:

(20) *stink, suck* ...

[+V,-N]

internal: CP

control: [+g,-d]

Some verbs, like *annoy, appeal, escape* and *elude*, require an NP or PP complement. Some verbs, like *stink*, do not permit an NP or PP complement. Other verbs, like *help* and *pay*, may or may not occur with an NP or PP complement. When *help* has an NP complement, then PRO is

PP_{to}, CP, and does not assign an external θ -role. In both (25i) and (25ii), PRO is [-g,+d] and is coreferential with *Joan*.

- (25) i. [it appeals to Joan_i [PRO_i to own a horse]]
 ii. [[PRO_i to own a horse] appeals to Joan_i]

Since the lexically defined reference of PRO is specified in the lexical entry for *appeal*, PRO in (26)-(29) must be coindexed with *Joan* and cannot be coindexed with *Mary*.

- (26) i. [[that [PRO[-g,+d]_i to own a horse] appeals to Joan_i] amused Mary]
 ii. [Mary was amused [that [PRO[-g,+d]_i to own a horse] appeals to Joan_i]]
- (27) i. [[that it appeals to Joan_i [PRO[-g,+d]_i to own a horse]] amused Mary]
 ii. [Mary was amused [that it appeals to Joan_i [PRO[-g,+d]_i to own a horse]]]
- (28) i. * [[that [PRO[-g,+d]_i to own a horse] appeals to Joan] amused Mary_i]
 ii. * [Mary_i was amused [that [PRO[-g,+d]_i to own a horse] appeals to Joan]]
- (29) i. * [[that it appeals to Joan [PRO[-g,+d]_i to own a horse]] amused Mary_i]
 ii. * [Mary_i was amused [that it appeals to Joan [PRO[-g,+d]_i to own a horse]]]

The lexical entry (19) for *escape* coindexes PRO with the internal argument:

- (30) i. [[how PRO[+g,+d]_i to tie his shoes] escaped Bill_i]
 ii. [it escaped Bill_i [how PRO[+g,+d]_i to tie his shoes]]

The lexically defined reference of PRO is internal to the lexical entry for *escape* and must be *Bill* and not *Tom*.

- (31) i. [[that [how PRO_[+g,+d]_i to tie his shoes] escaped Bill_i] surprised Tom]
- ii. [Tom was surprised [that [how PRO_[+g,+d]_i to tie his shoes] escaped Bill_i]]
- (32) i. [[that it escaped Bill_i [how PRO_[+g,+d]_i to tie his shoes]] surprised Tom]
- ii. [Tom was surprised [that it escaped Bill_i [how PRO_[+g,+d]_i to tie his shoes]]]
- (33) i. * [[that [how PRO_[+g,+d]_i to tie his shoes] escaped Bill] surprised Tom_i]
- ii. * [Tom_i was surprised [that [how PRO_[+g,+d]_i to tie his shoes] escaped Bill]]
- (34) i. * [[that it escaped Bill [how PRO_[+g,+d]_i to tie his shoes]] surprised Tom_i]
- ii. * [Tom_i was surprised [that it escaped Bill [how PRO_[+g,+d]_i to tie his shoes]]]

Verbs with the lexical marking (16), *annoy*, *please*, *frighten*, *scare*, etc., can occur in three different structures, two active as in (35i)-(35ii) and one passive as in (35iii).

- (35) i. [it annoys John_i [PRO_[-g,+d]_i to pay for himself]]
- ii. [[PRO_[-g,+d]_i to pay for himself] annoys John_i]
- iii. [John_i is annoyed e_i [PRO_[-g,+d]_i to pay for himself]]

Since, according to the lexical specifications, PRO is coindexed with the NP internal argument, PRO can only be coreferential with *John* and not *Bill* in these sentences.

- (36) i. [Bill thinks [it annoys John_i [PRO_[-g,+d]_i to pay for himself]]]
- ii. [Bill thinks [[PRO_[-g,+d]_i to pay for himself] annoys John_i]]
- iii. [Bill thinks [John_i is annoyed e_i [PRO_[-g,+d]_i to pay for himself]]]

- (37) i. [[that it annoys John_i [PRO_[-g,+d]_i to pay for himself]] surprised Bill]
- ii. [[that [PRO_[-g,+d]_i to pay for himself] annoys John_i] surprised Bill]
- iii. [[that John_i is annoyed e_i [PRO_[-g,+d]_i to pay for himself]] surprised Bill]

8.2.1 LITERATURE ON VERBS THAT ASSIGN NO EXTERNAL THETA ROLE

Manzini (1983) and Chomsky (1986a) have taken different positions concerning the interpretation of PRO with verbs like *help* and *annoy*.

Manzini offers the following views:

Consider on the other hand a PRO in a subject sentence of a sentence S. The PRO can have arbitrary reference, as in <26>; or it can corefer into S, as in <27>, into a phrase superordinate to S, as in <28>, or into a phrase subordinate to S, as in <29>:

- <26> [PRO to behave oneself in public] would help Bill.
- <27> [PRO to behave himself in public] would help Bill.
- <28> Mary knows that [PRO to behave herself in public] would help Bill.
- <29> [PRO to behave himself in public] would help Bill's development.

Thus, for a PRO in a subject sentence, the generalization <30> appears to hold:

- <30> A PRO in a subject sentence (co)occurs freely.

(Manzini 1983:424)

Chomsky (1986a) disagrees with the data offered by Manzini.

Chomsky states:

If the construction C with PRO as subject is itself a subject rather than a complement or adjunct, other complications arise, some already illustrated, and some illustrated in <153>, where β is some reflexive form:

- <153> (i) [PRO to have to feed β] would be a nuisance for John.
 (ii) [PRO to have to feed β] would annoy John.
 (iii) [PRO to have to feed β] would annoy John's friends.

- (iv) [PRO to have to feed β] would assist John's development.
- (v) [PRO to have to feed β] would cause John to be annoyed.

Choice of $\beta = \textit{oneself}$ is excluded throughout, meaning that PRO cannot have arbitrary interpretation and must be bound. Again, it is not the configuration but the presence of a potential binder in <153> that bars arbitrary interpretation, as <154> illustrates:

<154> [PRO to have to feed β] is a nuisance.

(Chomsky 1986a:128-129)

I would alter Chomsky's terminology. It is not the presence of a "potential binder" but the presence of a lexically specified internal argument that is coindexed with PRO that is the binder. My analysis yields the following distribution of data.

- (38) * [PRO to behave oneself in public] would help Bill.
- (39) [PRO_{*f*} to behave himself in public] would help Bill_{*f*}.
- (40) * [PRO to behave herself in public] would help Bill.
- (41) * [PRO to behave himself in public] would help Bill's development.
- (42) i. [PRO_{*f*} to have to feed himself] would be a nuisance for John_{*f*}.
 - ii. [PRO_{*f*} to have to feed himself] would annoy John_{*f*}.
 - iii. [PRO_{*f*} to have to feed themselves/*himself] would annoy John's friends_{*f*}.
 - iv. * [PRO to have to feed himself] would assist John's development.
 - v. [PRO_{*f*} to have to feed himself] would cause John_{*f*} to be annoyed.

verbs or adjectives can be *tough*-predicates. There are four possible cases in which the lexical item licenses a CP infinitive complement and assigns no θ -role.

- (46) ADJ, CP(+wh): It is unclear [how PRO to meet John]
[How PRO to meet John] is unclear
- (47) ADJ, CP(-wh): It is easy [PRO to meet John]
[PRO to meet John] is easy.
- (48) V, CP(+wh): * It VERBs [*wh*- PRO to meet John]
* [*wh*- PRO to meet John] VERBs.
- (49) V, CP(-wh): It stinks [PRO to eat meat raw]
[PRO to eat meat raw] stinks.

The empty operator cannot occur in cases (46) and (48) because the COMP node contains a *wh*- element. In fact, case (48) appears not to exist in English. There seems to be no verb in English which assigns no θ -role and requires a *wh*- element in the CP infinitive complement.

There are very few verbs in class (49), *stink* and *suck*, being the most common. All instances of predicates described by (49) seem to be pejorative. There are a large number of adjectives in class (47), and these have traditionally been called *tough* movement constructions. The empty operator can occur in the complements of adjectives like (47), as in example (50ii), but not in verbs like (49), as in example (51ii).

- (50) i. It is easy [PRO to eat meat raw]
ii. Meat is easy [O_i PRO to eat e_i raw]
- (51) i. It stinks [PRO to eat meat raw]
ii. * Meat stinks [O_i PRO to eat e_i raw]

The inability of verbs with no external θ -role to take *tough*-predicate analyses may have to do with the fact that they also have intransitive readings which assign an external θ -role.²¹

8.3.1 THE MOTIVATION FOR AN EMPTY OPERATOR.

The empty operator in both *tough* predicates, (52), and adjunct constructions, (53),²² can be separated from the NP position which it binds by several CP boundaries.

- (52) i. John is easy [O_i PRO to convince Bill to do business with e_i]
 ii. John is easy [O_i PRO to convince Bill to arrange for Mary to meet e_i]
- (53) i. John is too boring [O_i PRO to watch e_i]
 ii. John is too boring [O_i PRO to want to watch e_i]
 iii. John is too boring [O_i PRO to expect anyone to want to watch e_i]

The empty operator analysis makes specific claims about the interaction of *tough* predicates, *wh*- questions, and indirect questions since they all utilize the COMP node. Let us examine the possibilities for having the internal arguments of *take* bound to operators in a string like (54).

- (54) It is easy [PRO to take candy from a baby]

The sentences in (55)-(60) show that either of the internal arguments of *take* can be bound by a variable operator: *who*, *what*, or the empty operator.

21 There are some cases when so-called *tough*- movement constructions with verbs do not sound so bad, as in *This neighborhood stinks/sucks to park in*.

22 Examples (52i) and (52ii) are Chomsky's examples in <127b> in Chomsky 1977b, p.103, where Chomsky discusses similar phenomena.

- (55) Candy_i is easy [*O*_i PRO to take *e*_i from a baby]
- (56) A baby_i is easy [*O*_i PRO to take candy from *e*_i]
- (57) What_i is it easy [*e*_i PRO to take *e*_i from a baby]
- (58) Who_i is it easy [*e*_i PRO to take candy from *e*_i]
- (59) What_i is *e*_i easy [*O*_i PRO to take *e*_i from a baby]
- (60) Who_i is *e*_i easy [*O*_i PRO to take candy from *e*_i]

Sentences (61)-(64) are ill-formed because the specifier of the COMP node can contain either a trace bound to the *what/who* or an empty operator strongly bound to *baby/candy*, but not both. Sentences (61)-(62) are ill-formed for the same reason that (63)-(64) are ill-formed.

- (61) * What_i is a baby_j easy [*CP e*_i *O*_j [*IP* PRO to take *e*_i from *e*_j]]
- (62) * Who_j is candy_i easy [*CP e*_j *O*_i [*IP* PRO to take *e*_i from *e*_j]]
- (63) * What_i did John wonder [*CP who*_j *e*_i [*IP* PRO to take *e*_i from *e*_j]]
- (64) * Who_j did John wonder [*CP what*_i *e*_j [*IP* PRO to take *e*_i from *e*_j]]

If we assume that the specifier of a COMP can contain either an operator (65), or a trace bound to an operator (66), but not both (67), then the distribution of grammatical forms in *tough* predicate constructions follows from general principles of *wh*-operator binding.

- (65) [*NP*_i easy [*CP O*_i [*IP* PRO to V ... *e*_i]]]
- (66) [*wh*_i *NP* easy [*CP e*_i [*IP* PRO to V ... *e*_i]]]
- (67) * [*wh*_i *NP*_j easy [*CP O*_j *e*_i [*IP* PRO to V ... *e*_j *e*_i]]]

Let us consider sentences of three degrees of embedding, like (68)-(69).

- (68) i. It is easy to remember to send Bill money.
 ii. it is easy [PRO to remember [PRO to send Bill money]]

(69) i. Bill is easy to remember to send e money.

ii. $Bill_i$ is easy [0_i PRO to remember [e_i PRO to send e_i money]]

The sentences in (70)-(73) are ill-formed because the COMP would have to have both the empty operator and an overt *wh*- element.

(70) i. * Bill is easy to remember when to send e money.
ii. * $Bill_i$ is easy [0_i PRO to remember [CP when e_i [PRO to send e_i money]]]

(71) i. * Bill is easy to remember what to send e e.
ii. * $Bill_i$ is easy [0_i PRO to remember [CP what $_j$ e_i [PRO to send e_i e_j]]]

(72) i. * Bill is easy to remember what to send e to e
ii. * $Bill_i$ is easy [0_i PRO to remember [CP what $_j$ e_i [PRO to send e_j to e_i]]]

(73) i. * Money is easy to remember who to send e e
ii. * $Money_i$ is easy [0_i PRO to remember [CP who $_j$ e_i [PRO to send e_j e_i]]]

Sentences (74)-(75) indicate that a *wh*- operator can bind a variable internal to a CP but not internal to a NP since a CP contains a COMP node but a NP does not.

(74) i. Who do you believe John claims he saw?
ii. Who $_i$ do you believe [CP e_i [IP John claims [e_i he saw e_i]]]

(75) i. * Who do you believe John's claim he saw?
ii. * Who do you believe [NP [SPEC John's] [N' claim] [CP e_i he saw e_i]]]

In the following sentences, there is a COMP node on *plan* as a verb, but not on *plan* as a noun. Both (76), where *plan* is a verb, and (77), where *plan* is a noun, are well-formed with an infinitive complement. In (78)-(80), the matrix subject can bind an empty category in the object of the complement. However, when *plan* is a noun, as in (81)-(82), the matrix subject cannot bind an empty category in the object of the complement.

- (76) It is fun to remember to plan to cook a dinner.
 (77) It is fun to remember our plans to cook a dinner.
 (78) A dinner_{*i*} is fun to cook *e_i*.
 (79) A dinner_{*i*} is fun to plan to cook *e_i*.
 (80) A dinner_{*i*} is fun to remember to plan to cook *e_i*.
 (81) * A dinner is fun to remember the plan to cook *e*
 (82) * A dinner is fun to remember our plans to cook *e*

Arguments similar to those above hold for the CP infinitive adjunct complements. The adjectives *stingy* and *valuable* assign external θ -roles and can take an ADV degree modifier.

- (83) John is stingy/very stingy/extremely stingy/too stingy
 (84) This ring is valuable/very valuable/not so valuable/valuable enough

Consider the differences between these sentences:

- (85) i. John is too stingy to receive a gift from.
 ii. John is [_{AP} too [_{A'} [_{A'} stingy] [_{CP} *O_i* PRO to receive a gift from *e_i*]]]
 (86) i. John is too stingy to open a gift from.
 ii. * John is [_{AP} too [_{A'} [_{A'} stingy] [_{CP} PRO to open [_{NP} a [_{N'} gift from *e*]]]]]

In (85) and (87), *from NP* is in the verb complement of *receive*. In (86) and (88), *from NP* is in the noun complement of *gift*.

- (87) Who_{*i*} did you receive a gift from *e_i*
 (88) * Who did you open a gift from *e*
 (89) You received a gift from Bill.
 (90) You opened a gift from Bill.

The above distribution of data follows from two assumptions: (a) CP has a COMP node but NP does not, and (b) empty operator constructions

require a COMP node. A third assumption, (c) only one *wh*- operator or trace may be in a COMP node at one time, leads to the following distribution of forms.

- (91) This ring_{*i*} is too valuable [*O_i* PRO to forget [*O_i* PRO to put *e_i* in the safe at night]]
- (92) * This ring is too valuable [PRO to forget [when PRO to put *e* in the safe]]
- (93) This area_{*i*} is sensitive enough [*O_i* PRO to remember [*e_i* PRO to wash *e_i* everyday]]
- (94) * This area is sensitive enough [PRO to remember [when PRO to wash *e*]]
- (95) * What did you forget when to put *e e* in the safe?
- (96) * Where did you forget what to put *e e* ?

The complement of *easy* can be active or passive in an empty operator construction.

- (97) i. It is easy [PRO to introduce yourself to John]
 ii. It is easy [PRO_{*i*} to be introduced *e_i* to John]
 iii. John_{*j*} is easy [*O_j* PRO_{*i*} to be introduced *e_i* to *e_j*]
- (98) i. Her parties_{*i*} are fun [*O_i* PRO_{*j*} to be invited *e_j* to *e_i*]
 ii. Albania_{*i*} is difficult [*O_i* PRO_{*j*} to be allowed *e_j* [*e_i* PRO to visit *e_i*]]

Similarly, the ADV adjunct CP can be active or passive in empty operator constructions. The ADV adjunct is discussed in "Adverbial Adjuncts: Degree Adverbial Phrases" on page 296.

- (99) i. Her parties_{*i*} are too boring [*O_i* PRO_{*j*} to be invited *e_j* to *e_i*]
 ii. Her parties_{*i*} are too boring [*O_i* PRO to want [*e_i* PRO_{*j*} to be invited *e_j* to *e_i*]]
 iii. Her parties_{*i*} are too boring [*O_i* PRO_{*j*} to expect anyone [*e_j* PRO to want [*e_i* PRO_{*j*} to be invited *e_j* to *e_i*]]]

In considering verbs and adjectives which (a) assign no θ -roles and (b) license a CP infinitive complement, the distribution of empty

operators follows from the assumptions in (43). No adjective or verb is lexically marked to require an empty operator construction. For every empty operator construction, there is always a corresponding non-empty operator construction, e.g. (100) contain empty operators, (101) do not.

- (100) i. Bill is easy [O_i PRO to see e_i]
 ii. Mary is fun [O_i PRO to be with e_i]

- (101) i. It is easy [PRO to see Bill]
 ii. It is fun [PRO to see Mary]

Adverbial complements should be similar to constructions like (100)-(101), i.e. there should be an ADV adjunct infinitive clause which has PRO_{one} and no empty operator. In "Adverbial Adjuncts: Degree Adverbial Phrases" on page 296, we discussed these sentences:

- (102) Shakespeare's play_{*i*} is too sad [PRO_{*i*} to amuse anyone]

- (103) Shakespeare's play_{*i*} is too sad [O_i PRO_{*j*} to be amused e_j at e_i]

The sentences in (105), which contain neither a PRO or an empty operator, and sentences in (107) which contain PRO_{one}, are predicted by the assumptions in (43) to be well-formed. They are the ADV adjunct equivalent of (101). Sentences in (105) have the structure (104).

- (104) [NP is [AP too ADJ [CP for anyone_{*i*} to be amused e_i]]]

- (105) i. Shakespeare's play is too sad for anyone to be amused.
 ii. Poverty is too tragic for anyone to be amused.
 iii. The acting was too vulgar for anyone to be entertained.
 iv. This music is too melodious for anyone to be annoyed.
 v. The horror show is too beautiful for anyone to be frightened.

Sentences in (107) have the structure (106).

- (106) [NP is [AP too ADJ [CP for one_{*i*} to be amused e_i]]]

- (107) i. Shakespeare's play is too sad for one to be amused.
 ii. Poverty is too tragic for one to be amused.
 iii. The acting was too vulgar for one to be entertained.

- iv. This music is too melodious for one to be annoyed.
- v. The horror show is too beautiful for one to be frightened.

The sentences in (109) have the structure (108).

- (108) [NP is [AP too ADJ [CP PRO_{one-i} to be amused e_i]]]
- (109) i. Shakespeare's play is too sad PRO_{one} to be amused.
 ii. Poverty is too tragic PRO_{one} to be amused.
 iii. The acting was too vulgar PRO_{one} to be entertained.
 iv. This music is too melodious PRO_{one} to be annoyed.
 v. The horror show is too beautiful PRO_{one} to be frightened.

8.4 LEXICAL STRUCTURE OF ADJECTIVES

Adjectives in English generally do not have more than one θ -role to assign. Some adjectives assign an external θ -role, others do not. Since adjectives do not assign Case, when an adjective does assign an internal θ -role, the θ -role carrier is always in a prepositional phrase.

Adjectives which assign an external θ -role are generally intransitive. Adjectives with external θ -roles assign the features of Exclusive-PRO (PRO[-gen,+def]) to an infinitival complement. *Afraid* assigns an external θ -role and a clausal complement. Since there is only one θ -role, and PRO has the [+def] feature, the NP bearing the external θ -role is the controller of PRO.

- (110) *afraid* ...
 [+V,+N]
 external θ -role
 CP
 Control: [-gen,+def]

Ex. *John_i is afraid PRO_i to see the movie.*

I have discussed other adjectives which assign external θ -roles, but which do not categorize for clausal complements: *stubborn*, *pretty*, and *like* exemplify the range.

Stubborn assigns an external θ -role but not a clausal complement.

- (111) *stubborn* ...
 [+V,+N]
 external θ -role

Ex. *John is stubborn.*

Pretty, like *stubborn* assigns an external θ -role and no clausal complement. However, the external θ -role of *pretty* licenses the passive-patient adjunct. I will call this a passive-patient (or P) external θ role.

- (112) *pretty* ...
 [+V,+N]
 P-external θ -role

Ex. *Mary is pretty (to look at).*

I have proposed that *like* assigns an external θ -role as well as an object. However, it too does not assign a clausal complement. Rather it, like *pretty*, assigns the P-external θ -role.

- (113) *like*
 [+V,+N]
 P-external θ -role
 NP

Ex. *This coffee is like tar (to drink).*

8.4.1.1 Control in "Pretty" Adjuncts

I propose that the tree building mechanism would build trees for (114)-(116), and that the sentences (114) and (116) are subsequently marked as ill-formed by the control module.

(114) * *Mary is pretty to smile.*

(115) *Mary is pretty to look at.*

(116) * *Mary is pretty to look at John.*

Pretty does not assign a clausal complement. However, its external θ -role licenses a clausal adjunct. Thus the strings are assigned the following phrase markers by the tree building and binding mechanisms.

(114') Mary is [*AP* [*A'* [*A'* pretty] [*PRO* to smile]]]

(115') *Mary_i* is [*AP* [*A'* [*A'* pretty] [*O_i* *PRO* to look at *e_i*]]]

(116') Mary is [*AP* [*A'* [*A'* pretty] [*PRO* to look at John]]]

The *wh-PRO_{one}* licensing mechanism assigns the features [-gen,-def] to *PRO* in (115'), as in (115a), yielding the interpretation in (115b).

(115a) *Mary_i* is [*AP* [*A'* [*A'* pretty] [*O_i* *PRO*[+g,-d] to look at *e_i*]]]

(115b) *Mary_i* is pretty [*O_i* [*PRO_{one}* to look at *e_i*]]

Within the present theory, the structures (114') and (116') are ruled out because the external θ -role is lexically marked to prevent its being a controller.

Adjectives which have no external θ -role to assign may or may not assign an internal θ -role. Adjectives which assign neither external nor internal θ -roles can assign a clausal complement. Since there is no θ -role carrying NP to be the controller, control of *PRO* cannot be lexically specified. Adjectives with no possible controllers therefore assign the features of *PRO_{one}* (*PRO*[-def,+gen]) to the subject of the infinitive clause.

Adjectives which assign no θ -roles and license a clausal complement divide into two classes. The first is the class of *tough* movement predicates.

- (117) *easy, tough* ...
 [+V,+N]
 CP
 control: [+g,-d]

Ex. *John is easy to talk to.*

As we have seen, *tough* predicates license empty operator constructions.

In the second class, a non-expletive NP cannot appear in the subject position. *Unclear* is an example of this class.

- (118) *It is unclear when to visit John.*

- (119) **John is unclear when to visit.*

The empty operator analysis is blocked when the clausal complement is an indirect question because a lexical *wh-* is in COMP.

- (120) *unclear*
 [+V,+N]
 CP[+wh]
 control: [-g,+d]

Ex. *It is unclear what to do*

A non-external- θ -role assigning adjective may assign an internal θ -role. Since adjectives are not case assigners, the internal θ -role is assigned to a PP. Since the adjective carries a θ -role, it can lexically assign the control features to PRO. Typically, the features assigned are those of Exclusive-PRO (PRO[-gen,+def]).

- (121) *noble*
 [+V,+N]
 PP_{of}, CP
 Control: [-gen,+def]

Ex. *It is noble of John_i PRO_i to sacrifice his inheritance.*

8.4.1.2 Control in "Tough" Predicates

Consider the strings below:

- (122) *It is easy to talk to John.*

(123) * John is easy to talk.

(124) John is easy to talk to.

(125) * John is easy to talk to Mary.

The tree building mechanism will build phrase markers for (122) and (124). Phrase markers for (123) and (125) will not be built because *John* cannot be associated with a θ -role in either case.

The tree building mechanism builds the structure (122') for (122):

(122') [it is easy [[PRO to talk to John]]]

Based on the lexical structure of *easy*, the features of PRO_{one} ([+gen,-def]) are assigned to PRO in (122a), yielding the interpretation shown in (122b).

(122a) [it is easy [[PRO[+g,-d] to talk to John]]]

(122b) It is easy PRO_{one} to talk to John.

The tree building mechanism builds the structure (124') for (124):

(124') [$John_i$ is easy [O_i [PRO to talk to e_j]]]

In this lexical structure, both the *wh*- PRO_{one} licensing mechanism and the lexical structure for *easy* can assign the features [+gen,-def] to PRO, yielding (124a) for the interpretation (124b).

(124a) [$John_i$ is easy [O_i [PRO[+gen,-def] to talk to e_j]]]

(124b) $John_i$ is easy [O_i [PRO_{one} to talk to e_j]]]

9.0 THE DOMAIN OF LICENSING PRO

Chomsky (1980a) presents a grammar which generates (1) but blocks (2).

- (1) Mary promised John to go.
- (2) * John was promised to go.

Chomsky indicates one counterexample to his grammar in a footnote:

The point is now clear under the analysis of these forms presented above. There are some curious exceptions to this principle. Hust and Brame (1976) and Solan (1977) cite such double-passives as *John was promised to be allowed to leave*, which should be blocked but is not for reasons that are unclear. It seems that some syntactic property is involved, since substitution of a near synonym for the embedded passive changes the grammatical status to the expected *: e.g. **John was promised to get permission to leave*. (Chomsky 1980a:35 footnote 40)

The basic question investigated in this chapter is this: Why are sentences (2) and (3) ill-formed, but sentence (4) well-formed?

- (2) * John was promised to go.
- (3) * John was promised by Mary to go.
- (4) John was promised to be allowed to go.

I will consider the three following solutions:

The C-command Theory of control assumes that the correct distribution of forms follows from considerations of c-command and government. Manzini, Williams, and Koster offer versions of the c-command theory.

The Thematic Role Theory of control assumes that the relevant differences between the grammatical and ungrammatical forms is a consequence of the θ -roles assigned to the NP in the matrix clause and to PRO. Ruzicka (1983) offers a version of the thematic role theory.

The V_1 - V_2 Theory of control, presented in "The Domain of Licensing of PRO" on page 347, assumes that the correct distribution of grammatical and ungrammatical forms follows from lexical mechanisms, in particular, from the control specifications of lexical items and from the V_1 - V_2 mechanism which uses the information in lexical entries to assign reference and interpretation to PRO.

9.1 THE C-COMMAND THEORY OF CONTROL

Williams (1980) emphasizes that the passives of *strike* and *promise* are ill-formed for the same reason: The thematically specified antecedent does not c-command the complex predicate. Williams states:

The antecedent of OC must precede and c-command the controlled structure, since this is a property of predication. ...

The c-command restriction on predication is especially interesting. Just as we do not have a passive for *strike*, we do not have a passive for *promise*:

- <35> a. * John was struck by Bill as pompous.
 <35> b. * John was promised by Bill_i [PRO to leave]_i

The reason is the same in each case. *Promise*, being a case of OC, must be indexed in PS, and its thematically specified antecedent, *Bill*, does not c-command the complex predicate S.

The lack of a passive in each case is connected with predication--where predication is not involved, we find passives which are grammatical.

- <36> a. John was struck by Bill's pomposity.
 <36> b. John was promised.
 <36> c. John was promised the book.

(Williams 1980: 211)

Koster (1984) offers an analysis in which c-command and government play a role in blocking the passive of *promise*, *try*, and *strike*. Koster states:

With this minimal, rather conventional theory of control in mind, we can return to our original problem, the ungrammaticality of <43>:

<43> * *John* was tried [*e* to go]

The explanation appears to be quite simple. ... Since the infinitival complement of *try* in <43> is a reduced clause, its subject *e* (PRO) is governed and must therefore be bound in its minimal governing category. This means that there must be an obligatory c-commanding antecedent. *John* in <43> is the only NP that fulfills these conditions. But *John* is not the controller according to the independent theory of control, so the sentence is ruled out. In other words, <43> is ungrammatical because the combined requirements of the theory of control and the binding theory cannot be met.

The explanation of the ungrammaticality of <43> is analogous to the explanation of the ungrammaticality of <44>:

<44> * *Bill* was promised [*e* to go]

This sentence is ungrammatical because the designated controller, the underlying subject of *promise*, is absent. Again, the controller cannot be left implicit since *promise* is also a verb that selects reduced clauses (it does not select a *for*-complementizer). As before, this leads to government of the PRO-subject of the complement. Governed PRO is an anaphor subject to principle (A) of the binding theory. In other words, *e* must be bound by *Bill* in <44> (the only possible binder according to the binding theory). Again, this NP is not the underlying subject, which is the designated controller of *promise*. As in the case of <43>, the requirements of the binding theory and the theory of control are not compatible, hence the ungrammaticality. (Koster 1984:431)

The following distribution of data poses several obstacles for the c-command theory of control.

- (5) Columbus promised Queen Isabella [PRO to return with gold]
- (6) * Queen Isabella was promised by Columbus [PRO to return with gold]
- (7) Columbus' promise to Queen Isabella [PRO to return with gold]
- (8) The promise of/by Columbus to Queen Isabella [PRO to return with gold]
- (9) * Queen Isabella's promise of/by Columbus [PRO to return with gold]

Columbus does not c-command PRO in either (6) or (8), but in (8) *Columbus* is the referent of PRO. Further, although *Columbus* is the referent of PRO in (8), *Columbus* cannot be the referent of PRO in (9). However, the structural relation of *Columbus* to PRO is the same in (8) and (9).

An additional problem for the c-command theory of control, which I have already discussed, is that there are numerous cases in which PRO is coindexed with NPs which do not c-command PRO, see "C-command" on page 199.

9.2 THEMATIC ROLE THEORIES

The Thematic Role theory of control assumes that in a structure like (10), the interpretation and reference of PRO can be determined by considering the θ -roles of the NPs in the matrix sentence and the θ -role of the PRO. For example, let θ_1 , θ_2 , and θ_3 be the θ -roles of the matrix subject, matrix object, and PRO respectively. According to the thematic role theory, the interpretation and reference of PRO is a function of θ_1 , θ_2 , θ_3 and the verb class of V_m .

(10) $[CP \underset{\theta_1}{NP} V_m (P) \underset{\theta_2}{NP} [CP \underset{\theta_3}{PRO} \text{ to } V_e]]$

(11) control of PRO = function (θ_1 , θ_2 , θ_3 , verb class of V_m)

A version of the thematic role hypothesis is offered by Ruzicka (1983). Ruzicka assumes that there are two types of verbs possible in the matrix clause. If V_m is a *thematic identity verb*, it requires PRO to be coreferential with the NP in the matrix sentence which has the same θ -role as PRO. If V_m is a *thematically distinct verb* it requires PRO to be coreferential with the NP in the matrix sentence which does

not have the same θ -role as PRO. The details of his theory are not crucial. The only fact he treats in any detail is the one of immediate concern to us: the passive of *promise*. Ruzicka claims that *promise* is a thematic identity verb.

Ruzicka offers the following introduction to the basic ideas of the thematic role hypothesis. He indicates that [+SC], i.e. subject control, verbs in Chomsky (1980) are those which obey the thematic identity condition. Ruzicka states:

I will consider now a reinterpretation of the crucial feature [+SC] assigned to the verb *promise* in the lexicon in terms of a special kind of semantic relations; more precisely, in terms of *thematic relations*. I use the latter term in the sense of Jackendoff (1972) and, essentially, Fillmore's (1968) case roles ("deep cases"). In terms of thematic relations, the feature assignment [+SC] specifies that the person who gives the *promise* (more exactly, the NP expressing the "Agent" of *promise*, which is an argument in its semantic valence structure) must be identical with (alternatively, must control) the NP expressing the person responsible for the execution of the *promise* given, the Agent within the thematic relations of the infinitival construction.

[<96c> * John was promised t [PRO to VP]]

... I shall assume that a constraint is imposed on the control properties of verbs that are provided with the lexical feature [+SC]. The constraint posited restricts the control possibilities of those verbs to *thematic compatibility* of the two noun phrases involved. In particular for verbs of the [+SC] type like *promise* I will propose a condition which I tentatively call the *Thematic Identity Condition* (TIC). Clearly the TIC is violated in (71c), which therefore must be assigned *.
(Ruzicka 1983:311-312)

Ruzicka indicates that the subject of *promise* and the subject of *allow* have the same θ -role, and that the object of *promise* and the object of *allow* have the same θ -role. Ruzicka claims:

The TIC will help us in clarifying sentence <2>, which appears "curious" in Chomsky's treatment. To begin with, the TIC is naturally applicable to all thematic relations in principle.

Let us assume that *you* in <3> can be assigned the same thematic role as *you* in <4>.

- <2> John was promised to be allowed to leave.
 <3> I promise you to leave.
 <4> I allow you to leave.

Of course, in these examples *you* is the controller only in <4>, subject to later discussion ...

It is irrelevant whether or not we can give the same label (e.g. Receiver) to the thematic relations held by *you* in <3> and <4>. Their identity or, at least, their--intuitively--near similarity can easily be contrasted with the sharp distinction separating the thematic relations represented by *John* and by PRO in <96c>. (Ruzicka 1983:312)

Ruzicka indicates that control applies freely and the antecedent of PRO is determined by conditions of thematic identity. PRO obtains its θ -role from the object position of *allow*, and this θ -role agrees with the θ -role of the object position of *promise*. Ruzicka states:

Now let us look at <2> again, supplemented with the traces of NP Movement:

<2'> John_{*i*} was promised t_{*i*} [PRO_{*j*} to be allowed t_{*j*} to leave]

In <2> subject control can now operate correctly, because the Thematic Identity Condition is not violated. Nothing curious or unexpected remains. Control is freely applicable exactly as in <3>, for example. (Ruzicka 1983:312)

There are several arguments against any theory which that control is based on identity of thematic relations between the controller and PRO. In the analysis offered by Ruzicka, PRO_{*j*} in (12) (=Ruzicka's <2>) obtains its θ -role from the empty category that is the object of *allow*.

(12) John_{*i*} was promised e_{*i*} [PRO_{*j*} to be allowed e_{*j*} [PRO to leave]]]

Notice, however, that all of the following sentences, in which PRO_{*j*} likewise obtains its θ -role from the empty category object of *allow*, are ill-formed.

- (13) * John_i was promised e_i [PRO_j to be easy [θ_j [PRO to allow e_j [PRO to leave]]]]
- (14) * John_i was promised e_i [PRO_j to seem [e_j to be allowed e_j [PRO to leave]]]
- (15) * John_i was promised e_i [PRO_j to be certain [e_j to be allowed e_j [PRO to leave]]]
- (16) * John_i was promised e_i [PRO_j to be expected [e_j to be allowed e_j [PRO to leave]]]

A second argument against any theory which claims that the antecedent of PRO is the NP in the matrix sentence which bears the same θ-role as PRO hinges on the following verb pairs. Verb pairs like (*buy/sell* and *rent to/rent from*) share the same θ-roles. In the following examples, let us assume that *Bob*, *house*, and *John* have the θ-roles A, B, and C in sentences (17)-(19), and X, Y, and Z in (20)-(22). Perhaps there is only one set of θ-roles involved, but it is not important to our argument if A, B, and C are the same as X, Y, and Z.

- | | | |
|------|---------------------------------|----------------------|
| (17) | Bob rented a house to John. | A rented B to C |
| (18) | John rented a house from Bob. | C rented B from A |
| (19) | John was rented a house by Bob. | C was rented B by A. |
| (20) | Bob sold a house to John. | X sold Y to Z. |
| (21) | John bought a house from Bob. | Z bought Y from X. |
| (22) | John was sold a house by Bob. | Z was sold Y by X. |

Since sentences (18) and (19) have the same θ-role in the subject position, the θ-role hypothesis predicts they should both have the same properties of control in (23)-(24) when they have PRO as subject. But the sentences differ in PRO reference.

- (23) John_i promised Mary [PRO-C_i to rent a house on the beach]
- (24) John promised Mary_i [PRO-C_i to be rented a house on the beach]

Similarly, since sentences (21) and (22) have the same θ-role in the subject position, the θ-role hypothesis predicts they should both have

the same properties of control in (25)-(26) when they have PRO as subject. But the sentences have different control.

(25) John_f promised Mary [PRO-Z_f to buy a house on the beach]

(26) John promised Mary_f [PRO-Z_f to be sold a house on the beach]

According to the θ -role analysis, sentences (27)-(30) should all be well-formed with the PRO coindexed with *Mary*.

(27) * Mary_f was promised [PRO-C_f to rent a house on the beach]

(28) Mary_f was promised [PRO-C_f to be rented a house on the beach]

(29) * Mary_f was promised [PRO-Z_f to buy a house on the beach]

(30) Mary_f was promised [PRO-Z_f to be sold a house on the beach]

The implications of sentences (31) and (32) pose an additional problem for the Thematic Role Theory.

(31) John_f promised Mary [PRO_f to buy himself/*herself the house]

(32) John_f promised Mary [PRO_f to sell the house to her/*him=John]

According to the Thematic Identity Condition, the subject of *promise* and the subject of *buy* have the same θ -role in (31), and the subject of *promise* and the subject of *sell* have the same θ -role in (32). One might reason that this implies that the subject of *buy* and the subject of *sell* have the same θ -role.²³

²³ Ravin (1990) offers the following observations concerning the role of thematic-roles in syntax:

The semantic nature of theta-role representations is less clear. This is because the semantic content of theta-roles is never discussed by Chomsky in *LGB*. There is no mention of how many theta-roles need to be posited or of what semantic concepts they should represent. Elsewhere, this omission is acknowledged:

It's interesting that, although as you say the theta role has been brought back in, one thing has not been brought back in, and that is the difference between the different thematic relations. Nothing

The verb *ask* with a *wh*- complement is subject control, [+SC] in the Chomsky (1980) system discussed by Ruzicka, and therefore, according to Ruzicka, must obey the Thematic Identity Condition. Notice in the following that all are well-formed with PRO coreferential with the external argument of *ask*.

- (33) John_i asked Bill [how PRO-A_i to rent a house to him]
 (34) John_i asked Bill [how PRO-C_i to rent a house from him]
 (35) John_i asked Bill [how PRO-X_i to sell a house to him]
 (36) John_i asked Bill [how PRO-Z_i to buy a house from him]

If the TIC is satisfied in each case, then the verbs *buy* and *sell* have the same θ -role in subject position, i.e. θ -role X is identical to θ -role Z in sentences (35)-(36) and (20)-(22). Similarly for *rent to/rent from*, the θ -roles A and C are identical in (33)-(34) and (17)-(19). This can only be the case if *buy/sell* and *rent/rent* assign the same θ -role two times.

has been said here about the differences between Goal and Source for example. Part of the reason is I don't understand this very well. I never know how people are able to pick out thematic relations with such security, I can't. (Chomsky 1982b:88-9)

The refusal to discuss the semantic nature of theta-roles seems less surprising when one realizes that the real function of roles in the GB system is purely syntactic and has very little to do with semantics. Indeed, this is often confirmed in the GB literature. Van Riemsdijk and Williams (1986) write: "Though we will not go into the semantic content of such theta-roles, we will use the term to identify the arguments of predicates" (p.30); and "This terminology implies a system of argument *types*, in that, for example, it implies that the agent arguments of two different verbs have something in common. Although this may be true, theta-theory as outlined here is not committed to this idea" (p. 241).
 (Ravin 1990:65)

A final problem which is inherent in any thematic role theory is illustrated in the following sentences. In (37) and (39), PRO is coreferential with *John*, in (38) with *Bill*, and in (40) with *Mary*.

- (37) John_i offered Mary [PRO_i to put himself and his car at her disposal]
- (38) John offered Mary_i a thousand dollars [PRO_i to put herself and her car at his disposal]
- (39) John_i offered Mary [PRO_i to put himself/*herself in debt by cosigning her/*his note]
- (40) John offered Mary_i a lot [PRO_i to put herself/*himself in debt by cosigning his/*her note]

In any thematic role theory, PRO has the same θ -role in each case. It would appear that the θ -roles assigned to the NP (*John, Bill/Mary*) by *offer* are very similar, if not the same, in the four examples (37)-(40). The fact that PRO has different reference argues against the thematic role theory.

9.3 THE DOMAIN OF LICENSING OF PRO

I will examine the control properties of verbs which occur in constructions like (41) and have the lexical structures (42). For reasons which will become clear, I call these *V₂ verbs*.

- (41) [CP[NP John][IP will [VP promise [NP Mary] [CP PRO to leave]]]]
 persuade
 encourage
 allow
 cause
 ...

- (42) V₂ verb
 external θ -role
 NP (or PP), CP
 ...

V₂ verbs include: allow, invite, ask, remind, persuade, order, command, encourage, cause, train, convince, authorize, commission, bribe, send,

prepare, contract, drive, direct, elect, empower, engage, prepare, exhort, forbid, intend. This list is not exhaustive.

With the single exception of *promise*, all V_2 verbs are object control verbs.

- (43) i. John persuaded Mary_{*i*} [PRO_{*i*} to remind herself/*himself [PRO to study]]
 ii. John persuaded Mary_{*i*} [PRO_{*i*} to be reminded *e_i* [PRO to study]]
- (44) i. John encouraged Mary_{*i*} [PRO_{*i*} to remind herself/*himself [PRO to study]]
 ii. John encouraged Mary_{*i*} [PRO_{*i*} to be reminded *e_i* [PRO to study]]
- (45) i. John allowed Mary_{*i*} [PRO_{*i*} to remind herself/*himself [PRO to study]]
 ii. John allowed Mary_{*i*} [PRO_{*i*} to be reminded *e_i* [PRO to study]]
- (46) i. John caused Mary_{*i*} [PRO_{*i*} to ask herself/*himself [PRO to question Bill's motives]]
 ii. John caused Mary_{*i*} [PRO_{*i*} to be asked *e_i* [PRO to question Bill's motives]]

A subset of the V_2 verbs are the V_1 verbs. V_1 verbs may be a semantically definable subset. V_1 verbs involve some request (*beg, plead, pray to, ask*) or incur some obligation (*promise*). They do not impose one's will (*command, persuade, convince, etc.*).

V_1 verbs include: *promise, beg, plead, pray to, ask, appeal to, apply to, call upon, petition, pressure.*

Since V_1 verbs are a subset of V_2 verbs, V_1 verbs are object control, with the exception of *promise*.

- (47) John_{*i*} promised Mary [PRO_{*i*} to perjure himself/*herself]
 (48) John pleaded with Mary_{*i*} [PRO_{*i*} to perjure herself/*himself]
 (49) John begged Mary_{*i*} [PRO_{*i*} to perjure herself/*himself]

(50) John prayed to Mary_{*f*} [PRO_{*f*} to forgive him/*her(=Mary)]

Unusual control properties occur when V_1 verbs have V_2 verbs in an embedded passive clause. The Control Module can assign the features and reference to PRO only if it has information about the verb in the matrix clause (V_m), the verb in the embedded clause (V_e), and the presence or absence of the passive morpheme (*be+en*) in the embedded clause. By the V_1 - V_2 hypothesis in (52), control of PRO in the structure (51) is a function of whether the matrix verb, V_m , is in the verb class V_1 , whether the embedded verb, V_e , is in the verb class V_2 , and the presence or absence of the passive morpheme, *be+en*, in the embedded clause.

(51) [CP NP V_m (P) NP [CP [PRO to (be) V_e ...



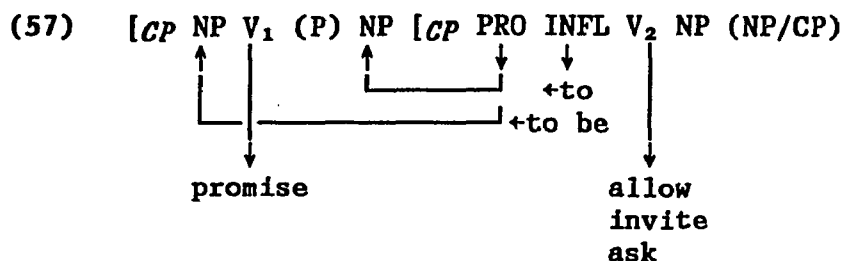
(52) V_1 - V_2 Control Hypothesis:

Control properties of PRO = function(V_m , V_e , *be+en*).

In the structure (51), when $V_m=V_1$ and $V_e=V_2$, control of V_1 switches if V_2 is passive, i.e. (*be+en*) is present in IP₂:

- (a) if V_1 is lexically subject control, it is object control when V_2 is passive.
- (b) if V_1 is lexically object control, it is subject control when V_2 is passive.

Consider the role of the passive morpheme in the embedded sentence. The V_1 verbs *plead with*, *beg*, and *pray to* are object control if V_2 is active and subject control if V_2 is passive.



- (58) i. John_i promised Mary [PRO_i to perjure himself/*herself]
 ii. John promised Mary_i [PRO_i to be allowed e_i [PRO_i to perjure herself/*himself]]
 iii. John_i promised Mary [PRO_i to allow himself/*herself [PRO_i to have a good time]]

The control properties of the V₁ verbs shift for any V₂ verb that is in the complement. In the following sentences, V₂ is *allow*, *invite*, *ask*, and *remind*.

- (59) i. John prayed to Saint Peter_i [PRO_i to allow sinners [PRO to enter]]
 ii. John_i prayed to Saint Peter [PRO_i to be allowed e_i [PRO to enter]]
- (60) i. John_i promised Mary [PRO_i to invite Susan [PRO to sing]]
 ii. John promised Mary_i [PRO_i to be invited e_i [PRO to sing]]
- (61) i. John begged Mary_i [PRO_i to ask herself/*himself [PRO to question Bill's motives]]
 ii. John_i begged Mary [PRO_i to be asked e_i [PRO to question Bill's motives]]
- (62) i. John pleaded with Mary_i [PRO_i to remind herself/*himself [PRO to study]]
 ii. John_i pleaded with Mary [PRO_i to be reminded e_i [PRO to study]]

Since V₁ verbs are a subset of V₂, V₁ verbs also function as V₂ verbs. For example, the V₁ verb *beg* functions as a V₂ verb, in (64) and (65). Control switches in the (ii) examples when V_e contains passive morphology.

- (64) i. John_f promised Mary [PRO_f to beg her/*him(=John) [PRO to give money]]
 ii. John promised Mary_f [PRO_f to be begged e_f [PRO to give money]]
- (65) i. John_f promised Mary [PRO_f to beg her/*him(=John) for money on the subway]
 ii. John promised Mary_f [PRO_f to be begged e_f for money on the subway]

The adjacent pair V₁, V₂ can be embedded any depth in the sentence:

- (66) i. John promised Mary [PRO to beg Sue_f [PRO_f to allow Bill [PRO to go]]]
 ii. John promised Mary_f [PRO to beg Sue [PRO_f to be allowed e_f [PRO to go]]]
- (67) i. John persuaded Mary [PRO to beg Sue_f [PRO_f to allow Bill [PRO to go]]]
 ii. John persuaded Mary_f [PRO_f to beg Sue [PRO_f to be allowed e_f [PRO to go]]]

The V₁-V₂ mechanism operates only if the embedded sentence contains the passive element *be+en*; it does not function if the embedded sentence contains *be* in an adjective construction. Notice that (68) contains *be* in the embedded sentence, but only (69) contains *be+en*. Hence, only in (69) does the V₁-V₂ mechanism mark control in the phrase marker the reverse of the lexical entry.

(68) John_f promised Mary [PRO_f to be honest]

(69) John promised Mary_f [PRO_f to be invited e_f]

9.4 THE LEXICAL V₁-V₂ MECHANISM OF CONTROL

Verbs in the V₂ class, which license internal and external θ -roles and a CP infinitive complement, play a crucial role in deciding how the grammar must be organized to generate and interpret sentences containing

infinitives with PRO as subject. The control module assigns PRO its coreference and features in the phrase marker. The lexical mechanisms of the control module answer the question: How is information in the lexical entries of items used to assign PRO its reference and features? The lexical entry of an item indicates the information about the control properties of the item. The V_1 - V_2 mechanism indicates how information in the the lexical entries is utilized to enter control information about PRO into the phrase marker.

When we embed one verb as the complement of another, we obtain sentences like these:

- (70) i. John will ask Mary to encourage Bill to go.
 ii. John will be allowed to ask Bill to go.
 iii. John will be asked how to be allowed to go.
 iv. John will be promised to be allowed to go.
 v. John will plead with Mary to be asked to go.
 vi. John will beg Mary to be promised to go.

We assume that the lexicon contains the following entries for verbs like those in (70). For each entry, the lexicon contains the information indicated in (71).

(71) word
 external θ -role
 internal arguments: unordered complement(s)
 Control: verb class if any, e.g. V_1 or V_2 ;
 PRO coindexed with internal or external argument;
 Extrinsic feature assignment to PRO

(72) *promise*
 external θ -role
 internal: NP, CP
 Control: V_1 ; external; PRO[-gen,+def]

(73) *plead*
 external θ -role
 internal: PP_{with}, CP
 Control: V_1 ; internal; PRO[-gen,+def]

- (74) *beg*
 external θ -role
 internal: NP, CP
 Control: V_1 ; internal; PRO[-gen,+def]
- (75) *pray*
 external θ -role
 internal: PP_{to}, CP
 Control: V_1 ; internal; PRO[-gen,+def]
- (76) *allow*
 external θ -role
 internal: NP, CP
 Control: V_2 ; internal; PRO[-gen,+def]
- (77) *ask₁*
 external θ -role
 internal: NP, CP
 Control: V_1 ; internal; PRO[-gen,+def]
- (78) *ask₂*
 external θ -role
 internal: NP, CP_{wh}-
 Control: V_2 ; external; PRO[+gen,+def]
- (79) *explain*
 external θ -role
 internal: PP_{to}, CP
 Control: V_2 ; internal; PRO[+gen,+def]
- (80) *invite*
 external θ -role
 internal: NP, CP
 Control: V_2 ; external; PRO[+gen,+def]

The tree building mechanisms of the grammar which assign structure to strings will assign the labeled bracketing (81), representing the more detailed phrase marker in Figure 37 on page 355. In this figure, I follow the convention of representing the passive participial as a V. However, it may well be the case that the lexical entry for the passive morphology is itself a V with no external θ -role assignment and which licenses an adjectival passive participle.

- (81) [CP NP (be) V_m (P) NP [CP (wh-) [IP PRO to (be) V_e (P) NP CP]]]
- | | |
|-----------|-----------|
| ↓ | ↓ |
| ask | ask |
| allow | allow |
| encourage | encourage |
| promise | promise |
| beg | beg |

The phrase marker assigned by the grammar to any one, S_x , of the sentences listed in (70) is identical to the phrase marker assigned to any other, S_y , of the sentences in (70) with three possible differences: (I) The constituents of the terminal string of S_x may be in a different order than those of S_y . (II) There may be empty categories in one sentence that are not in another. (III) The element *be+en* may be present in one sentence and not the other.

The control module uses the information about the control properties of a lexical item to decide on the properties of PRO in the matrix sentence. For sentences like those in (70), the control module uses the V_1 - V_2 Mechanism.

Lexical Mechanisms of the Control Module: The lexical entry of an item indicates the reference and assigns features to PRO in complements it licenses. In utilizing lexical information, the control module will first mark in the phrase marker any coindexing of PRO and second will assign PRO in the phrase marker its feature combination. If it is impossible to satisfy the required lexical coindexing for an item, lex_n , the features of lex_n will not be marked on PRO in the phrase marker.

V_1 - V_2 lexical interpretation mechanism: In the structure Figure 37, produced by the tree building mechanisms:

- IF:**
- (a) if V_m is V_1 and V_e is V_2 , and
 - (b) if *beten* is present in IP_e ,
- THEN:** mark PRO for the reverse the lexical specification of control for V_m :
- (c) PRO is lexically coindexed with the external argument of V_m , it must be coindexed with the internal argument,
 - (d) if PRO is lexically coindexed with the internal argument of V_m , it must be coindexed with the external argument.
- OTHERWISE:** In the phrase marker, mark the indices on PRO and the NP to satisfy the lexical markings on V_m which require PRO to be coindexed with the internal or external argument of V_m .

Basically, the claim of the V_1 - V_2 Mechanism is that control of PRO is not a property of the matrix verb alone, nor is it a property of the matrix verb and the embedded verb, nor is it a property of the matrix verb and grammatical formatives like *beten*. In order to determine the antecedent of PRO in a construction where a V_1 verb is in the complement of a V_2 verb, the control module must simultaneously consider three specific pieces of information about the structure: (a) whether V_e is V_2 , (b) whether V_m is V_1 , and (c) whether the embedded CP contains the passive morpheme *beten*.

Looked at another way, suppose that the mechanism which determined the antecedent of PRO were a computer program which had available to it two types of information: (a) the categorial structure given in phrase marker Figure 37 on page 355, and (b) a small window which permitted the scanner to consider one lexical item in the terminal string at a time. Such a scanner would never be able to determine the reference of PRO. In order for a computer program to determine the antecedent of PRO, it must consider (a) the categorial structure given in phrase marker

Figure 37 on page 355, and (b) at least two verbs, V_m and V_e and the grammatical formative *be+en* in the terminal string.

Let us consider the steps by which the grammar assigns control to the following strings:

- (82) John promised Mary to allow Sue to go.
- (83) John promised Mary to be allowed to go.

- (84) John allowed Mary to promise Sue to go.
- (85) John allowed Mary to be promised to go.

- (86) John begged Mary to allow Sue to go.
- (87) John begged Mary to be allowed to go.

- (88) John allowed Mary to invite Sue to go.
- (89) John allowed Mary to be invited to go.

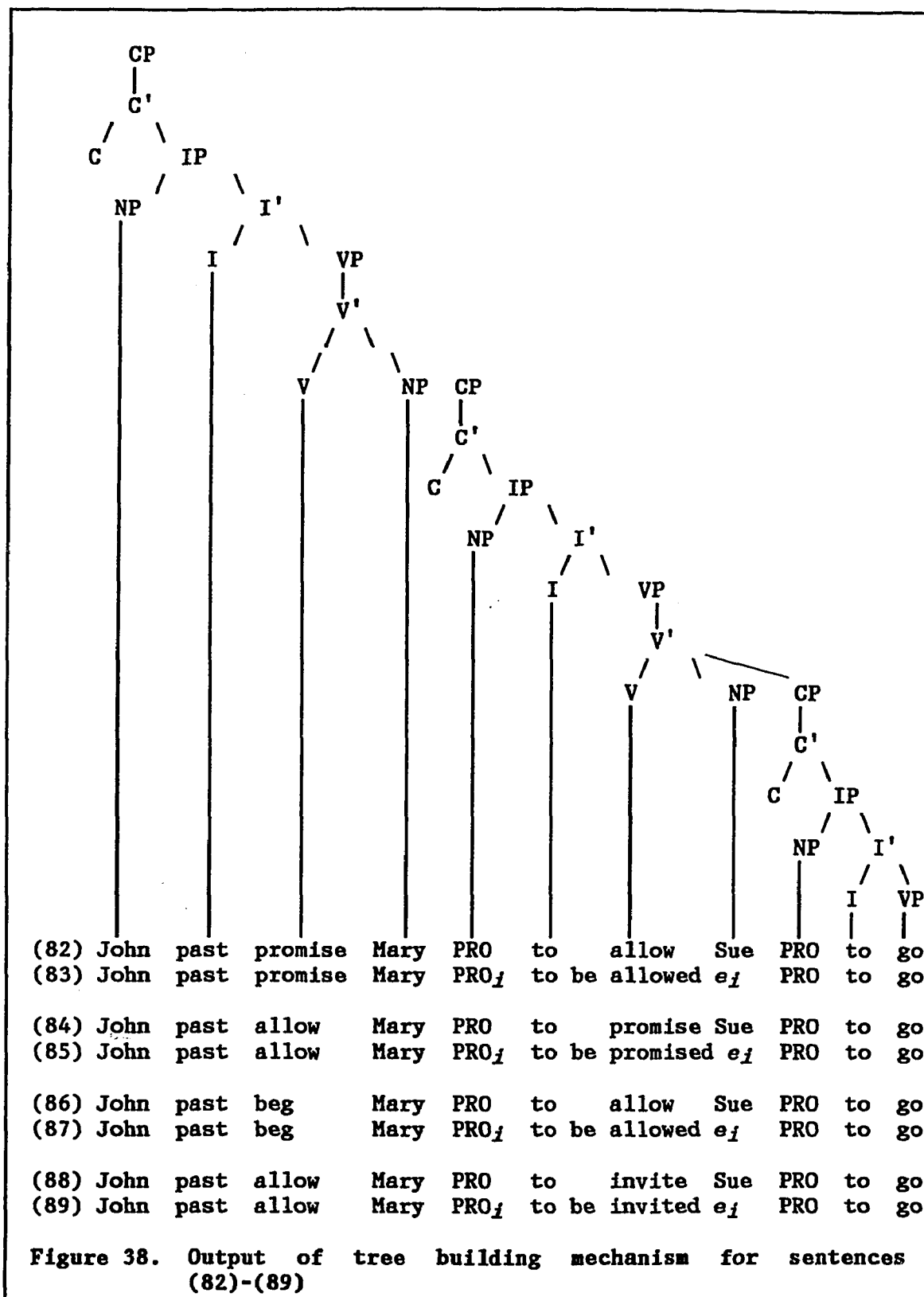
The tree building mechanisms utilize information in the lexicon and principles of universal grammar to construct a phrase marker and assign it to a string. The tree building mechanism will assign the phrase marker in Figure 38 on page 359 to the above strings. Binding will operate to indicate that some of the empty categories are coindexed with an NP. The element $PRO_{[+a,+p]}$ will be inserted in ungoverned positions.

The input to the control module is a string containing $PRO_{[+a,+p]}$ and the phrase marker constructed for that string by the tree building and binding mechanisms. The control module is only concerned with certain aspects of the tree, in particular, those structural aspects indicated in (82')-(89'):

- (82') [John promised Mary [PRO to allow Sue [PRO to go]]]
- (83') [John promised Mary [PRO_i to be allowed e_i [PRO to go]]]

- (84') [John allowed Mary [PRO to promise Sue [PRO to go]]]
- (85') [John allowed Mary [PRO_i to be promised e_i [PRO to go]]]

- (86') [John begged Mary [PRO to allow Sue [PRO to go]]]



- (87') [John begged Mary [PRO_i to be allowed e_i [PRO to go]]]
 (88') [John allowed Mary [PRO to invite Sue [PRO to go]]
 (89') [John allowed Mary [PRO_i to be invited e_i [PRO to go]]]

In all of these cases, the mechanism of the control module which assigns PRO its reference and features is the V₁-V₂ mechanism. Each of these sentences meets the structural condition of the V₁-V₂ mechanism:

(90) [CP NP (be) V_m (P) NP [CP (wh-) [IP PRO to (be) V_e (P) NP CP]]]

(91) control of PRO = function(V_m, V_e, be+en)

(a) if V_m = V₁, V_e = V₂, and IP₂ contains be+en, then mark PRO in phrase marker the reverse of the lexical markings of V_m,

(b) otherwise, mark PRO to match the lexical entry of V_m.

Sentence (82) is assigned the structure (82') by the tree building mechanism.

(82) John promised Mary to allow Sue to go.

(82') [John promised Mary [PRO to allow Sue [PRO to go]]]
 V_m V_e

V_m = V₁, V_e = V₂, but be+en is not in IP₂, so the V₁-V₂ mechanism marks PRO in the phrase marker to match the lexical entry of V_m. V₁-V₂ first coindexes PRO with the external argument of *promise*, yielding (82a).

(82a) [John_i promised Mary [PRO_i to allow Sue [PRO to go]]]

V₁-V₂ then marks the features from the entry for V_m on PRO, yielding the output of the V₁-V₂ mechanism, shown in (82b).

(82b) [John_i promised Mary [PRO_i[-g,+d] to allow Sue [PRO to go]]]

The structure [PRO_i[-g,+d] to allow Sue [PRO to go]] is examined by the V₁-V₂ mechanism with the result: V_m (*allow*) = V₂, V_e (*go*) ≠ V₂, IP_e does not contain be+en. Therefore, PRO is coindexed to accord with the

(86b) [John begged Mary [PRO to allow Sue_i [PRO[-g,+d]_i to go]]]

V_1 - V_2 analyzes the structure [John begged Mary [PRO to allow Sue_i]], where $V_m = V_1$, $V_e = V_2$, and IP_e does not contain *be*+*ten*. Hence, PRO is coindexed with the internal argument of *beg*, yielding (86c).

(86c) [John begged Mary_j [PRO_j to allow Sue_i [PRO[-g,+d]_i to go]]]

Since V_1 - V_2 successfully indexed PRO with an NP, V_1 - V_2 copies the features from the lexical entry of *beg* onto PRO, yielding (86d) as the output of the control module.

(86d) [John begged Mary_j [PRO[-g,+d]_j to allow Sue_i [PRO[-g,+d]_i to go]]]

In sentence (87), the tree building mechanism assigns the indexed structure in (87').

(87) John begged Mary to be allowed to go.

(87') [John begged Mary [PRO_i to be allowed _{V_m} e_i [PRO to go]]] _{V_e}

$V_m = V_2$, $V_e \neq V_2$, and IP_e does not contain *be*+*ten*, hence V_1 - V_2 coindexes the PRO in the phrase marker with the internal argument of *allow*. Since this argument already has an index, the index is copied onto PRO, yielding (87a).

(87a) [John begged Mary [PRO_i to be allowed e_i [PRO_i to go]]]

Since V_1 - V_2 was successful in coindexing PRO, it copies the features from the lexical specification of *allow* onto the phrase marker, yielding (87b).

(87b) [John begged Mary [PRO_i to be allowed e_i [PRO[-g,+d]_i to go]]]

V_1 - V_2 analyzes the structure [John begged Mary [PRO_i to be allowed e_i]] to obtain $V_m = V_1$, $V_e = V_2$, and IP_e contains *be*+*ten*. Hence, V_1 - V_2 coindexes PRO in the phrase marker with the external argument of *beg*.

Since PRO already has an index, this index is copied onto *John*, yielding (87c).

(87c) [John_{*i*} begged Mary [PRO_{*i*} to be allowed *e_i* [PRO[-*g*,+*d*]_{*i*} to go]]]

Since V_1 - V_2 successfully coindexed PRO, it copies the features from the lexical entry for *beg* onto PRO in the phrase marker yielding (87d) as the output of the control module.

(87d) [John_{*i*} begged Mary [PRO[-*g*,+*d*]_{*i*} to be allowed *e_i* [PRO[-*g*,+*d*]_{*i*} to go]]]

Sentence (88) is assigned the structure (88') by the tree building mechanism.

(88) John allowed Mary to invite Sue to go.

(88') [John allowed Mary [PRO to invite Sue [PRO to go]]]
 V_m V_e

The V_1 - V_2 mechanism to indicate $V_m = V_2$, $V_e \neq V_2$, IP_e does not contain *be+en*. PRO in the phrase marker is coindexed with the internal argument of *invite* to accord with the lexical specification of *invite*.

(88a) [John allowed Mary [PRO to invite Sue_{*i*} [PRO_{*i*} to go]]]

The features are copied from the entry for *allow* to yield (88b).

(88b) [John allowed Mary [PRO to invite Sue_{*i*} [PRO[-*g*,+*d*]_{*i*} to go]]]

V_1 - V_2 analyses [*John allowed Mary [PRO to invite Sue_{*i*}]*] obtaining: $V_m = V_2$, $V_e = V_2$, IP_2 contains *be+en*. Hence, PRO in the phrase marker is coindexed with the internal argument of *allow* to match the lexical specification of *allow*, as in (88c).

(88c) [John allowed Mary_{*j*} [PRO_{*j*} to invite Sue_{*i*} [PRO[-*g*,+*d*]_{*i*} to go]]]

Since V_1 - V_2 is successful on the coindexing, it assigns the features to PRO yielding (88d) as the output of the control module.

(88d) [John allowed Mary_j [PRO[-g,+d]_j to invite Sue_i [PRO[-g,+d]_i to go]]]

Sentence (89) is assigned the indexed structure in (89').

(89) John allowed Mary to be invited to go.

(89') [John allowed Mary [PRO_i to be invited _{V_m} e_i [PRO to go]]] _{V_e}

$V_m = V_2$, $V_e \neq V_2$, IP_e does not contain *be+en*, hence PRO in the phrase marker is coindexed with the internal argument of *invite* according to the lexical entry for *invite*. Since the internal argument of *invite* bears an index, this index is copied onto PRO yielding (89a).

(89a) [John allowed Mary [PRO_i to be invited e_i [PRO_i to go]]]

Successful on the coindexing, V_1-V_2 copies the features from the entry of *invite* onto PRO yielding (89b).

(89b) [John allowed Mary [PRO_i to be invited e_i [PRO[-g,+d]_i to go]]]

V_1-V_2 now applies to the structure [John allowed Mary [PRO_i to be invited e_i]] with the analysis $V_m = V_2$, $V_e = V_2$, IP_2 contains *be+en*. According to V_1-V_2 , PRO in the phrase marker must be coindexed to meet the lexical specifications of *allow*. PRO is already coindexed with the internal argument of *allow*. Since the PRO already bears an index, this is copied onto the internal argument of *allow* yielding (89c).

(89c) [John allowed Mary_i [PRO_i to be invited e_i [PRO[-g,+d]_i to go]]]

With coindexing successful, V_1-V_2 copies the PRO features of the lexical item *allow* onto the phrase marker yielding (89d) as the output of the control module.

(89d) [John allowed Mary_i [PRO[-g,+d]_i to be invited e_i [PRO[-g,+d]_i to go]]]

The V_1 - V_2 mechanism scans (92') with the result that $V_m = V_1$, $V_e = V_2$, and IP_e does not contain *be+en*. This is the condition under which V_1 - V_2 must coindex PRO to match the lexical specification of the matrix verb. The lexical entry of the matrix verb *promise* requires PRO to be coindexed with the external θ -role. But in structure (92'), *promise* does not license an external θ -role.²⁴ The passive element *be+en* suppresses external θ -role assignment. Hence, V_1 - V_2 cannot coindex PRO with the lexically defined external θ -role of *promise*. Since V_1 - V_2 cannot coindex PRO, V_1 - V_2 cannot copy the features of the lexical entry onto PRO in the phrase marker.

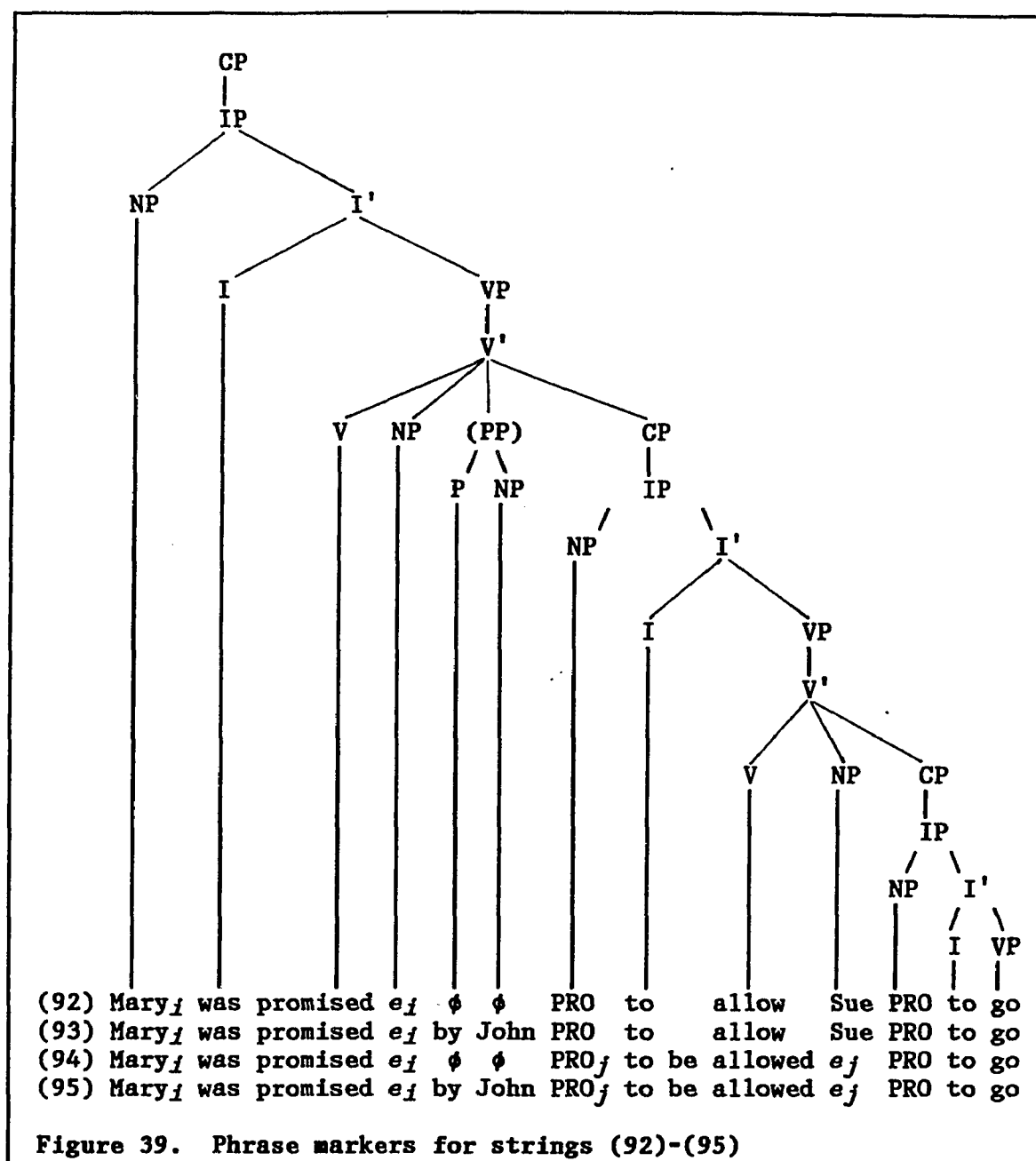
V_1 - V_2 is a mediation mechanism between the phrase marker produced by the tree building mechanisms and the control information stored in lexical entries. V_1 - V_2 cannot coindex PRO in the phrase marker because the lexical entry of *promise* does not license any coindexing in structure (92'). Thus, PRO has no features and is uninterpretable, or ill-formed according to the control module.

Let us now look at sentence (93) and its indexed structure (93').

(93) * Mary was promised by John to allow Sue to go.

(93') [Mary_{*i*} was promised V_m *e_{*i*}* by John [PRO to allow Sue [PRO to go]]] V_e

²⁴ We must distinguish between whether a verb is lexically marked to permit/license an external θ -role, and whether a verb in a particular phrase marker (active or passive) licenses an external θ -role. We must also differentiate between (a) external theta role and (b) subject. Every sentence must have a subject, but every verb does not need an external θ -role. In some phrase marker's (passive) a verb may not license an external θ -role, although the sentence must have a subject.



The V_1 - V_2 mechanism finds the same situation as with (92). $V_m = V_1$, $V_e = V_2$, and IP_e does not contain *be+en*. Therefore, V_1 - V_2 must coindex PRO to match the lexical specification of the matrix verb. V_m (*promise*) requires PRO to be coindexed with the external θ -role of *promise*. However, the grammatical formative *be+en* in the matrix sen-

tence causes the external θ -role of *promise* to be suppressed. Thus, as in (92), V_1 - V_2 cannot coindex PRO, and PRO cannot be assigned features. PRO with no features is uninterpretable.

V_1 - V_2 cannot coindex PRO with *John* in the *by*-phrase because V_1 - V_2 only uses information in the lexical specification of V_m and V_e , and the lexical entry of a verb does not license *by*-phrases. The *by*-phrase is licensed by *beten* in the matrix sentence, not by V_m . *John* in (93) obtains case and a θ -role from the element *by*.

Both the V_1 - V_2 theory and the Thematic theory make the claim that information from the lexical entries of *promise* and *allow* is relevant in deciding the grammaticality and the interpretation of PRO in sentences (92)-(93). But they make very different claims about what information is relevant and how that information is used by the grammar. In particular, the two theories take very different approaches to the questions in (96).

(96)

1. What is the structure and content of lexical information concerning the coindexing and interpretation of PRO?
2. What are the mechanisms which relate PRO in the phrase marker to the lexical entries?
3. What is the relevance of θ -roles for control theory?

Concerning the questions in (96), the V_1 - V_2 theory claims the following:

1. The lexical information about PRO is stored in terms of the internal and external θ -roles of the item.
2. The V_1 - V_2 mechanism is concerned with whether there is an internal or external θ -role linked to the verb, and whether the lexical entry of the verb licenses coindexing PRO with the internal or external θ -role.

3. It is crucial whether a verb is lexically marked for an external θ -role and whether the verb licenses an internal θ -role. However, the actual content or actual θ -role assigned as an internal or external argument (agent, patient, instrument, etc.) is irrelevant.

Concerning questions in (96), the Thematic theory makes these claims.

1. The lexical information relevant to determining the antecedent of PRO is stored in terms of specific θ -roles assigned to lexical items. Each item (*promise, allow*) must be lexically marked to indicate the specific θ -roles assigned (agent, instrument...). Whether the θ -roles are internal or external is not crucial to the determination of the antecedent of PRO.
2. The mechanisms which determine the antecedent and interpretation of PRO are concerned with (1) the θ -role assigned to PRO, (2) the θ -role assigned to each of the NPs in the matrix sentence, and (3) whether the verb requires thematic identity or thematic distinctness.
3. The actual θ -roles assigned by V_m and V_e are crucial. However, whether any particular role is internal or external, or whether a V licenses an external θ -role in a particular phrase marker, is irrelevant.

The implications of these different approaches are illustrated with the following sentences.

(97) John promised Mary to allow Sue to go.

(98) * Mary was promised by John to allow Sue to go.

(99) * Mary was promised to allow Sue to go.

The V_1 - V_2 theory claims that the correct distribution of grammaticality and interpretation follows from lexical specification of internal and external θ -roles, independent of what those θ -roles are. The lexical entry of *promise* requires PRO to be coindexed with the external argument of *promise*. In the phrase marker of sentence (97), *promise* licenses an external θ -role, and the sentence is well-formed: V_1 - V_2 can coindex PRO with the external argument. In the phrase markers assigned to (98) and (99), *promise* does not license an external θ -role (as per the definition of passive), hence V_1 - V_2 cannot coindex PRO with the external argument of *promise*. The correct data follows from the assumptions of the V_1 - V_2 theory. If sentence (98) were well-formed with the same interpretation as (97), this would be a strong argument against the V_1 - V_2 theory.

The Thematic Role Theory claims that the correct distribution of data follows from lexical specification of precise θ -roles independent of whether they are internal or external. This theory predicts that (97) and (98), in which *John* presumably has the same θ -role in each case, should be well-formed with the same interpretation, but that (99), which has no appropriate θ -role, should be ill-formed. Sentences like (98), with an overt *by*-phrase, appear to be counterexamples to the Thematic Theory since they should behave exactly like the corresponding actives. Ruzicka does not discuss these examples. He discusses only passives with no *by*-phrase.

Let us look at the discussion by Chomsky (1980) which began this chapter (emphasis added by me):

The point is now clear under the analysis of these forms presented above. There are some curious exceptions to this prin-

ciple. Hust and Brame (1976) and Solan (1977) cite such double-passives as *John was promised to be allowed to leave*, which should be blocked but is not for reasons that are unclear. *It seems that some syntactic property is involved*, since substitution of a near synonym for the embedded passive changes the grammatical status to the expected *: e.g. **John was promised to get permission to leave*. (Chomsky 1980:35 fn.40)

Chomsky hits the nail on the head. The correct distribution of data follows from syntactic properties in the phrase marker: the presence or absence of the passive *be+en* in CP_e , and the presence or absence of an external θ -role for V_m . Whether sentences are near synonyms or have the same θ -roles is not relevant.²⁵

Consider sentence (94) which is assigned the indexed structure (94') by the tree building mechanism.

(94) Mary was promised to be allowed to go.

(94') [$Mary_i$ was promised e_j [PRO_j to be allowed e_j [PRO to go]]]
 V_m V_e

The control module advances through these steps:

$V_m = V_1$, $V_e = V_2$, and IP_e contains *be+en*. This is the condition under which V_1 - V_2 must coindex PRO to be the reverse of the lexical

²⁵ By syntactic properties I refer to the principles and mechanisms internal to the tree building mechanisms, binding theory, and control theory, plus those elements of lexical entries which are used by these three modules. Syntactic properties include: (1) The requirement that CP must contain a subject and predicate, but NP does not need to contain a subject and predicate. (2) A verb may license an external θ -role in its lexical entry. (3) A verb, which lexically is permitted to license an external θ -role, may not license an external θ -role in some phrase marker, e.g. passive constructions.

Semantic properties certainly exist and are relevant for the characterization of the principles of sentence construction. For instance, V_1 verbs are a subset of the V_2 verbs. It may prove that V_1 verbs are a semantically definable subset involving some concepts of requesting action or requesting information. It is possible that the concepts differentiating V_1 from V_2 verbs might be formalizable in terms of a general theory of the semantics of *request-type* verbs.

specification of the matrix verb. The lexical entry of the matrix verb *promise* requires PRO to be coindexed with the external θ -role. V_1 - V_2 reverses this lexical marking and attempts to coindex PRO in the phrase marker with the internal θ -role of *promise*.

In structure (94') the internal θ -role of *promise* is carried by e_j . Since the internal θ -role already has an index (i), this index must be copied onto the PRO subject of *allow*. Since the PRO subject of *allow* bears an index (j), the indexing mechanism must indicate that $i = j$. Let us assume that all j 's are replaced by i , yielding (94a).

(94a) [Mary _{i} was promised e_i [PRO _{i} to be allowed e_i [PRO to go]]]

Since V_1 - V_2 was successful in coindexing PRO, the features of PRO are copied onto PRO in (94a) from the entry of *promise*, yielding (94b).

(94b) [Mary _{i} was promised e_i [PRO[- g ,+ d] _{i} to be allowed e_i [PRO to go]]]

V_1 - V_2 analyses the structure [PRO to be allowed e_i [PRO to go]]. $V_m = V_2$, $V_e \neq V_2$, and IP_e does not contain *be+en*, so V_1 - V_2 coindexes PRO to accord with the lexical specifications of *allow*: PRO is coindexed with the internal argument of *allow*. Since coindexing is successful, the features from the entry of *allow* are copied onto PRO yielding (94c) as the output of the control module.

(94c) [Mary _{i} was promised e_i [PRO[- g ,+ d] _{i} to be allowed e_i [PRO[- g ,+ d] _{i} to go]]]

Sentence (95) is virtually identical to sentence (94). As discussed above, the presence of a *by*-phrase does not affect the steps of the control module. The input to the control module is (95'). The output of the control module for (95) is (95a):

(95) Mary was promised by John to be allowed to go.

(95') [Mary_i was promised e_i by John [PRO_j to be allowed e_j [PRO to go]]]

(95a) [Mary_i was promised e_i by John [PRO[-g+d]_i to be allowed e_i [PRO[-g,+d]_i to go]]]

9.6 THE V₁-V₂ CONTROL THEORY VERSUS THE THEMATIC ROLE THEORY.

How much information, and what type of information, must the Control Module have from a string containing PRO to determine the interpretation and reference of PRO? The Thematic Role Theory and the V₁-V₂ Theory make different claims about the factorization of a string required by the Control Module.

The Thematic Role Theory claims that the Control Module can obtain sufficient information to determine the interpretation and reference of PRO from the following factorization of a string containing PRO. The presence or absence of the passive morpheme and the presence or absence of *wh*- plays no role. The only relevant factor is the identity or non-identity of the θ -role of PRO, θ_3 , with the θ -roles assigned by V_m in the matrix sentence.

(100) Factorization of a string according to the Thematic Role Theory

$$\begin{array}{cccc} [CP & NP & V_m & (P) & NP & [CP & PRO & to & V_e]] \\ & \theta_1 & & & \theta_2 & & \theta_3 & & \end{array}$$

control of PRO = function (θ_1 , θ_2 , θ_3 , verb class of V_m)

The V₁-V₂ Hypothesis claims that the Control Module can obtain sufficient information to assign PRO its interpretation and reference from factorization (101). The V₁-V₂ Control Module will use information about the presence or absence of *be+en* from (101) plus information from the lexical entries of V_m and V_e concerning their verb class (V_1 or

V₂), the PRO features, reference indicated in the lexical entries, and the presence or absence of *wh*-.

(101) Factorization of a string according to the V₁-V₂ Theory

[CP NP V_m (P) NP [CP (*wh*-) [PRO to (be) V_e

control of PRO = function (V_m, V_e, *be*+en)

The following sentences provide some data which is relevant in choosing between these two theories of control. As predicted by the V₁-V₂ hypothesis, control in (102)-(103) shifts when *be*+en is present in the embedded sentence. Since PRO in each case has the same θ -role, the Theta Role hypothesis predicts, contrary to fact, that control should be the same in the (102)-(103).

(102) i. John pleaded with Mary_{*f*} [PRO_{*f*} to rent a house on the beach]

ii. John_{*f*} pleaded with Mary [PRO_{*f*} to be rented a house on the beach]

iii. John pleaded with Mary_{*f*} [PRO_{*f*} to buy (herself/*himself) a house on the beach]

iv. John_{*f*} pleaded with Mary [PRO_{*f*} to be sold a house on the beach]

(103) i. John begged Mary_{*f*} [PRO_{*f*} to rent a house on the beach]

ii. John_{*f*} begged Mary [PRO_{*f*} to be rented a house on the beach]

iii. John begged Mary_{*f*} [PRO_{*f*} to buy (herself/*himself) a house on the beach]

iv. John_{*f*} begged Mary [PRO_{*f*} to be sold a house on the beach]

In the sentences in (104)-(106), PRO obtains its θ -role from the complement of *allow*. The sentences are sometimes odd and benefit from being given a context, but control is as indicated. As expected by the V₁-V₂ theory, when the embedded sentence adjacent to V_m (which is V₁) contains a V₂ verb and the passive morpheme, as in the (i) examples,

control is different than when the embedded sentence adjacent to V_m does not contain a V_2 verb and a passive morpheme, as in the (ii)-(iv) examples. The Thematic Theory predicts, contrary to the facts, that control will be the same in each case.

- (104) i. North_{*f*} pleaded with Poindexter [PRO_{*f*} to be allowed e_f [PRO_{*f*} to go]]
- ii. North pleaded with Poindexter_{*f*} [PRO_{*f*} to seem [e_f to be allowed e_f [PRO_{*f*} to go]]]
- iii. North pleaded with Poindexter_{*f*} [PRO_{*f*} to be certain [e_f to be allowed e_f [PRO_{*f*} to go]]]
- iv. North pleaded with Poindexter_{*f*} [PRO_{*f*} to be expected [e_f to be allowed e_f [PRO_{*f*} to go]]]
- (105) i. North_{*f*} begged Poindexter [PRO_{*f*} to be allowed e_f [PRO_{*f*} to go]]
- ii. North begged Poindexter_{*f*} [PRO_{*f*} to seem [e_f to be allowed e_f [PRO_{*f*} to go]]]
- iii. North begged Poindexter_{*f*} [PRO_{*f*} to be certain [e_f to be allowed e_f [PRO_{*f*} to go]]]
- iv. North begged Poindexter_{*f*} [PRO_{*f*} to be expected [e_f to be allowed e_f [PRO_{*f*} to go]]]
- (106) i. North promised Poindexter_{*f*} [PRO_{*f*} to be allowed e_f [PRO_{*f*} to go]]]
- ii. North_{*f*} promised Poindexter [PRO_{*f*} to seem [e_f to be allowed e_f [PRO_{*f*} to go]]]
- iii. North_{*f*} promised Poindexter [PRO_{*f*} to be certain [e_f to be allowed e_f [PRO_{*f*} to go]]]
- iv. North_{*f*} promised Poindexter [PRO_{*f*} to be expected [e_f to be allowed e_f [PRO_{*f*} to go]]]

The above examples show that the properties of control in verbs like *promise*, *beg*, etc. depend on the presence or absence of *beten* in the lower sentence and have nothing to do with identity of θ -roles between PRO and any element in the matrix sentence.

9.7 THE V1-V2 CONTROL THEORY VERSUS THE C-COMMAND THEORY.

The c-command control theory claims that the control module (or those mechanisms which determine control) can obtain sufficient information to assign PRO its interpretation and reference from factorization (107).

(107) Factorization of a string according to the c-command theory.

[CP NP V_m ((P) NP) [CP PRO to ...]]

control of PRO = function (c-command and government relations of PRO to the NPs and V in the matrix sentence)

In the following sentences, the c-command theory would consider only the part of the phrase marker which dominates the italicized elements in these strings:

(108) i. [*John begged Mary_i* [*PRO_i* to invite Sue [PRO to sing]]]

ii. [*John_i* begged Mary [*PRO_i* to be invited *e_i* [PRO to sing]]]

(109) i. [*John_i* promised Mary [*PRO_i* to allow Sue [PRO to sing]]]

ii. [*John* promised *Mary_i* [*PRO_i* to be allowed *e_i* [PRO to sing]]]

The factorization of the c-command theory (107) indicates that the sentences in (108) both have the factorization (110), and that the sentences in (109) both have the factorization (111). How can information about c-command and government of PRO, considering only the structures in (110) and (111) correctly indicate control in (108)-(109)?

(110) [CP John begged Mary [PRO ...

(111) [CP John promised Mary [PRO ...

There are two basic failings to the c-command approach to control: First, there is not enough information contained in the factorization (107) upon which the theory attempts to determine control of PRO. Second, the type of information which is relevant for determining con-

trol of PRO has little, if anything, to do with government and c-command in cases like (108)-(109).

The V1-V2 theory makes the claim that control of PRO can only be determined by a mechanism which reads lexical information from two entries and accesses information about INFL of the embedded sentence. Control is determined by the mechanism which relates the information in the lexicon to information in the phrase marker in order to coindex and assign features to PRO.

10.0 CONCLUSION

I have examined a current model of grammar within Chomsky's Move- α framework in relation to its implications for parsing natural language. I have demonstrated that Chomsky's arguments for the movement of elements are not conclusive. From the computational point of view, the model can be greatly improved by eliminating movement operations entirely.

I have explored how a lexically based parser would operate in an analysis of infinitive constructions in English. The grammatical properties of infinitives can be accounted for by assuming that control and complement information are encoded lexically, and that these lexical structures are available to the parser.

The tree building mechanisms of the Parser as Licenser Model incorporate the following:

1. Access to lexical information.
2. Insertion of empty categories and assignment of the features λ anaphor, λ pronominal. These result in three kinds of empty categories in English: anaphors (NP-trace), variables (wh - trace, and empty operator) and PRO.
3. Binding of anaphors and variables.
4. X-bar theory with the X-bar adjunct mechanism.
5. θ -theory
6. Case theory

These tree building mechanisms are also driven by the Projection Principle and the Licensing Principles of Full Interpretation.

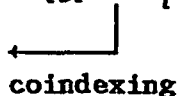
The phrase marker of control structures contains PRO_[+a,+p] in ungoverned positions, in particular, as the subject of infinitive constructions. The control model utilizes structural information in the phrase marker and lexical information about the elements in the string in order to (a) coindex PRO with its antecedent, if any exists, and (b) assign extrinsic features to PRO. The properties of PRO in a string are the features and reference assigned by the last mechanism which applies. Later mechanisms can overwrite or change features of earlier mechanisms. There are four basic PRO interpretation/reference mechanisms and they apply in order.

The input to the control module is a string containing PRO_[+a,+p] which has been inserted into the phrase marker by the tree building mechanisms. Binding Theory has applied to coindex the relevant reflexives, etc. in the tree input to the Control Module.

- (1) [CP NP V ((P) NP) [CP PRO_[+a,+p] to VP]]

The output of the control module is the phrase marker and string in which PRO (a) may be coindexed with some NP in the string and (b) PRO is assigned features [*±gen,±def*].

- (2) [CP NP V (P) NP [CP PRO_[+a,+p,αg,βd] to VP]]



coindexing

A basic assumption of this theory is that if the control module cannot assign the features [*α generic, β definite*] to PRO, then the string is uninterpretable.

The *wh*-PRO_{one} licensing mechanism operates on a string which has PRO with no features in a CP which has a *wh*-operator in COMP. The claim

is that a *wh*-operator in COMP licenses assigning the features [+gen,-def] to PRO. This mechanism cannot overwrite any features.

The V_1 - V_2 lexical interpretation mechanism first assigns indices to PRO and its referent, and then assigns features to PRO. This assignment of features after coindexing is a technical detail of my formulation more than a real process. If the V_1 - V_2 mechanism cannot coindex PRO with an appropriate referent, and the lexical entry indicates that such coindexing is required, then the V_1 - V_2 mechanism will not assign the indexing and not assign the features to PRO. The V_1 - V_2 lexical interpretation mechanism can overwrite the features assigned by the *wh*-PRO_{one} licensing mechanism.

There is some redundancy between the assignment of indices to indicate coreference of PRO and the definition of the feature [+def], i.e. a [+def] PRO must have a single node which is indicated as coreferential with PRO. If PRO does not have an antecedent in the sentence, then by definition it cannot be [+def]. The feature [+def] might be marked on PRO redundantly to indicate that PRO is coindexed. Or perhaps the feature might be dispensed with and the coindexing alone provides the interpretation, the feature [+def] being redundant. The issue seems to be simply a technical problem.

The long-distance control mechanism operates on a phrase marker which already has the features PRO[+gen,-def]. Binding Theory indicates that PRO[+gen,-def] is bound to a definite third person reflexive (*himself, herself, themselves*). The definite features of the reflexive, including number and gender, project onto the PRO changing the features to PRO[+gen,+def,*afeminine,βplural*]. The feature [+def] requires PRO

to have an NP antecedent in the sentence. Since PRO is [+anaphor,+pronoun], binding theory cannot locate the antecedent. The features [+anaphor,+pronoun] determine the reference: PRO, being [+anaphor], must have a least one NP in the sentence which is an antecedent. PRO, being [+pronoun], is satisfied by any antecedent that could be a possible referent for *he*, *she*, or *them*. This mechanism can overwrite features placed on PRO by the V₁-V₂ mechanism.

The adjunct mechanism operates to assign a referent to PRO in adjunct structures where the maximal projection is one of the major categories defined by ±V, ±N: V, NP, ADV, ADJ. PRO in adjuncts is PRO_{one} in constructions with an operator and PRO_{exclusive} in non-operator constructions. The mechanisms which determine the referent can utilize information from lexical entries, structural information in the phrase marker, adverb modifier (ADV-mod) links, and agreement-selection (Agr-SR) links.

A summary of the operations of the Control module follows:

1. *wh*-PRO_{one} Licensing Mechanism

If there is a structure like (a) in which PRO has no features, then the presence of a *wh*-element in COMP will license the assignment of the features [+gen,-def] to PRO, as in (b).

(a) [CP *wh*- [IP PRO to VP]] →

(b) [CP *wh*- [IP PRO[+gen,-def] to VP]]

2. Lexical Licensing Mechanism

The lexical entry of an item indicates the reference and assigns features to PRO in complements it licenses. In utilizing lexical information, the control module will first mark in the

phrase marker any coindexing of PRO and second will assign PRO in the phrase marker its feature combination. If it is impossible to satisfy the required lexical coindexing for an item, $word_n$, the features of $word_n$ will not be marked on PRO, and PRO is uninterpreted.

The contents of a lexical entry follows:

word: textual θ -role
 internal arguments: unordered complements
 Control (if applicable): verb class, antecedent of PRO,
 PRO features

3. V_1 - V_2 Lexical Interpretation Mechanism

IF: (a) if V_m is V_1 and V_e is V_2 , and
 (b) if $beten$ is present in IP_e ,

THEN: mark PRO for the reverse the lexical specification of
 control for V_m :

- (c) PRO is lexically coindexed with the external argument of V_m , it must be coindexed with the internal argument,
- (d) if PRO is lexically coindexed with the internal argument of V_m , it must be coindexed with the external argument.

OTHERWISE: In the phrase marker, mark the indices on PRO and the NP to satisfy the lexical markings on V_m which require PRO to be coindexed with the internal or external argument of V_m .

4. Long-Distance Control Mechanism

This mechanism can rewrite features. The basic principles for determining the antecedent of $PRO[+gen,+def]$ in the case of long-distance control are these:

- a. The fact that PRO is [$+anaphor,+pronoun$] means that coindexing PRO and its antecedent cannot be done by binding theory.

- b. Binding assigns the features [\pm feminine, \pm plural] to PRO[+gen,-def] when PRO binds a definite reflexive pronoun:

PRO[+gen,-def]_i ... himself_i → PRO[+gen,+def,-fem,-plu]
 PRO[+gen,-def]_i ... herself_i → PRO[+gen,+def,+fem,-plu]
 PRO[+gen,-def]_i ... themselves_i → PRO[+gen,+def,+plu]

- c. The feature [+def] means that the PRO must have an antecedent NP in the sentence.
 d. The feature [+gen] means that the PRO need not be identical with the antecedent, but only include the antecedent.
 e. The features [\pm feminine, \pm plural] indicate what is a possible NP antecedent.

With long-distance control configurations, the number feature need not match the number feature of the antecedent, as in *Mary thinks that it is unclear how to behave themselves*. Such is not the case when control is lexically assigned as in **Mary wonders how to behave themselves*.

5. The Adjunct Mechanisms

There is a possible adjunct infinitive clause for each of the major categories: V, ADV, NP, ADJ.

The VP-Adjunct Control Mechanism

WHEN [+V]_m licenses an external argument in its lexical entry,

IF CP_m contains an external argument (does not contain *be+en*),

THEN Assign the features [-gen,+def] to PRO and coindex PRO[-g,+d] with the external argument. (CASE I)

OTHERWISE Assign PRO_{one}. (CASE II)

OTHERWISE Assign no coindexing or features to PRO, which will result in the PRO being uninterpreted by the control module. (CASE III)

The ADV-Adjunct Indirect Control Mechanism:

- i. If the adverbial CP is in an AP, then coindex PRO with the NP pointed to by the Agr-SR link for that adjective.

[John is too [AP too stubborn [PRO to talk to Bill]]]
 └──Agr-SR──┘

- ii. If the adverbial CP is in an ADVP, and the head of the ADVP is linked to a verb via an ADV-modifier link, then coindex PRO with the NP pointed to by the Agr-SR link for that verb.

[John is stubborn [ADVP enough [PRO to talk to Bill]]]
 └Agr-SR┘ └──ADV-mod──┘

The NP-Adjunct Control Mechanism:

In an NP infinitive complement, assign the features [-g,+d] and coindex PRO with the head noun of the NP to which the adjunct is attached, i.e. to structure (i) assign the features and indices (ii).

- i. [CP ... [NP [N' [N' N] [CP PRO to VP]]] ...] →
 ii. [CP ... [NP [N' [N' N_i] [CP PRO[-g,+d]_i to VP]]]
 ...]

The ADJ-Adjunct Control Mechanism:

The NP that is the referent of the empty operator is the NP pointed to by the Agr-SR link of the adjective head of the AP which contains the adjunct.

APPENDIX A. TYPES OF ADJECTIVES1. *afraid* predicates+external θ -role, no operator:

- i. * it is afraid [PRO to talk to John]
- ii. John_f is afraid [PRO_f to talk]
- iii. * John_f is afraid [ei [PRO to talk to ei]]
- iv. John_f is afraid [PRO_f to talk to Mary]

able (oblig INF)
 afraid
 anxious
 ashamed
 careful
 contented
 destined (oblig INF)
 determined
 disappointed
 disinclined (oblig INF)
 disposed (oblig INF)
 due (oblig INF)
 eager
 eligible
 engaged (oblig INF)
 enraged
 enraptured
 enthralled
 entranced
 fated (oblig INF)
 frightened
 game
 glad
 happy
 hell-bent (oblig INF)
 horror-stricken
 impatient
 impotent
 inclined (oblig INF)
 incompetent
 indisposed
 keen
 licensed (oblig INF)
 loath (oblig INF)
 overjoyed
 poised
 powerless
 prepared
 prone (oblig INF)
 proud
 qualified

quick
 raring (oblig INF)
 relieved
 reluctant
 rumored (oblig INF)
 scared
 sorry
 swift
 thankful
 unable (oblig INF)
 welcome
 willing
 wont (oblig INF)

2. *pretty* predicates

non-agentive external θ -role, obligatory operator

Lasnik & Fiengo: *beautiful, cacophonous, delicious, fragrant, graceful, melodious, pretty, pungent, scratchy, slippery, tasty*

Interesting cases are those which take human subjects

attractive
 beautiful
 graceful
 handsome
 pretty
 winsome

3. *tough* predicates

-external θ -role, obligatory operator: missing object obligatory when a non-expletive is in subject position:

- i. it is easy [PRO_{arb} to talk to John]
- ii. * John is easy [PRO to talk]
- iii. John is easy [O_i [PRO_{arb} to talk to ei]]
- iv. * John is easy [PRO to talk to Mary]

amusing
 boring
 clever
 competent
 convenient
 difficult
 disappointing
 easy
 elating
 enraging
 enrapturing
 entertaining
 enthralling
 entrancing

frightening
 gratifying
 hard
 impossible
 infuriating
 intelligent
 interesting
 overjoying
 possible
 relieving
 simple
 stimulating
 sufficient
 tough
 unhealthy
 uninteresting

4. Raising constructions
 (S'-deletion: Chomsky 1982a:22)

-external θ -role, no EO.

about
 apt
 bound
 certain
 liable
 likely
 set
 supposed
 sure
 unlikely

5. Mixes

illegal
 ready
 crazy (insane or eager)
 curious (odd or interested in)

APPENDIX B. NOMINALIZATIONS

Ruzicka claims that he offers an analysis of "the verb *promise* in the lexicon in terms of a special kind of semantic relation: more precisely, in terms of *thematic relations*, I use the latter term in the sense of Jackendoff (1972) and, essentially, Fillmore's (1968) case roles ("deep cases")."(Ruzicka:311). Fillmore's system considers *Columbus* in (1), (2), (3) and (4) to have the same θ -role.

- (1) Columbus_{*f*} promised Queen Isabella [PRO_{*f*} to return with gold]
- (2) * Queen Isabella was promised by Columbus_{*f*} [PRO_{*f*} to return with gold]

Noun phrase complements are beyond our discussion, but they seem to obey the general principles of our theory. Consider the complements of the noun *promise*.

- (3) Columbus'_{*f*} promise to Queen Isabella [PRO_{*f*} to return with gold]
- (4) The promise of/by Columbus_{*f*} to Queen Isabella [PRO_{*f*} to return with gold]

Structures (1)-(4) seem to pose a problem for any Thematic Role Theory which attempts to account for the distribution and interpretation of PRO solely on the basis of agreement of θ -roles between PRO and an NP in the matrix phrase.

A possible avenue for research in our framework would be to assume that the lexical entry of a noun can include an unordered specification for at most one NP and one or more PP and CP. Since nouns cannot assign case, a mechanism permits *of* as a case-assigner for NP in the complement. Control information might be specified whether PRO is coindexed with

NP or PP. A noun has no external θ -role. A noun entry might be something like (5), or (6).

- (5) *noun*₁ internal arguments: NP, PP, CP
 control information: PRO is coindexed with N,
 PRO features
- (6) *noun*₂ internal arguments: NP, PP, CP
 control information: PRO is coindexed with PP,
 PRO features

Possible entries might be these:

- (7) *promise* internal arguments: NP, PP
 control information: PRO is coindexed with NP,
 PRO[-g,+d]
- (8) *advice* internal arguments: NP, PP
 control information: PRO is coindexed with PP,
 PRO[-g,+d]

These entries would permit the control module to assign these interpretations:

- (9) the promise of Columbus_{*f*} to Queen Isabella [PRO[-g,+d]_{*i*} to return with gold
- (10) the advice of Columbus to Queen Isabella_{*f*} [PRO[-g,+d]_{*i*} to spend it wisely]

In the N complements, as in the V complements, the basic mechanisms seem to be syntactic (i.e. structural, configurational, and lexical) and not semantic (i.e. specific theta roles). Notice that there is no agreement of θ -roles between PRO and the referent in the matrix sentence:

- (11) The promise of Columbus to Isabella to buy Spanish ships.
- (12) The promise of Columbus to Isabella to sell Spanish ships.
- (13) The advice of Columbus to Isabella to sell the ships.
- (14) The advice of Columbus to Isabella to buy the ships.

BIBLIOGRAPHY

- Abney, S. (1986). "Licensing and Parsing". Unpublished manuscript, MIT, Cambridge.
- Abney, S. and J. Cole (1986). "A Government-Binding Parser". Unpublished manuscript, MIT, Cambridge.
- Bach, E. (1982). "Purpose Clauses and Control", in P. Jacobson and G.K. Pullum (eds.), *The Nature of Syntactic Representation*, 35-57. Foris, Dordrecht.
- Berwick, R.C. and S. Fong (1990). "Principle-Based Parsing: Natural Language Processing for the 1990's" in P.H. Winston and S.A. Shellard (eds.) *Artificial Intelligence at MIT: Expanding Frontiers*. MIT Press, Cambridge.
- Chierchia, G. (1983). "Outline of a Semantic Theory of (Obligatory) Control", in M. Barlow et al (eds.), *Proceedings of the West Coast Conference on Formal Linguistics*, Vol. 2, Stanford Linguistic Association, 19-31.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.
- Chomsky, N. (1968). *Language and Mind, enlarged edition*. Harcourt Brace Jovanovich, New York.
- Chomsky, N. (1977a). *Essays on Form and Interpretation*. North-Holland, Amsterdam.
- Chomsky, N. (1977b). "On *wh*-Movement", in P.W. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax*. Academic Press, New York.
- Chomsky, N. (1980a). "On Binding." *Linguistic Inquiry*, 11.1:1-46.
- Chomsky, N. (1980b). *Rules and Representations*. Columbia University Press, New York.
- Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Foris, Dordrecht. (Second revised edition, 1982).
- Chomsky, N. (1982a). *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge.
- Chomsky, N. (1982b). *The Generative Enterprise: A discussion with Riny Huybregts and Henk van Riemsdijk*. Foris, Dordrecht.
- Chomsky, N. (1986a). *Knowledge of Language: Its Nature, Origin, and Use*. Praeger, New York.

- Chomsky, N. (1986b) "Changing Perspectives on Knowledge and Use of Language" in M. Brand and R.M. Harnish, eds., *The Representation of Knowledge and Belief*. University of Arizona Press, Tucson.
- Chomsky, N. (1986c) *Barriers*. MIT Press, Cambridge.
- Chomsky, N. (1987). *Language in a Psychological Setting*, Working Papers in Linguistics Number 22. Sophia University, Tokyo.
- Chomsky, N. (1988a). *Language and Problems of Knowledge: The Managua Lectures*. MIT Press, Cambridge.
- Chomsky, N. (1988b). "A Personal View", paper presented at conference session *Linguistics and Adjacent Fields: the State of the Art*, Israel.
- Chomsky, N. (1988c). "Prospects for the Study of Language and Mind", paper presented at conference session *Linguistics and Adjacent Fields: the State of the Art*, Israel.
- Chomsky, N. (1988d). "Some Notes on Economy of Derivation and Representation", paper presented at conference session *Linguistics and Adjacent Fields: the State of the Art*, Israel.
- Chomsky, N. and H. Lasnik (1977). "Filters and Control." *Linguistic Inquiry*, 8.3:425-504.
- Correa, N. (1988). *Syntactic Analysis of English with respect to Government-binding Grammar*. Doctoral dissertation, Syracuse University.
- Curme, G.O. (1931). *A grammar of the English Language in Three Volumes, Volume III. Syntax*. D.C. Heath, Boston.
- Epstein, S.D. (1989) "Quantification in Null Operator Constructions". *Linguistic Inquiry*, 20:647-658.
- Farkas, D.F. (1988) "On Obligatory Control". *Linguistics and Philosophy*, 11:27-58.
- Hornstein, N., and D. Lightfoot (1987) "Predication and PRO" *Language*, 63.1:23-52.
- Jespersen, O. (1909-1949). *A Modern English Grammar on Historical Principles*. George Allen and Unwin, London.
- Koster, J. (1978). "Conditions, Empty Nodes, and Markedness". *Linguistic Inquiry*, 9:551-593.
- Koster, J. (1984). "On Binding and Control". *Linguistic Inquiry*, 15:417-459.
- Koster, J. (1987). *Domains and Dynasties: The Radical Autonomy of Syntax*. Foris, Dordrecht.

- Lasnik, H. and R. Fiengo (1974). "Complement Object Deletion." *Linguistic Inquiry* 5.4:535-571.
- Lasnik, H. and J. Uriagereka (1988). *A Course in GB Syntax: Lectures on Binding and Empty Categories*. MIT Press, Cambridge.
- Lees, R.B. (1960). "A Multiply Ambiguous Adjectival Construction in English." *Language* 36.2:207-221.
- Manzini, M.R. (1983). "On Control and Control Theory." *Linguistic Inquiry* 14.3:421-46.
- Quirk, R. (1977). "A Tough Object to Trace." *Journal of Linguistics* 13:99-102.
- Ravin, Y. (1990). *Lexical Semantics without Thematic Roles*. Clarendon Press, Oxford.
- Riemsdijk, H.C. van, and E. Williams (1986). *Introduction to the Theory of Grammar*. MIT Press, Cambridge.
- Ruzicka, R. (1983). "Remarks on Control." *Linguistic Inquiry* 14.2:309-324.
- Wehrli, E. (1988). "Parsing with a GB-Grammar", in U. Reyle and C. Rohrer, eds., *Natural Language Parsing and Linguistic Theories* D. Reidel, Dordrecht.
- Williams, E. (1980). "Predication." *Linguistic Inquiry*, 11:203-238.