GRAMMATICAL CONFIGURATIONS AND GRAMMATICAL RELATIONS

by

YEHUDA NAHUM FALK

B.A., Brandeis University (1980)

SUBMITTED TO THE DEPARTMENT OF
LINGUISTICS AND PHILOSOPHY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS OF THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1984

© Yehuda Nahum Falk 1984

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

Signature of Author

Department of Linguistics and Philosophy

23 May 1984

Certified by

James Higginbotham
Thesis Supervisor

Accepted by

Samuel Jay Keyser
Chairman, Departmental Graduate Committee

JUN 27 1984
ABSTRACT

This dissertation presents and motivates a lexicalist syntactic framework called Configurelational Grammar, which, like Lexical-Functional Grammar, recognizes the primacy of both (surface) grammatical configurations and (surface) grammatical relations, but unlike Lexical-Functional Grammar represents both of these at a single level of syntactic representation.

The initial chapter sets out the basic goals of this dissertation and outlines the basic assumptions about syntax. Chapter 2, the heart of this study, presents evidence from constituent order, subcategorization, the interface between syntax and the other major components (phonology and semantics), and anaphora that there has to be a single level of representation in which both grammatical configurations and grammatical relations are encoded.

Chapter 3 presents an analysis of anaphora, including control and syntactic binding, within the formalism of Configurelational Grammar. It is also shown that anaphora gives no support to the claim that syntax must be able to refer to nonsurface grammatical relations. Chapter 4 makes the formalisms of the theory of grammatical relations (based on Lexical-Functional Grammar) and the theory of grammatical configurations (a version of the so-called X-Bar Convention) more explicit, in addition to discussing the organization of the lexicon. The final chapter addresses the question of whether or not an autonomous level of relational representation on the order of Lexical-Functional Grammar's level of functional structure is needed, and answers in the negative.

The Appendix proposes a set of categorial distinctive features which assumes a theoretical framework such as Configurelational Grammar. It is shown to be superior to other proposed systems, and a proposal is put forward for the evolution of the category "modal" in English.
ACKNOWLEDGMENTS

In a field as varied as generative syntax is today, having the opportunity to work with people who have very different outlooks is very important. I have been lucky to be able to have a thesis committee which represents this variety. In addition, each member of my committee has brought particular talents which have helped this thesis reach its conclusion. My advisor, Jim Higginbotham, has the unique ability to say what I want to say better than I can; this has proved to be very helpful. Luigi Rizzi's invaluable insights into the workings of language and Raj Ross's ability to come up with a counterexample to anything have kept me on my toes. Last, but certainly not least, Jane Grimshaw's ability to put up with me never ceases to amaze me. She supervised my undergraduate thesis at Brandeis, and still agreed to serve as an outside member on my committee now. Much of what I have done is due to her encouragement to follow through on my half-baked ideas.

I must acknowledge two people who, being in California this year, were unable to have any direct impact on this dissertation, but who have been instrumental in making a linguist out of me. These are Ray Jackendoff and Joan Bresnan. Ray was my first linguistics teacher, and it was his enthusiasm about the relationship between generative grammar and the mind that made me decide to major in linguistics as an undergraduate. It was from Ray that I learned the importance of explicit analyses and of considering all options. Joan reinforced these basic lessons, and her statement in class once that unless you have a theory of exceptions calling
something a "marked exception" is simply jargon for "counterexample" has stayed with me. The theoretical influence of both Ray and Joan should be transparent in this dissertation.

Other professors, both at MIT and at Brandeis, who have inspired and influenced me, either through class lectures, private discussions, or both, include Noam Chomsky, Ken Hale, Paul Kiparsky, and Joan Maling. Among fellow students, the first one to come to mind is Jeri Kisala, with whom I have had many fruitful discussions over the past four years. Other students who have been of particular help either with the thesis or in other ways include (but are not limited to) Mark Baker, Nigel Fabb, Sharon Flank, Lori Levin, K. P. Mohanan, Malka Rappaport, and Lisa Travis. A special thank-you is due to Jay Keyser and especially Morris Halle for helping me survive the recent faculty shuffle and put together a committee.

Finally, I would like to thank my family: my wife Brandel for putting up with me and occasionally serving as an informant, my older son Eli for reminding me of the wonder of language acquisition, and my younger son Ioni for being one of the two cutest babies I have ever seen.

Blessed be He who has kept us alive, sustained us, and enabled us to reach the momentous occasion of the completion of this dissertation.
This dissertation is dedicated to the memory of my grandfather Louis Klima.
TABLE OF CONTENTS

ABSTRACT .................................................................................................................. 2
ACKNOWLEDGMENTS ............................................................................................... 3
DEDICATION ................................................................................................................ 5
TABLE OF CONTENTS ............................................................................................... 6

CHAPTER ONE: INTRODUCTION ............................................................................. 10
  1.1. Syntactic Metalanguage .................................................................................. 10
  1.2 Background Assumptions ............................................................................... 14
Footnotes ................................................................................................................... 19

CHAPTER TWO: EVIDENCE FOR CONFIGURELATIONALITY ............................... 21
  2.1. Ordering Generalizations .............................................................................. 21
  2.2. Subcategorization .......................................................................................... 32
  2.3. The Syntax-Semantics Interface .................................................................... 48
  2.4. The Syntax-Phonology Interface .................................................................. 55
  2.5. Configurelationality in Anaphora ................................................................. 63
    2.5.1. Overview .................................................................................................. 63
    2.5.2. Anaphora in General .............................................................................. 65
    2.5.3. Reflexives and Reciprocals ...................................................................... 68
    2.5.4. C-command ............................................................................................. 72
    2.5.5. Island Constraints ................................................................................... 93
  2.6. Summary ........................................................................................................ 93
Footnotes ................................................................................................................... 101
CHAPTER THREE: THE THEORY OF ANAPHORA

3.1. Introduction ................................................. 105
3.2. Personal Pronouns .......................................... 107
3.3. R-anaphors .................................................. 112
3.4. Pseudo-reflexives ......................................... 118
3.5. Subjecthood of Antecedents ............................... 125
  3.5.1. Subjective Anaphors .................................. 125
  3.5.2. Nonsubjective Pronouns ............................... 127
  3.5.3. Nonsurface Subjects .................................. 130
3.6. Lexical Reflexivization .................................... 142
3.7. Raising and Control Phenomena .......................... 147
  3.7.1. Division of the Phenomena ............................ 147
  3.7.2. Raising .................................................. 150
  3.7.3. Obligatory Equi ....................................... 156
  3.7.4. The Anaphor PRO ..................................... 160
3.8. Syntactic Binding .......................................... 166
3.9. Island Constraints ......................................... 176
  3.9.1. Overview .............................................. 176
  3.9.2. The Role of the Focusing GRs ........................ 177
  3.9.3. Configurational Constraints: A First Approximation... 180
  3.9.4. Configurationality Revisited ........................ 193

Footnotes .................................................................. 197
CHAPTER FOUR: An Explicit Theory of Configurelational Grammar........201

4.1. Grammatical Relations........................................201
4.2. Grammatical Configurations..................................207
   4.2.1. Grammatical Categories..................................207
   4.2.2. Exocentricity..............................................210
   4.2.3. Syntactic Rules...........................................218
4.3. The Lexical Component.........................................221

Footnotes.............................................................223

CHAPTER FIVE: The Level of Functional Structure..................224

5.1. Structural Considerations......................................225
   5.1.1. Nonisomorphism of Constituent Structure and Functional Structure.............................................226
   5.1.2. Definitions and Features at Functional Structure....232
      5.1.2.1. Definitions............................................232
      5.1.2.2. Features.................................................236
   5.2. Discontinuous Expressions..................................244
   5.3. Sluicing........................................................247
   5.4. Conclusion....................................................251

Footnotes.............................................................253

APPENDIX: Towards an Analysis of Grammatical Categories...........254

A.1. Distinctive Features...........................................256
A.2. Past Approaches................................................258
   A.2.1. Chomsky and Stowell.....................................258
   A.2.2. Bresnan.....................................................260
   A.2.3. Jackendoff................................................261
A.3. A New Attempt

A.3.1. The General Framework

A.3.2. The Major Division Features

A.3.2.1. [predicate]

A.3.2.2. [grammaticalized]

A.3.3. The Argument List Features

A.3.3.1. \[\text{pure G(rammatical) R(elations)}\]

A.3.3.2. \[\text{tO(bligatory) A(rguments)}\]

A.3.3.3. Division of the Major Categories

A.3.4. Additional Features

A.3.4.1. [tbasic]

A.3.4.2. [tnominal]

A.3.5. Summary

A.4. Syntax

A.4.1. Modern English Syntax

A.4.1.1. Distribution of S and S Complements

A.4.1.2. Subjects

A.4.1.3. Degree Modifiers and Determiners

A.4.1.4. Adjectives and Adverbs

A.4.2. The Evolution of the Category "Modal"

A.5. Conclusion

Footnotes

BIBLIOGRAPHY

BIOGRAPHICAL NOTE
CHAPTER ONE: INTRODUCTION

1.1. Syntactic Metalanguage

In traditional grammatical description, syntax is dealt with using two kinds of vocabulary: grammatical categories and configurations, and grammatical relations. The problem of relating these two levels of description has been the topic of much debate among generativists over the past decade.

We can isolate five possible views on the relationship between grammatical configurations and grammatical relations.

(1) a. Grammatical configurations are basic. Grammatical relations (if they have any theoretical significance at all) are derived from grammatical configurations.¹

b. Grammatical relations are basic. Grammatical configurations (if they have any theoretical significance at all) are derived from grammatical relations.

c. Grammatical configurations and grammatical relations are both basic, and are represented at distinct levels of representation.

d. Grammatical configurations and grammatical relations are both basic, and are represented at the same level(s) of representation.

e. Neither grammatical configurations nor grammatical relations are basic.

(Here and for the remainder of this dissertation, we use the term "configurations" to include categories.) Position (1a) is the classical position of transformational grammar, as represented by Chomsky (1965). It persists in the system of "GFs" in the Government/Binding (G3) theory of Chomsky (1981, 1982) and in the Generalized Phrase Structure Grammar
(GPSG) framework of Gazdar (1981, 1982). The position described in (le), while it does not express the view of any complete theory, does describe the view of Generative Semantics with respect to underlying structure. Until the advent of Relational Grammar (RG) in the mid-1970s, these were the only positions in (1) which were realized in generative theory. What relational grammarians have argued (Johnson and Postal 1980; Perlmutter, ed., 1983; Perlmutter and Rosen, eds., in press) is that (lb) is the correct position (although, as Marantz 1981 suggests, this is not true of the initial stratum in some versions of RG, such as Johnson and Postal's "L-graph," which is actually described by (le)).

Putting (le), which as a complete view of Universal Grammar (UG) amounts to denying the importance of syntax, aside, (la,b) each propose one type of universal primitive in syntax. (lc,d), on the other hand, represent a more complex view of UG by specifying two different kinds of syntactic primitives. For this reason (la,b) may seem preferable as theories of language. In this dissertation, however, we will assume the necessity of both types of primitives, as recognized in recent work. (lc) has been realized in two very different ways: Lexical-Functional Grammar (LFG; Bresnan, ed., 1982; Levin, Rappaport, and Zaenen, eds., 1983) and the GB-based theory of Marantz (1981). This dissertation will argue for (ld). A theory that conforms to (ld) will be termed a configurational theory, and the specific one to be outlined here will be called Configurelational Grammar (CRG). To my knowledge, there is only one configurational theory that has been previously proposed: GB's Case theory. If, as argued by Marantz (1981, 1982a), the abstract Case system of GB can be interpreted as a representation of surface grammatical relations, then S-structure after the assignment of Case is
a level of configurational representation. However, due to the overall structure of the theory, it is unclear if it is possible to predict any interesting interaction between grammatical configurations and abstract Case.  

An explanation is in order as to why we will pursue (ld) in this dissertation if, as suggested above, it is a priori less preferable because it proposes two types of primitives. For this explanation, we turn to Chomsky (1972, 69), who addressed the same question with respect to the notion of a syntactic level of deep structure.

I will try to show that the status of deep structure, though conceptually somewhat on a par with that of autonomous phonemics (as has been observed, correctly, by McCawley, Postal, and others), differs in a fundamental way in that, in the case of deep structure, the questions [relating to the existence of the level, based on the kinds of generalizations that are possible with it and without it] receive positive answers, whereas in the case of autonomous phonemics, they do not. My point at the moment is that the question, in both cases, is an empirical one. There is almost nothing to say of any significance on a "methodological" level. One can point out, a priori, that further elaboration of linguistic theory, in both cases, complicates linguistic theory. Although one wants the "simplest" linguistic theory, [footnote omitted,] in some sense that cannot be specified in a useful way, elaborations and complications of linguistic theory are all to the good insofar as they narrow the choice of grammars and the range of admissible languages.

Thus, we can take the empirical evidence of the need for both kinds of primitives as overriding notions of a priori preferability of the form of linguistic theory. The advances that have been made over the years in various forms of transformational grammar, and in particular in research on the so-called X-Bar Convention, show that the vocabulary of grammatical configurations and categories is needed. The new insights provided by RG and LFG show that primitive grammatical relations are needed as well.

The general approach to be taken here is based on the theory of Lexical-Functional Grammar. As mentioned above, LFG conforms to the
view of (lc). It postulates two levels of representation: a representation of surface grammatical configurations, called *c(ons constituent) -structure*, and a representation of surface grammatical relations, called *f(unctional)-structure*. There are no abstract levels of c-structure or f-structure; in fact, they are excluded in principle from the theory by what Bresnan (1982b) and Kaplan and Bresnan (1982) call the principle of "direct syntactic encoding," which states that syntactic rules may not change structure. We agree with LFG about the lack of non-surface levels of representation; where we disagree is the separation of c-structure and f-structure as distinct levels of representation. In CRG, we recognize a single level of syntactic representation, with no syntactic derivation or strata, corresponding roughly to the "annotated c-structures" of Kaplan and Bresnan (1982).

The empirical content of our claim about syntactic representation is that we predict that there are configurational phenomena in language, i.e. phenomena whose rules involve simultaneous reference to grammatical configurations and grammatical relations. According to LFG, these should only exist in the case of island constraints, which are handled as conditions on the instantiation of long distance domination metavariables in the mapping between c-structure and f-structure (Kaplan and Bresnan 1982), as well as constituent order if one assumes phrase structure rules of the standard sort, as has been done in most of the LFG literature. If one assumes a rule system such as that proposed by Falk (1983a), as suggested in passing by Bresnan (1982a), even constituent order cannot be included, because ordering rules would be stated at the level of c-structure. Any other configurational phenomena would require either global constraints⁴ or a weakening of LFG.
In the next section, we will outline our basic representational assumptions. Chapter 2 presents evidence that there are configurational phenomena by examining constituent order, subcategorization, anaphora, and the interface between syntax and the other major components. Chapter 3 analyzes anaphora from the point of view of CRG, supporting the evidence from anaphora for configurationality given in Chapter 2, and adding to it. The remaining chapters deal with other formal issues: making the representational system more explicit in Chapter 4, examining the status of LFG's level of functional structure in Chapter 5, and proposing a set of categorial distinctive features in the Appendix.

1.2. Background Assumptions

We will assume that there is a single level of syntactic representation at which all the relevant configurational, relational, and lexical information is expressed. Here we will discuss the basic properties of such representations; a more complete and explicit account is deferred until Chapter 4.

Configurational properties are represented in fairly conventional form throughout Chapters 2 and 3 and will be formalized in an explicit version of the so-called X-Bar Convention in §4.2. The only innovation is that the projections (in the sense of the X-Bar Convention) of lexical items will be vertically above them. We treat $\tilde{S}$ as a projection of the complementizer (and, by treating to as a complementizer, we treat what is sometimes called $\overline{VP}$ as a case of $\tilde{S}$), $S$ as a projection of the AUX if there is one and the finite verb otherwise. These identifications of $S$ and $\tilde{S}$, which are similar in spirit to the identification of $\tilde{S}$ as the maximal projection of COMP and $S$ as the maximal projection of INFL
in GB (Stowell 1981), will be discussed further in §4.2.

Every nonhead in the syntactic representation of a sentence will have a relational label, representing the surface grammatical relation that the constituent bears. The set of grammatical relations to be assumed here is roughly that of LFG (Bresnan 1982a); in §4.1 we will discuss some aspects of our choice of grammatical relations and compare it both with LFG and RG. We call the grammatical relations specified in lexical entries (the "subcategorizable" ones) Lexical Relations, and the others Syntactic Relations. "Arguments" which serve as semantic heads ("objects" of oblique PPs, the verb following supportive do) will be marked with the relational label HEAD and be referred to as functional heads.

By "lexical information" we mean what is usually included under the term "subcategorization." In order to understand it, it is necessary to understand the rudiments of the lexical component, which operates roughly as in LFG.

One of the major tasks of the syntactic component of the lexicon is to provide a mapping between thematic relations and grammatical relations. In LFG, the representation of syntactically-relevant semantic argument structure is called the predicate-argument structure and the list of grammatical relations borne by the arguments is called the grammatical function assignment; in CRG we call these the semantic argument list and the syntactic argument list respectively. The nature of the mapping between them is discussed by Bresnan (1980, 1982a). In a notation only slightly different from that used in the latter article, the following examples drawn from both articles illustrate the mapping. The lexical form is an abbreviated notation used in LFG but not CRG for the links between the two argument lists.
(2)  

a. sample sentence: Sara inquired who was coming.  

semantic argument list: <Agent/Source, Propositional Theme>  
syntactic argument list: {SUBJ, COMP}  
lexical form: 'inquire<(SUBJ)(COMP)>' 

b. sample sentence: Max wrote to Sara.  

semantic argument list: <Agent/Source, Theme, Goal>  
syntactic argument list: {SUBJ, OBL}  
lexical form: 'write<(SUBJ) Ø (OBL,Goal)>' 

c. sample sentence: Max believes hamster food to taste bad.  

semantic argument list: <Location, Propositional Theme>  
syntactic argument list: {SUBJ, OBJ, XCOMP}  
lexical form: 'believe<(SUBJ)(XCOMP)>(OBJ)'

d. sample sentence: The hamster seemed upset to me.  

semantic argument list: <Goal, Propositional Theme>  
syntactic argument list: {SUBJ, OBL, XCOMP}  
lexical form: 'seem<(OBL,Goal)(XCOMP)>(SUBJ)'

Much of the linkage is regular and therefore defined by lexical rules. For instance, the linking of an Agent to SUBJ and the Theme of a transitive verb to OBJ is a characteristic of "accusative" languages, while true "ergative" languages link Agent to OBJ and Theme to SUBJ. (On the distinction between true ergative (or syntactically ergative) languages and languages with ergative case marking (or morphologically ergative languages), see Marantz 1981 and B. Levin 1983.) (2b) is the result of the application of the lexical rule of Intransitivization, stated by Bresnan (1980) as (3).

(3) OBJ $\rightarrow$ Ø

(2c,d) illustrate nonthematic SUBJs and OBJs. Only the Pure GRs (SUBJ,
OBJ, OBJ2) can remain unlinked to thematic arguments. These syntactic argument lists are also due to lexical rules — corresponding to the Raising transformation(s) of transformational grammar. We will not state the rules here. Other relationally-based lexical rules include passivization, middle formation, etc.

(4) a. Passivization (based on Bresnan 1982b)\(^7\)

Relational change: \(\text{OBJ} \leftrightarrow \text{SUBJ}\)
\(\text{SUBJ} \leftrightarrow \text{OBL}_{by}/\emptyset\)

Morphological: past participle

b. Middle Formation (Bresnan 1980)

\(\text{SUBJ} \leftrightarrow \emptyset\)
\(\text{OBJ} \leftrightarrow \text{SUBJ}\)

The relation-changing rules are LFG's (and CRG's) equivalent of NP-movement rules in transformational theory. As part of the lexical (word-formation) component, the output of lexical rules can enter further word-formation processes, as discussed by Bresnan (1978, 1982b) with respect to passivization. This will not be pursued further here, but it is an important piece of evidence for treating rules like passivization as lexical. To illustrate the operation of relation-changing rules, the passive version of (2c) is (5).

(5) sample sentence: Hamster food is believed (by Max) to taste bad.

semantic argument list: \(<\text{Location, Propositional Theme}>\)

syntactic argument list: \{OBL\(_{by}/\emptyset\), SUBJ, 'XCOMP'\}

lexical form: 'believe<(OBL\(_{by}/\emptyset\),(XCOMP))(SUBJ)'

The relationship between active and passive forms is captured by this lexical account just as well as traditional and current transformational accounts. For discussion, see Bresnan (1982b).
The lexical information that is represented in the syntax is the syntactic argument list. The configurational syntactic representation of sentence (6a), for example, is (6b).

(6) a. Miriam gave hamsters to Max.

b. 

```
S
  |<SUBJ>  VP
  NP Miriam
     | V gave {SUBJ, OBJ, OBL} hamsters
        |<OBJ> PP
        NP to Max
        <OBL> P
```
Footnotes for Chapter 1

1 Related to this is the view that grammatical relations derive from compositional semantics (Dowty 1982), a view that is taken in an informal way in GB. For discussion, see Marantz (1981) and Bresnan (1982a).

2 As pointed out by Marantz, the term "primitive" is somewhat misleading. In his theory, grammatical relations are represented structurally as in transformational grammar (i.e. SUBJ is the NP immediately dominated by S and OBJ is the NP immediately dominated by VP), but this is at an autonomous level of representation, one not directly related to the surface arrangement of morphemes. Thus, it is unclear whether or not one can say that Marantz views grammatical relations as primitives, but he does see an autonomous level of representation at which all that is expressed is surface grammatical relations. It is for this reason that the discussion here uses the term "basic" instead of "primitive."

3 It is not clear in GB to what extent abstract Case is to be identified with morphological case. If there is not a one-to-one correspondence between them, then abstract Case, a primitive of Government/Binding theory, does not meet Chomsky's (1981) criterion of epistemological priority, and, in fact, it is unclear how abstract Case differs from surface grammatical relations.

4 The arguments against global constraints should be familiar from the Extended Standard Theory (EST) polemics of the early 1970s. It is interesting that this issue is showing up again in a totally different context, and it serves to illustrate the fact that there was a misconception in some of the literature which argued in favor of global rules. The misconception is that one can speak of global phenomena independently of the representational system. There is no such thing as a global phenomenon, only a phenomenon which the existing representational system cannot handle without a global rule. What people on the EST side of the argument were saying was that a more constrained theory can be obtained if global rules are disallowed; the alternative was to change the representations. Thus, it is misleading when Bach (1977, 135) says, concerning Trace Theory:

   One of the hotly debated issues of the sometimes more calorific than phosphorescent "interpretive-generative-semantics" battles of recent memory was the question of globality. G. Lakoff and others ... claimed that the translation relationship necessarily involved nonlocal constraints (derivational constraints). Despite Chomsky's disclaimer ..., it seems to me that this claim is recognized as correct and explicitly incorporated into the "revised extended standard theory" (REST).

In LFG and CRG, which recognize the essential correctness of trace theory, there is no sense in calling traces global devices. In REST and GB they can be thought of as encoding globality, but need not be, especially in GB as it is presented in Chomsky (1981), where D-structure is seen as an abstraction from S-structure and therefore conceptually dependent on it.
Our notation differs from that of Bresnan (1982a) in the following respects:

(i) In the semantic argument list, we have omitted the name of the predicate and labelled the arguments with the thematic relations they bear. These thematic relation names are a convenient description based on the semantic representation, which we assume is in the code that the mind uses to represent concepts. On thematic relations, see Gruber (1965) and Jackendoff (1972); on thematic relations as they relate to cognition, see Jackendoff (1976, 1983).

(ii) In the syntactic argument lists we have omitted the parentheses around the names of grammatical relations.

(iii) In the syntactic argument lists, we have not put the non-thematic arguments last.

(iv) We have drawn in the links between the semantic and syntactic arguments.

Hale (1983) proposes that in "Lexical Structure," a representation akin to the syntactic argument list, the SUBJ is represented as an "external argument." In our notation, we could duplicate this as (i).

(i) \{SUBJ, [OBJ, OBL]\}

Hale's primary reason is that the reflexive in the Australian language Warlpiri, which is an AUX clitic, can never represent a SUBJ coreferential with an OBJ, only an OBJ coreferential with a SUBJ, as in other languages. By proposing a "lexical" VP, the c-command relation can be invoked even in extreme nonconfigurational languages. However, see fn 4 of Chapter 4.

We leave it as an open question whether both relational changes need to be stated in this rule. It is possible that one will follow as a consequence of the other. For example, Baker (1983a) argues that passivization only demotes the SUBJ, and that an independent rule (analogous to RG's Unaccusative Advancement) promotes the OBJ as a result. Alternatively, it has been argued in the RG literature (Perlmutter and Postal 1977, in press) that passivization is essentially the advancement of OBJ to SUBJ, and that the demotion of the original SUBJ is a by-product of this promotion. A more fully articulated theory of the lexicon will have to address this question. (I currently lean towards Baker's view.)

Note also that the specification of the relation that the former SUBJ bears as "OBL_{by}" is nothing more than a makeshift notation; there is actually no such relation as distinct from any other OBL. Bresnan (1982b) seems to suggest that by is simply the preposition that marks Agents, but as Marantz (1981) points out, passive NPs are not always Agents. I suspect that this rule has to specify what we call in §5.1.2.2 a noncategorical feature.
CHAPTER TWO: EVIDENCE FOR CONFIGURELATIONALITY

In this chapter, we demonstrate that there are configurelational phenomena. We examine every type of rule that has access to the syntactic representation of a sentence.

2.1. Ordering Generalizations

Traditional phrase structure rules (and the annotated phrase structure rules of Lexical-Functional Grammar) define two kinds of language-specific constraints on well-formed syntactic structures: dominance constraints and ordering constraints. Falk (1983a) proposes that these two roles of phrase structure rules should be separated into separate specifications of dominance (c(onsituency)-rules) and precedence (o(dering)-rules). We will discuss these two rule types in more detail in §4.2.3. In this section, we will argue that o-rules are configurelational. The basic notation for o-rules is (1).

\[
\begin{align*}
(1) & \quad a. \ A >> B \ (''A \ precedes \ B'') \\
& \quad b. \ A \ initial \\
& \quad c. \ A \ final \\
& \quad d. \ A \ second
\end{align*}
\]

We will assume in this discussion that an o-rule may specify the dominating category as a marked option. We adopt this assumption for the sake of argument; it will transpire that none of the rules that actually make use of this option (all of them are rules which exclude relational specifications) are adequate, and a theory which excludes this possibility can make stronger predictions about the nature of ordering generalizations.
For discussion, see §4.2.3. Note also that we do not recognize "discontinuous structures" (J. McCawley 1982) or "virtual structures" (Zubizaretta and Vergnaud 1982), and that o-rules cannot supersede the dominance relations specified by the c-rules. What these facts mean is that o-rules can specify ordering and positioning only with respect to sister nodes. Thus, for example, (1b) means only that A is initial within the immediately dominating category. It cannot mean initial in the clause or the entire structure because structure-mutilating rules are not permitted in this framework. We also allow what Falk (1983a) calls preferred o-rules. These are o-rules which, unlike ordinary o-rules, can be violated for extrasyntactic reasons. The clearest case is Heavy NP Shift, which is local reordering to facilitate processing; the rules which place the shiftable NP before items it shifts over are preferred o-rules which are violated when the Shift "applies."

Let us consider the configurationality of o-rules. Clearly, we would have a stronger theory if we allowed o-rules to refer only to grammatical relations or only to grammatical categories. However, this seems impossible, and we must allow o-rules to be configurational. Relational grammarians have on occasion (see, e.g., Perlmutter 1983) attempted to show that grammatical relations are sufficient for describing word order, but they generally only discuss rather superficial aspects of ordering, such as the relative positions of subjects, objects, verbs, obliques, and chômeurs, and often only in vague terms. For instance, J. R. Ross (class lectures, 1983, MIT) has pointed out that relational grammarians cannot explain why 2-chômeurs cannot follow passive chômeurs. What about the other side: categories only? Until recently, transforma-
tional grammar has not considered generalizations about ordering. In Government/Binding theory, ordering is held to be largely due to parameters of universal principles, with perhaps some arbitrary language-specific stipulations. The major effort within GB to study the question of ordering is Stowell (1981). In part, Stowell resorts to Case theory, the GB equivalent of surface grammatical relations, to get the ordering generalizations. For instance, the fact that (abstracting away from Heavy NP Shift) OBJs immediately follow the verb in English is due to a restriction on Case assignment that requires the assigner and assignee to be adjacent. Thus, the GB account of ordering is configurational. Aside from fully specified phrase structure rules, no purely configurational account of ordering has ever been proposed.

Consider first $\tilde{S}$s which are daughters of the S node. These are of two types: subjects (2a) and nonrestrictive (appositive) relative clauses (2b), according to Jackendoff (1977).

(2) a. That the earth is round bothers Max.
   b. Miriam is a home-birth midwife, which bothers obstetricians who are afraid of competition.

Appositive relatives are constituent final and $\tilde{S}$ SUBJs are constituent initial. Stating the o-rules solely on the basis of category is inadequate.¹

(3) a. $\tilde{S}$ final (preferred)
    b. $\tilde{S}$ initial / $S$

There are two ways to interpret (3), neither of which yields the correct results. The most obvious is to simply interpret the two rules as contradictory with respect to daughters of S. Under this interpretation, no $\tilde{S}$
could ever appear as a daughter of S because no matter where it appears in the S it would violate at least one of the rules in (3). The other interpretation uses the Elsewhere Condition on rule application, originally proposed by Kiparsky (1973), and stated by Kiparsky (1982) as (4).

(4) **Elsewhere Condition**

Rules A,B in the same component apply disjunctively to a form $\phi$ if and only if

(i) The structural description of A (the special rule) properly includes the structural description of B (the general rule).

(ii) The result of applying A to $\phi$ is distinct from the result of applying B to $\phi$.

In that case, A is applied first, and if it takes effect, then B is not applied.

Here (3b) is a more specific form of (3a) in terms of the structural description of the rule, so if the Elsewhere Condition applies, then all $\bar{S}$s under S will be initial while $\bar{S}$s under all other nodes will be final. This is also incorrect. Thus, the relevant o-rules cannot be stated in terms of grammatical categories alone. What about grammatical relations? Assuming, for the purposes of this discussion, that nonrestrictive relative clauses bear the relation NRR, the rules would be (5).

(5)  

a. NRR final  
b. SUBJ initial

(5a) is correct, although it misses the generalization that, aside from SUBJs, all $\bar{S}$s are constituent-final, regardless of dominating node or grammatical relation. (5b), on the other hand, is simply incorrect. For instance, sentential adverbs can freely precede NP SUBJs; however, they are odd before $\bar{S}$ SUBJs.²
(6)  
a. Apparently, the fact that the earth is round bothers Miriam.

b. ??Apparently, that the earth is round bothers Miriam.

Secondly, consider (7).

(7)  
a. Miriam, the fact that the earth is round seemed to bother ___.

b. *Miriam, that the earth is round seemed to bother ___.

The generally accepted constituent structure for (7b) is (8).

(8)  

With a structure like (8) there is no way to rule out (7b) with local constraints like o-rules. However, there is some evidence that the correct structure for (7b) is actually (9).

(9)  

Consider the following.

(10)  
a. Max knows who I like ___ and you hate ___.

b. *Max thinks that bagels I like ___ and you hate ___.

That is, sequences following wh phrases can be coordinated, but ones
following topicalized phrases cannot. Also note the following Right Node Raising data.

(11)  
   a. Max knows who, and Miriam knows what, I like __.
   b. Max knows that bagels, and Miriam thinks that lox, I like __.

The impossibility of (11b) as compared with (11a) indicates that the NP-VP sequence following a topicalized phrase does not form a separate constituent, while the one following a wh phrase does. It is interesting to note that Maling and Zaenen (1981) argue that in Icelandic wh phrases are daughters of S but topicalized phrases are daughters of S on the basis of the "verb second" constraint. In embedded clauses with a topicalized phrase the finite verb immediately follows this phrase (i.e. a topicalized phrase counts as the first constituent for the "verb second" rule), but if there is a wh phrase, the SUBJ precedes the verb (i.e. the wh phrase does not count as the first constituent). Given structure (9), the contrast in (7) is essentially the same as the one in (6). A further contrast is that NP SUBJs, but not sentential SUBJs, can follow the first auxiliary in direct questions. In transformational theory, this is accomplished by a transformation called Subject-Auxiliary Inversion, but in the analysis of Falk (1983a, 1983b, to appear) this construction is analyzed by allowing free ordering of the SUBJ and AUX, with a special interpretation rule applying to the "inverted" form. Thus, there are various items which can precede a nominal SUBJ but not a sentential one. The conclusion is that the o-rules, or at least the o-rule for sentential SUBJs, must be configurational. In order to capture the fact that all ſs except for SUBJs are final, we will assume the Elsewhere Condition and state the o-rules as (12).
(12)  
   a.  \( \tilde{S} \) final (preferred)  
   b.  \( \tilde{S} \)\textsubscript{SUBJ} initial  
   c.  \text{SUBJ} \gg \text{VP}  

(12a) is crucially not configurational, because it applies to every \( \tilde{S} \) not covered by another rule. The only other rule for \( \tilde{S} \) is (12b), and it also refers to the grammatical relation SUBJ. (12c) is configurational in a different way: one term is a grammatical relation and the other term is a category. \( \tilde{S} \) SUBJs actually satisfy both (12b) and (12c), as pointed out to me by Jane Grimshaw (personal communication). We note in passing that these facts about \( \tilde{S} \) SUBJs have led Koster (1978) to propose that they are not really SUBJs at all, but rather topicalized clauses binding an empty category in subject position. This approach makes the prediction that these clauses should behave like topicalized phrases. One well-known fact about topicalization is that it rarely applies in embedded clauses. It can under certain complex pragmatic conditions (Green 1976) but in most embedded contexts topicalization is decidedly bad. Consider the following.

(13)  
   a.  *I guess that the bagels Max won't eat.  
   b.  I guess that for the earth to be round would bother Max.  

Here, \( \tilde{S} \) SUBJs do not behave like topicalized phrases, contrary to the expectations raised by Koster's analysis.

There is a second, more striking case of configurationality in o-rules. We begin by looking at V-NP-NP structures. In the sentence

\[
I \text{ gave Max a hamster} \text{ the first of the two NPs (Max) is, in LFG and CRG terms, the OBJ, and the second (a hamster) is OBJ2.}^5
\]

Thus, we need an o-rule of the form (14).

(14)  \( \text{OBJ} \gg \text{OBJ2} \)
Furthermore, in *The nation elected a former actor president*, the first NP (*a former actor*) is OBJ and the second NP (*president*) is an XCOMP (predicative complement). Thus, we also need a rule (15).

(15) OBJ $\gg$ XCOMP

Thus far, the discussion has been dependent on the correctness of the proposal to split phrase structure rules into c-rules and o-rules. However, the ramifications of the ordering of double NPs transcend that particular proposal. Consider (16) and (17).

(16) a. *I gave my good friend who I introduced to you yesterday a hamster.

b. *I gave a hamster my good friend who I introduced to you yesterday.

(17) a. *The nation elected a former actor who had decided to go into politics president.

b. The nation elected president a former actor who had decided to go into politics.

The b. examples illustrate Heavy NP Shift over another NP. The Shift is grammatical when the second NP is a predicative complement, as in (17), but not in double object constructions. In the framework of Falk (1983a) this means that (14) is an ordinary o-rule but (15) is a preferred o-rule which can be violated for extrasyntactic reasons (such as what is probably a processing constraint that prefers long phrases finally).

An alternative analysis of Heavy NP Shift must likewise be capable of distinguishing between (16) and (17), and, we would like to suggest, such a distinction must involve explicit reference to the grammatical relations borne by the relevant NPs. One could, of course, assume different structures for the two constructions, such as Chomsky's (1981) "small clause" analysis of predicative complements, but such a move must be supported by (or at least be consistent with) the empirical evidence. In this case,
the evidence seems to favor the more traditional view of constituency: Bresnan (1982a) shows that traditional tests for constituency, such as Right Node Raising, show that the two NPs do not form a small clause constituent. If traditional tests are to be rejected, then new tests for constituency must be proposed in their place to keep constituent structure claims from being vacuous and untestable. Furthermore, the small clause analysis is motivated by universal principles, yet there is some very convincing evidence from other languages that "Raising to Object" constructions, including the double NP construction in question, must be allowed (e.g. Cole and Hermon 1981 on Imbabura Quechua and Mohanan 1982c on Malayalam). It thus seems that whatever the analysis of Heavy NP Shift, it will have to make reference to grammatical relations.

(14) and (15) are both relational rules; no categorial rule could order two NPs. There is another rule applying to XCOMP's, however, that cannot be stated purely relationally or categorically. Consider the facts exemplified by (18) and (19), due to Bresnan (class lectures, 1980, MIT).

(18)  
  a. Max seems foolish to me.
  b. *Max seems to me foolish.

(19)  
  a. *Max seems to be foolish to me.6
  b. Max seems to me to be foolish.

There is no way to derive these data from any universal principles in Government/Binding theory. Consider the representations of (18) and (19) at S-structure.
"Government" and Case-marking relations are exactly the same in both cases. The true generalization seems to be that to infinitives, which we are treating as Ss, are subject to (12a) (S final), while other XCOMPs are subject to a separate rule. What is this rule? We give a possible relational statement in (21a). A categorial statement requires a distinctive feature specification that collapses all non-S XCOMP categories (NP,
VP, AP, PP); in the system proposed in the Appendix they are all [+basic].

The categorial hypothesis is (21b).

(21)  
   a. XCOMP >> OBL
   b. [+basic]'' >> PP

Neither of these rules is adequate. (21a), not being in an Elsewhere Condition relation with (12a), creates conflicting requirements on to infinitives which are XCOMPs. As to (21b), note that PPs are [+basic] in our feature system. The rule predicts that there will not be an obligatory relative ordering between XCOMP PPs and OBL PPs. This is incorrect.

(22)  
   a. Max seems [XCOMP out of his mind] [OBL to me].
   b. *Max seems [OBL to me] [XCOMP out of his mind].

Thus, neither rule in (21) is correct. Once again, we need a configurational rule.

(23)  
   [+basic]'' XCOMP >> OBL

It should be noted that most of the rules we have motivated in this section are configurational. That is, we have not just shown that ordering is sensitive to both grammatical configurations (or categories) and grammatical relations, but that it is sensitive to both simultaneously. Thus adding a notion of f-[structure]-precedence to LFG, as suggested by Bresnan (class lectures, 1981, MIT, and unpublished research), would not suffice in any of the cases except for the ordering of double NPs. Any theory in which grammatical configurations and grammatical relations are represented at distinct levels would require global power to account for all the other ordering restrictions discussed here.
2.2. Subcategorization

The term "subcategorization" refers to the lexical specification of the syntactic context in which a given lexical item may appear. In the Standard Theory, this was realized as a specification of required sister categories. That is to say, subcategorization has traditionally been viewed as a configurational property. This view is maintained in Generalized Phrase Structure Grammar, although it is formalized differently. It has on occasion been claimed that subcategorization is a consequence of semantic properties (this seems to be the position of Chomsky 1981), but Grimshaw (1979) shows that complement selection occurs autonomously at both the semantic and the syntactic levels. In CRG, the specification of subcategorization is the syntactic argument list, discussed in §1.2; like the "lexical form" in LFG, it specifies the relational context that a lexical item must appear in, not the configurational/categorial context. In this section, we will first review the evidence for relational subcategorization and then discuss the issue of configurationality.

The first argument for relational subcategorization comes from Grimshaw (1982a,b). In the Romance languages (Grimshaw deals specifically with French), OBJs and indirect objects may be realized as full NPs and PPs respectively, or as preverbal pronominal clitics.

    sees 'Jean sees Pierre.'

b. Jean le voit.
    him sees 'Jean sees him.'

    talks to 'Jean talks to Pierre.'

b. Jean lui parle.
    him talks 'Jean talks to him.'
In the traditional transformational view, \( le \) in (24b) is an NP in deep structure while \( lui \) in (25b) is a PP. What category they are in surface structure is not addressed, but it is coincidental that they at least look as if they belong to the same surface category; as Grimshaw (1982a) points out, the ordering restrictions on the clitics are easier to account for if they are all the same category. In lexicalist theories, like CRG and LFG, NPs are NPs, PPs are PPs, and clitics are clitics. There are two positions which are assigned the OBJ relation, the NP and the clitic, and either one can be filled (or both in languages that allow clitic doubling). Transitive verbs are simply specified for an OBJ, and whether this requirement is fulfilled by an NP or a clitic is an accident of the syntax.

Another argument from Grimshaw (1982a) comes from Spanish. In Spanish, full (nonclitic) OBJs can be either NPs or PPs headed by the preposition \( a \) 'to'.

(26)  
\[ \begin{align*}  
\text{a.} & \quad \text{Juan vio algo.} \\
& \quad \text{saw something} \quad \text{'Juan saw something.'} \\
\text{b.} & \quad \text{Juan vio a alguien.} \\
& \quad \text{saw to someone} \quad \text{'Juan saw someone.'} 
\end{align*} \]

Grimshaw shows that \( a+NP \) OBJs behave like PPs, not NPs. However, under clitic doubling, \( a+NP \) which fills an OBJ slot shows up as an "accusative" clitic \( (lo, la) \) but \( a+NP \) which represents an indirect object shows up as a "dative" clitic \( (le) \).

(27)  
\[ \begin{align*}  
\text{a.} & \quad \text{Juan le / *lo habló a Pedro.} \\
& \quad \text{dat acc spoke to} \quad \text{'Juan spoke to Pedro.'} \\
\text{b.} & \quad \text{Juan *le / lo vio al perro.} \\
& \quad \text{dat acc saw to-the dog} \quad \text{'Juan saw the dog.'} 
\end{align*} \]

Thus, the choice of clitic is triggered by relational structure. Furthermore, Grimshaw points out that under a categorial system of subcategori-
zation the verb *ver* would have the subcategorization frame (28).

(28) \[
\begin{array}{c}
\text{NP} \\
\text{PP}
\end{array}
\]

There is nothing in principle to prevent a verb from selecting only NP, yet there are no verbs that do. In LFG and CRG, which recognize subcategorization as relational, transitive verbs like *ver* are subcategorized for an OBJ without regard for category.

A final argument from Grimshaw (1982a) deals with passivization (and also provides evidence in favor of a lexical treatment of the passive as opposed to RG's syntactic treatment). The basis of the argument is that English allows both NP and S SUBJs, but arguably only allows NPs as OBJs. In the terminology used here, based on Bresnan (1982a), clausal "objects" bear the relation COMP[lement]. Passivization, it will be recalled from Chapter 1, is a relationally-based rule which has the effect of promoting an OBJ to SUBJ. It is repeated as (29).

(29) **Passivization**

Relational change:

\[
\begin{array}{c}
\text{OBJ} \leftrightarrow \text{SUBJ} \\
\text{SUBJ} \leftrightarrow \text{OBL}_{\text{dy}} / \emptyset
\end{array}
\]

Morphological:

past participle

Passivization of a COMP is ungrammatical.º

(30) a. Linguists think that Hamsterian syntax is very complex.

b. *That Hamsterian syntax is complex is thought (by linguists).

Grimshaw notes that the assumption of relational subcategorization forces the view that Ss cannot be OBJs in English because there are verbs that can take propositional arguments either of both categories S and NP (e.g. *believe*) or of only one category (e.g. *pray*, which only takes S). What is interesting is the behavior of verbs which select propositional OBJs
but not COMPs (Ss). Grimshaw cites two such verbs: express and reflect.

(31)  
  a. The grammar expresses \( \text{[NP the fact that the rule is obligatory]} \).
  
  b. *The grammar expresses \( \text{[S that the rule is obligatory]} \).

(32)  
  a. This formulation of the rule reflects \( \text{[NP the fact that all NPs behave uniformly]} \).
  
  b. *This formulation of the rule reflects \( \text{[S that all NPs behave uniformly]} \).

The syntactic argument list for both these verbs is (33). Passivization of this argument list yields (33b).

(33)  
  a. {SUBJ, OBJ}
  
  b. {OBL\textsc{by}/\emptyset, SUBJ}

A comparison of (33a,b) predicts that the second argument, which has to be realized as an NP in the active, can be an \( S \) in the passive. The reason is that the only way that the second argument is constrained to be an NP in the active is by virtue of being an OBJ; since SUBJs can be \( S \)s, there is no reason to expect \( S \)s to be excluded in the passive. This prediction is borne out.

(34)  
  a. \( \text{[S That the rule is obligatory]} \) is expressed by the grammar.
  
  b. \( \text{[S That all NPs behave uniformly]} \) is reflected by this formulation of the rule.

The problem with relational subcategorization arises when one considers the class of verbs that take predicative complements (XCOMPs).

As is well known, predicates that take XCOMPs do not necessarily allow the full categorial range allowed by the syntax. XCOMPs in English can be NPs, PPs, APs, VPs (participles and bare infinitives), and VPs (to infinitives). Given the assumption of relational subcategorization, we would expect all verbs that take XCOMPs to allow all five categories.
In fact, very few do. There are three possible explanations for this.

(35)  

a. XCOMP is really a cover term for NCOMP, PCOMP, etc.

b. What look like categorial restrictions on XCOMPs are really semantic restrictions on the nature of the complement due to the semantics of the governing predicate.

c. Subcategorization has to be able to make reference to categories in addition to grammatical relations.

The position in (35a) was adopted in early work on LFG.3

...various types of [X]COMP function (e.g. NCOMP, ACOMP, PCOMP, VCOMP) can be marked by categorial feature. [Bresnan (1982b)]

The symbol XCOMP is used to designate the set of adjectival complements (ACOMPs), noun complements (NCOMPs), prepositional complements (PCOMPs), and verbal complements (VCOMPs). [Neidle (1982)]

The problem with this view is that it represents a conflation of relational and configurational (categorial) information. There are relational differences between SUBJs, OBJs, OBLs, etc., but NCOMPs are no different from PCOMPs in any way other than categorial identity. Just as a difference between OBJ and COMP is motivated on the basis of relational behavior (such as passivization), NCOMP and PCOMP cannot be distinct relations unless they are relationally distinguishable. This view has therefore been rejected in the more recent LFG literature, such as Bresnan (1982a).

The current LFG view is basically the one stated in (35b). Under this view, the apparent categorial selection is an illusion, due to the fact that different categories tend to map into different kinds of meanings (see Jackendoff 1983 for a description in one semantic framework of the kinds of meanings that different categories usually map into). This view is clearly stated by Maling (1983, 225–6).

Subcategorization frames have traditionally been stated in terms of syntactic categories, as shown in [(36a)], on the basis of contrasts like those shown in [(36b)]:
b. [36] seem, [__ AP]
consider, [__ NP AP]

b. Sandy seems {^out of town._}
clever.

We consider Sandy {^out of town._}
clever.

But it is imply not the case that only adjectives can occur in
these contexts, as shown by the impeccable sentences under [(37)].

[(37)] a. Robin looks a bit under the weather today.
b. Kim was acting out of character.
c. That suggestion seemed completely off the wall.
d. Everyone considers Kennedy out of the running.
e. We found the patient in good spirits.

The internal structure of the complement phrases in [(37)] is
obviously that of PP. Their ability to occur in what are considered
adjective frames cannot be attributed to some idiosyncratic property
of a particular preposition, e.g. out of, since the very same
prepositions are sometimes good and sometimes not, as illustrated
by the contrasts in [(38)]....

[(38)] a. Lee sure seems {^out of town._}
out of it.

b. Lee sure seems {^out of the house._}
out of his mind.

c. Lee sure seems {^onto the roof._}
onto something.

d. Lee sure seems {^under the old apple tree._}
under the weather.

Such contrasts show clearly that what is relevant is not the
syntactic category of the complement, but rather the distinction
between locative and directional PPs on the one hand, which are bad
in these contexts, and what I will call the 'metaphorical' PPs,
which have nonliteral, nonlocative readings, and which are good
in these contexts.

(35b) is surely the explanation of at least some cases of apparent cate-
gorical subcategorization. Consider (39).

(39) a. *Max began a hamster vendor. (N)
b. *Max began happy. (A)
c. *Max began in the kitchen / down the road. (P)
d. (i) *Max began being a hamster vendor. (V)
(ii) Max began selling hamsters.
e. (i) ??Max began to be a hamster vendor. (to)
(ii) Max began to sell hamsters.

Begin apparently does not allow NP, AP, or PP arguments. However, there
is a semantic generalization to be captured, one that is especially clear
from (39d(i)) and (39e(i)), which are also bad ((39e(i)) is slightly
better for unknown reasons). The rather transparent generalization is that the complement to begin cannot be stative. A similar verb is illustrated in (40).

(40)  
   a. *Max continued a hamster vendor. (N)  
   b. *Max continued happy. (A)  
   c. (i) *Max continued in the kitchen. (P)  
      (ii) Max continued down the road.  
   d. (i) *Max continued being happy.  
      (ii) Max continued running down the road.  
   e. (i) *Max continued to be a hamster vendor. (to)  
      (ii) Max continued to sell hamsters.

The difference between begin and continue is (40c(ii)): continue allows a directional PP. Presumably, this means that continue has some inherent physical motion as part of its meaning which begin does not. Even in the absence of an explicit semantic analysis, it is clear that the paradigms in (39) and (40) do not require appeal to grammatical categories. A less obvious case involves proclaim and declare.

(41)  
   a. The king proclaimed the hamster the national rodent. (N)  
   b. *The dictator proclaimed the banks nationalized. (A)  
   c. ?*The witch proclaimed Max under a curse. (P)  
   d. *The king proclaimed the hamster being the national rodent. (V)  
   e. (i) The king proclaimed the hamster to be the national rodent.  
      (ii) *The king proclaimed the banks to be nationalized.  
      (iii) The witch proclaimed Max to be under a curse.

The verbs proclaim and declare involve naming in some broad sense of the term which we will not formalize here. It is pointed out by Simpson (1983) that, in general, only NPs can act as names, and therefore verbs like name typically take only NP XCOMP$s$. As to the ability of proclaim to take VPs, note (42).

(42) The president named him to be on the new commission.

The ungrammaticality of AP, PP, and VP XCOMP$s$ in (41) can once again be reasonably brought under the rubric of (35b).
The point of the foregoing is that (35b) covers a lot of ground. However, it must be kept from being a vacuous claim; that is, it must be testable. What we have attempted to show is that plausible semantic accounts can be built up even in the absence of an explicit semantic theory. The question that must be answered is this: can all apparent cases of categorical selection of XCOMPs be reduced to semantics? If not, then we must accept (35c) and, in doing so, reject LFG in favor of a configurational theory such as CRG.

The structure of the argument will be to examine several verbs that take XCOMPs and see which of the impossible categories, if any, cannot be ruled out on semantic grounds. The problem is that, for the most part, different grammatical categories map onto different types of meanings. Thus, the "metaphorical" PPs cited by Maling are very important. In addition, we will appeal to several "minimal pairs" such as cold_{AP} / a cold day_{NP}, green_{AP} / a pale shade of green_{NP}, worthless_{AP} / worth nothing_{PP}, near_{NP} / near to near_{AP}. The category identity of the two AP/NP pairs should be obvious, but the AP/PP pairs are less so. Maling (1983) argues that despite its adjectival meaning, worth is a preposition in Modern English. On the other hand, she argues that near is a locative adjective. One problem she notes with her analysis is that the prepositional modifier right, which cannot be an AP specifier, can occur with near, but she neglects to point out that this is true only if near is followed by a bare NP.

(43) a. Put it right near the door!
   b. *Put it right near to the door!

Also, nearer and nearest, which are adjectival comparative and superlative forms, sound a lot less natural when followed by a bare NP than plain
near does. There is, of course, no reason to expect that near has to
always be the same category: Jackendoff (1977) suggests that we/us and
you are ambiguous between N and DET, and Falk (to appear) shows that
treating need and dare as ambiguous between V and M accounts for their
peculiar syntactic and semantic properties.

The major problem is the fuzziness of the semantic restrictions
involved. Consider the following, due to Grimshaw (personal communication).

(44)  a. The Martian turned green.

b. The Martian turned a pale shade of green.

(45)  a. It turned cold at noon

b. *It turned a cold day at noon.

Turn, when used with colors, can take both NPs and APs, but for weather
only APs are acceptable. Similarly, consider (46).

(46)  a. Max went insane/crazy.

b. *Max went sane.

We will proceed despite these problems. We will deal with several verbs,
grouped into rough semantic classes to highlight the syntactic differen-
tces. The verbs are:

(47)  a. Verbs of continuing: resume, keep (on)
b. Verbs of maintenance of state: stay, remain
c. be
d. Verbs of becoming: become, get, grow
e. Verbs of appearance: appear, seem
f. make
g. prove
h. consider
i. Verbs of permission: let, allow

We begin with verbs of continuing. Like continue and begin, these
verbs do not allow stative complements.
(48)  a. *Max resumed a hamster vendor.  (N)
    b. *Miriam resumed happy.  (A)
    c. *Max resumed down the road  (P)
    d.  (i)  *Miriam resumed being happy.  (V)
        (ii) Miriam resumed running down the road.
    e. *Max resumed to run down the road.  (to)

(49)  a. *Max kept (on) a hamster vendor.  (N)
    b. *Miriam kept (on) happy.  (A)
    c. *Max kept (on) down the road.  (P)
    d.  (i)  ??Miriam kept (on) being happy.  (V)
        (ii) Miriam kept (on) running down the road.
    e. *Miriam kept (on) to run down the road.  (to)

We can account for the inability of these verbs to occur with NP, AP, and
PP complements on semantic grounds, as already demonstrated. However,
these verbs display an additional inability to occur with VP. It should
be noted that no such inability exists with respect to the semantically
related verbs: continue and begin, as well as start (which behaves like
continue) and cease and commence (which behave like begin). This suggests
very strongly that the constraint here is a categorial one, not semantic.
Thus, from the categorial point of view, these verbs allow NPs, APs, PPs,
and VPs, but not VPs.

The two verbs of maintenance of state present an interesting contrast.
As the descriptive name suggests, these verbs require stative complements.
Keeping that restriction in mind, there should be no constraint on what
can appear as an XCOMP. Let us first consider remain.

(50)  a.  (i) The Martian remained a pale shade of green.  (N)
    (ii) It remained a cold day despite the southerly breeze.
    b.  (i) The Martian remained green.  (A)
    (ii) It remained cold despite the southerly breeze.
    c. Max remained in the giant hamster cage.  (P)
    d. *Max remained being in the giant hamster cage.  (P)
    e. *Max remained to be in the giant hamster cage.  (to)

Initially, the ungrammaticality of (50d,e) may look puzzling and therefore
appear to be another case of categorial subcategorization. In this case,
however, this is incorrect. Jackendoff (1976, 1983) claims that STA' is
a semantically primitive predicate which takes two arguments: a Theme and (in abstract semantic fields, an abstract) Location. *Remain* only takes nonpropositional locations. Sentences in which *be* is present have the primitive predicate *BE* in their semantic representations, and are therefore propositional. The semantic representations of the grammatical sentences in (50) have the matrix predicate *STOP* and an abstract nonpropositional location. The ungrammatical cases involve a propositional location with the predicate *BE*. Thus, if Jackendoff is right in treating *be* as a meaningful verb and not just a placeholder, the ungrammaticality of (50d,e) follows. Now consider *stay*.

(51)  
   a. (i) *The Martian stayed a pale shade of green.* (N)  
       (ii) *It stayed a cold day despite the southerly breeze.*  
   b. (i) ?The Martian stayed green. (A)  
       (ii) It stayed cold despite the southerly breeze.  
   c. *Max stayed in the giant hamster cage.* (P)  
   d. *Max stayed being in the giant hamster cage.* (V)  
   e. *Max stayed to be in the giant hamster cage.* (to)  

The paradigm in (51) is almost identical with the one in (50). The only difference is that (at least in this author's dialect) *stay* does not allow an NP XCOMP. Thus, *remain* allows everything while *stay* allows APs, PPs, VPs, and VPs.

The verb *be* itself is not free of categorial subcategorization.

(52)  
   a. Miriam is an accomplished neurosurgeon. (N)  
   b. Max is proud of his pet hamster. (A)  
   c. Max is at the hamster farm. (P)  
   d. Max is lecturing on the syntax of Hamsterian. (V)  
   e. Miriam is to lecture on the syntax of Hamsterian. (to)  
       (grammatical, but different verb *be*)  

(52a-d) are all perfect sentences with the verb (Helping Verb in the terminology of Falk, to appear) *be*; (52e), while it seems to involve the same verb, has been recognized by several researchers on the English auxiliary as a totally different verb which, unlike ordinary *be*, can never appear in nonfinite form and does not correspond to Jackendoff's (1976)
primitive predicate BE. If this is right, then ordinary be has to be lexically marked to allow only NPs, APs, PPs, and VPs.

The tree verbs of becoming listed in (47d) are very interesting. They all take stative nonlocative XCOMP's. The easiest one is become.

(53)  a. Miriam became a midwife despite harassment from reactionary medical authorities. (N)
     b. Due to the flooding of the market by gerbils, hamsters became worthless. (A)
     c. *Due to the flooding of the market by gerbils, hamsters became worth nothing. (P)
     d. *Miriam became being a midwife. (V)
     e. *Miriam became to be a midwife. (to)

(53d,e) might be explainable on grounds similar to remain and stay (except that we are dealing with a GO verb, not a STAY verb). However, it is not clear that the impossibility of a PP can be explained: the PP in (53c) is not locative, and it is synonymous with the AP in (53b). It is thus a candidate for exclusion on categorial grounds. Become is thus categorically subcategorized for NP, AP, VP, and \( \overline{VP} \) XCOMP's. Next consider get and grow.

(54)  a. *Miriam got a good linguist. (Ungrammatical if a good linguist is an XCOMP; grammatical on the irrelevant OBJ reading) (N)
     b. Miriam got near to the candidate at the rally. (A)
     c. (i) Miriam got near the candidate at the rally. (P)
        (ii) Miriam's hamster got under the cage.
     d. Miriam got accepted by an exclusive club. (V)
     e. Miriam got to teach linguistics. (to)

(55)  a. *Max grew a good hamster vendor. (see note under (54a)) (N)
     b. Due to the flooding of the market by gerbils, hamsters grew worthless. (A)
     c. *Due to the flooding of the market by gerbils, hamsters grew worth nothing. (P)
     d. *Max grew being a good hamster vendor. (V)
     e. Max grew to be a good hamster vendor. (to)

First of all, it should be noted that our explanation of the ungrammaticality of (53d,e) means that become is fundamentally different semantically from get and grow; this is not implausible. It does not seem that any of
the ungrammatical readings in (54) and (55) can be excluded on semantic
grounds, and if they cannot we have further proof for configurationality.

One important question about these data is the following: are all the
instances of get in (54) the same verb? In particular, (54e) has an
implication of success absent in the other cases. This researcher's
feeling is that this get is a distinct lexical item. If so, get is
subcategorized for AP, PP, and VP; if (54e) is the same verb, then VP
must be added to the list. Grow selects AP and VP.

The two verbs of appearance, seem and appear, display an interesting
contrast.

(56)  a. *The Martian seems a pale shade of green. (N)
b. The Martian seems green. (A)
c. The car seems out of control. (P)
d. *Max seems sell(ing) hamsters. (V)
e. Max seems to sell hamsters. (to)

(57)  a. *The Martian appears a pale shade of green. (N)
b. The Martian appears green. (A)
c. *The car appears out of control. (P)
d. *Max appears sell(ing) hamsters. (V)
e. Max appears to sell hamsters. (to)

Both verbs exclude NPs, but appear excludes PP as well. On the seman-
tic side, all locatives are out. Thus, seem takes XCOMPs which are APs,
PPs, and VPs, and appear, APs and VPs.

Next we consider (58).

(58)  a. I made Max a hamster vendor. (N)
b. I made Max crazy. (A)
c. *I made Max out of his mind / under a curse / worth nothing. (P)
d. I made Max sell hamsters. (V)
e. *I made Max to sell hamsters. (to)

As (58c) shows, even "metaphorical" PPs are excluded by make, which is
subcategorized for NP, AP, and VP.

(59) illustrates the facts about prove.
(59)  a. Miriam proved Max a competent linguist.  (N)  
b. Miriam proved Max worthless.  (A)  
c. *Miriam proved Max worth nothing.  (P)  
d. *Miriam proved Max be(ing) a competent linguist.  (V)  
e. Miriam proved Max to be a competent linguist.  (to)

Prove's XCOMP can be an NP, AP, or VP.

We conclude with a pair of verbs about which we will be less con-
clusive.

(60)  a. *I let Max a hamster vendor.  (N)  
b. *I let Max happy.  (A)  
c. (i) ?Miriam let Max in her home state.  (P)  
    (ii) Miriam let Max into her home state.  
d. (i) I let Max be a hamster vendor.  (V)  
    (ii) I let Max sell giant hamsters.  
e. (i) *I let Max to be a hamster vendor.  (to)  
    (ii) *I let Max to sell giant hamsters.

(61)  a. *I allowed Max a hamster vendor.  (Grammatical if a hamster 
    vendor is OBJ)  (N)  
b. *I allowed Max happy.  
c. (i) ?Miriam allowed Max in her home state.  (P)  
    (ii) ?Miriam allowed Max into her home state.  
d. (i) *I allowed Max be(ing) a hamster vendor.  (V)  
    (ii) *I allowed Max sell(ing) giant hamsters.  
e. (i) I allowed Max to be a hamster vendor.  (to)  
    (ii) I allowed Max to sell giant hamsters.

(60a,b) and (61a,b) seem to be excluded on semantic grounds, particularly
if we consider be to be meaningful (Jackendoff 1976) as the (d(i)) cases
are grammatical. Although, as the c. cases show, the semantics of these
verbs are not identical, they are presumably similar enough for an
adequate semantic theory to rule out NPs and APs. But note:

(62)  a. (i) ?I let the hamster near to me.  (A)  
    (ii) ?I let the hamster near me.  (P)  

b. (i) *I allowed the hamster near to me.  (A)  
    (ii) I allowed the hamster near me.  (P)  

From (62b(i)), it would appear that allow excludes APs categorically. Un-
fortunately, this leaves the status of NP unclear; there is no locative
noun to test. We will assume that NPs are ruled out the same way as APs:
semantically for let and categorially for allow.
This concludes our survey of XCOMP-taking verbs. It should be noted that we have omitted any discussion of verbs which only take VPs because the existence of this class is well known. In the case of many of these verbs, it is possible that semantic constraints can be found. It should also be noted that we have assumed no semantic difference between bare infinitives (VPs) and to infinitives (VPs); this may not be correct in all cases.

There is another potential source of evidence for categorial subcategorization. The groups of categories allowed by each verb we have examined should, minimally, be expressible in terms of an independently motivated system of categorial distinctive features. We have such a feature system in the Appendix to this dissertation. In the theory of categories adopted there and in Chapter 4, VP is the maximal projection of the category C[omplementizer]. Keeping that in mind, we present in (63) all the possible combinations of XCOMP categories and their feature decompositions.

(63) a. Attested

<table>
<thead>
<tr>
<th>Category</th>
<th>Feature Decomposition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A/V</td>
<td>[-gram]</td>
<td>make, consider, prove</td>
</tr>
<tr>
<td>N/A/C</td>
<td>[αPGR, -abasic]</td>
<td>be, keep (on), let, resume</td>
</tr>
<tr>
<td>N/A/P/V</td>
<td>[+basic]</td>
<td>become</td>
</tr>
<tr>
<td>N/A/V/C</td>
<td>[agram, -abasic]</td>
<td>appear, grow</td>
</tr>
<tr>
<td>A/C</td>
<td>[αPGR, -αOA, -abasic]</td>
<td>get</td>
</tr>
<tr>
<td>A/P/V</td>
<td>[agram, -αOA, +basic]</td>
<td>seem</td>
</tr>
<tr>
<td>A/P/C</td>
<td>[αPGR, -αOA]</td>
<td>stay, get(?)</td>
</tr>
<tr>
<td>P/C</td>
<td>[+gram]</td>
<td>allow</td>
</tr>
</tbody>
</table>

b. Unattested

<table>
<thead>
<tr>
<th>Category</th>
<th>Feature Decomposition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A/P/C</td>
<td>unstatable</td>
<td>N/A</td>
</tr>
<tr>
<td>N/P/V/C</td>
<td>unstatable</td>
<td>N/P</td>
</tr>
<tr>
<td>N/P/V</td>
<td>unstatable</td>
<td>N/V</td>
</tr>
<tr>
<td>N/V/C</td>
<td>unstatable</td>
<td>N/P/C</td>
</tr>
<tr>
<td>N/C</td>
<td>[αPGR, -OA, -abasic]</td>
<td>A/V</td>
</tr>
<tr>
<td>N/A/P</td>
<td>[agram, αPGR, +basic]</td>
<td>A/V/C</td>
</tr>
<tr>
<td>A/P</td>
<td>[αPGR, -αOA, +basic]</td>
<td>P/V</td>
</tr>
<tr>
<td>V/C</td>
<td>[+PGR, αOA, abasic]</td>
<td>P/V/C</td>
</tr>
</tbody>
</table>
The first thing we note is that the four combinations that are listed in (63) as unstatable are also unattested. We also note that simple one-feature specifications for the features [grammaticalized] and [basic] are attested (the only [-basic] category which can be an XCOMP is C, or VP, which is specified for many verbs). The only problem is that [+PGR], [-PGR], [+OA], and [-OA] are unattested. We have no explanation, but we note that this is not the only such case in the lexicon. There are lexical items that can be seen as ambiguous between N and V: level 1 zero-derived deverbal nouns (Kiparsky 1982). Similarly, we have seen that near is ambiguous between A and P, and Maling (1983) shows that several lexical items have changed from A to P as English has evolved. However, there are no lexical items that are ambiguous between V and P ([+PGR]), N and A ([-PGR]), V and A ([+OA]), or N and P ([-OA]).

There is a problem raised by the evidence for categorial selection of XCOMPs: how do we reconcile this with all the evidence for relational subcategorization? Why, for instance, are there no transitive verbs in Spanish that specify that the OBJ has to be an NP? I suspect that further research on the internal structure of the lexicon may provide an answer. Unlike constituent order, which is purely syntactic, subcategorization has to meet the constraints imposed on it by two components. This study is not the place to examine this question in detail. The fact that categorial subcategorization exists at all, no matter how it needs to be constrained, is evidence for a configurational theory of syntax.
2.3. The Syntax-Semantics Interface

In LFG, all semantics is read off of the level of functional structure. There is, of course, a sense in which grammatical relations are associated with semantics. The sense in which the LFG view is correct is the same as the sense in which the Standard Theory was right about deep structure being the input to semantic interpretation: it determines how the predicate's argument positions are filled (Jackendoff 1972). This is to be expected. To see why, consider the role of the Lexical (subcategorizable) Relations: they define the mapping between configurational position and thematic relations. Consider the structure of the sentence Miriam gave hamsters to Max, shown in §1.2 and repeated here.

(64)

The lexical entry of gave is (65).

(65) 

The lexical entry specifies, for example, that the Goal is realized as an OBL, which the constituency rules specify as a PP in VP. The point is that grammatical relations are the medium that the lexicon uses to map between semantics and syntax, so it is not surprising that it is the representation of grammatical relations that seems to serve as the
input to that aspect of semantic interpretation that relates to the filling of argument positions. We adopt the following convention.

(66) **Argument Feature Convention**

Annotate the lexical features of each constituent which bears a Lexical Relation to the argument list position which it fills.

Among the lexical features are agreement features and some representation of semantics, as well as the anaphoric features which will be discussed in Chapter 3.

With respect to how argument positions are filled, LFG and CRG are essentially identical. However, LFG claims that all semantics is derived from the f-structure representation, while the structure of CRG strongly suggests that both configurational and relational information are available to the semantic component. Our conception of what kind of information is relational and what kind is configurational suggests two places to look for the effects of configurational information: items bearing the same relation occurring at different "bar" levels, and the effects of ordering.

Of these two areas, the harder one is the first. There are two reasons for this. In the first place, theories of bar-level constituency are still fairly controversial; for concreteness we will assume the basic constituencies proposed by Jackendoff (1977). Secondly, given the current state of our ignorance concerning the Syntactic Relations (those not associated with argument positions), it is always possible to propose an additional Syntactic Relation to account for differences in interpretation. We could evade this second problem by restricting our attention to the Lexical Relations, but this turns out not to be very fruitful. In English, for example, PPs and 5s which are arguments (XCOMP,
OBLs, COMPs) can appear under either X' or X". There is some sort of vague markedness associated with the X" position, but nothing tangible. Therefore, it is necessary to look at the Syntactic Relations. It is not surprising that the Lexical Relations are less reliable, because their interpretation is based on the lexical entry of the governing predicate, whereas the interpretation of Syntactic Relations is determined directly from the syntax.

We will examine adjeuncts of category S. At the X" level, Ss are relative clauses, comparative clauses, and degree clauses; at the X' level they are nonrestrictive (appositive) relative clauses. The X" clauses share the property of having an effect on ("restricting" or "modifying") the "meaning" of the phrase in which they are located, where by "meaning" we mean things like reference of an NP, the truth conditions on an S, etc. Nonrestrictive relatives, on the other hand, are more on the order of parenthetical comments which do not change the "meaning" of the phrase. For example, (67a) means something like "that specific hamster vendor that I hate" while (67b) means "a particular hamster vendor about whom it happens to be true that I hate him."

(67)  
a.  the hamster vendor who I hate  
b.  the hamster vendor, who I hate

It is not a general property of daughters of X" that they do not effect the "meaning" of their phrase. Most obviously, English POSS and SUBJ are daughters of NP and S respectively, both of which are maximal projections (X""). Less obviously, determiners and degree modifiers, which bear the Syntactic Relation SPEC, effect the "meaning" in a very important way, often related intimately to relatives and degree clauses when they are present. Even when there is no relative clause, it is obvious that, for
instance, the distinction between the determiners the and a(n) makes an important difference on the reference of an NP. There are two ways to account for the distinction: we can invent a new relation NRR for nonrestrictive clauses (68a) or we can do it configurationally (68b).

(68)  
   a. ADJs effect the "meaning" of the phrase in which they are located; NRRs do not.

   b. ADJs which are daughters of X" effect the "meaning" of the phrase in which they are located; those which are daughters of X"" do not.

If we adopt (68a) we are forced to explain why it is that all adjuncts at the X"" level are NRRs and all adjuncts at the X" level are ADJs, for it is not only Ss that exhibit this contrast, but PPs, and, especially, adverbs, where one way of viewing the distinction between VP adverbs and S adverbs is as related to the restrictive/nonrestrictive distinction.

Nonrestrictive modifiers are best exemplified by appositive relative clauses in NPs..., but they also occur as appositives in other major phrasal categories and as sentential adverbials.... For present purposes, it is sufficient to note that they do not contribute to the identification of a referent, but instead comment on a referent already identified by the rest of the major phrasal category. [Jackendoff (1983, 73-4)]

It seems that (68b) represents a better generalization, and eliminates the need for an otherwise unmotivated Syntactic Relation. If we are correct, this is a case of configurational semantic interpretation.

The effects of constituent order on semantic interpretation are very familiar. Minimally, many (though not all) free constituent order languages assign one order unmarked (or neutral) status and associate other orders with emphasis on specific arguments. A case in English which is intuitively clear but hard to pin down comes from sentential adverbs, which can appear initially, between the SUBJ and the first auxiliary (AUX or M), between the AUX and the VP, and finally with comma
However, the meaning of the sentence changes slightly with the adverb shifted. Roughly, in initial position and between the SUBJ and AUX, the adverb is assertive. One indication of this is that these positions are impossible in questions, and nonassertive adverbs like *possibly* sound rather unnatural in these positions.

(69)  
   a. *Evidently, is Max going to win the election?  
   b. *Is evidently Max going to win the election?  
   c. Is Max evidently going to win the election?  
   d. ??Possibly, Max will adopt a hamster.  
   e. ?Max possibly will adopt a hamster.  
   f. Max will possibly adopt a hamster.

(69c,f) show that these adverbs are not inconsistent with the sentences in question, as there is no strangeness in the position immediately preceding the VP, the most neutral position. Thus, semantic interpretation has to have access to information about constituent order. Another case of interpretation based on ordering is evident in Icelandic (Maling and Zaenen 1981) and English (Falk 1983a) and has to do with the position of the head of S in direct (main clause) questions. In Icelandic, the finite verb is initial in the S according to the analysis of Maling and Zaenen, and in English the AUX precedes the SUBJ NP. The rules for interpreting questions in these languages must be able to refer to order.

A note is in order at this point on arguments about the role of order in phenomena that are f-structural in LFG. There is a weakness in our line of argumentation, which is that it is possible to define a notion of *f-precedence*, i.e. precedence defined at f-structure. Such a definition has been provided by Bresnan (class lectures, 1981, MIT): it turns out to be inadequate for present purposes. Essentially the definition requires assigning a more important role to the variables involved in the
instantiation procedure of Kaplan and Bresnan (1982), which defines an algorithm for mapping between c-structure and f-structure. Familiarity with the algorithm, which is discussed in §§4.2-4.4 of Kaplan and Bresnan (1982), will be assumed here. Bresnan's proposal can be summarized as (70).

(70)  

(a) **Convention**

*When assigning actual variables for metavariables, assign them sequentially in the tree left-to-right and top-down.*

(b) **Definition**

*For f-structure constituents \(\alpha\) and \(\beta\), we say that \(\alpha\) \text{*f-precedes*} \(\beta\) iff the variables associated with \(\alpha\) are lower than those associated with \(\beta\).*

For the purposes of this discussion, we will follow Bresnan (1982a) as opposed to Kaplan and Bresnan (1982) in annotating the features of a lexical item to the lexical item itself rather than to the lexical category node which dominates it. Consider (71a). Its annotated c-structure with variables assigned according to Convention (70a) is (71b), and the f-structure is (71c).

(71)  

(a) *Will Miriam sell the hamster?*

(b)  

\[
\begin{align*}
S & \rightarrow \text{NP}\text{VP} \\
\text{NP} & \rightarrow \text{DET}\text{N} \\
\text{VP} & \rightarrow \text{AUX}\text{NP}\text{VP} \\
\text{AUX} & \rightarrow \text{\'will\'} \\
\text{DET} & \rightarrow \text{\'the\'} \\
\text{N} & \rightarrow \text{\'hamster\'}
\end{align*}
\]
The problem is that the AUX will does not f-precede the SUBJ. The reason is that, in LFG terminology, only the values of functions are involved in f-precedence, not features. By virtue of the very definitions supplied by the algorithm for transforming c-structures into f-structures, the f-structure associated with a clause is the same as the one associated with its head. Thus, f-precedence cannot be invoked in any case where the position of a phrasal head is in question.

There is a further point that must be raised vis-à-vis the notion of f-precedence. This is an embellishment to LFG that must not be taken lightly. Precedence is certainly not required in the representation of surface grammatical relations, and to the extent that f-structure is supposed to be universal, it is undesirable. To add a notion of f-precedence is to weaken the distinction between c-structure and f-structure by allowing the formal representation of a c-structure notion in f-structure. It is thus a tacit recognition of configurationality. One could equally well propose a notation for bar level or grammatical category and turn f-structure into a bizarre notation for configurational X-bar-based structures.
2.4. The Syntax-Phonology Interface

The major role of constituent structure in LFG is as the input to the phonological component. The idea that surface phrase structure is rich enough to determine the phonetic realization of a sentence is certainly not unique to LFG; it dates back to early transformational syntax and generative phonology (Chomsky 1957, 1965; Chomsky and Halle 1968) and persists in current transformational (GB) theory. It is certainly true to some extent that phonology relies heavily on surface grammatical configurations, just as it is essentially true that semantics is based on relational information (§2.3). The reason for this is easy to see: one of the primary uses that phonology makes of syntactic information is the creation of metrical trees. It is to be expected that the starting point for a metrical tree would be an already branching structure, and the only branching structure recognized in CRG is configurational. In addition, it is in the phonetic interpretation that precedence is realized, and precedence is a configurational notion. As with semantics, however, it is not enough to show that grammar tends to meet LFG predictions; it must be the case that all phonologically relevant facts are recoverable from the configurational representation alone. We will see that this is not the case.

The first example we will discuss comes from Selkirk's (1972) account of French liaison. In discussing what she calls Style I, colloquial spoken French, the style with the fewest liaisons, the following footnote appears (pp. 289-90, fn 13).20
A further fact which also escapes a purely syntactic explanation is the non-occurrence of liaison between a monosyllabic auxiliary [i.e. a monosyllabic form of avoir 'to have' or être 'to be'] and an adverb in conversation familiale. In the (a) sentences, liaison occurs in Style I, but in the (b) sentences it doesn't:

[(72)]
   a. Elie et Yolande ont accueilli leurs petits enfants.
   b. Elie et Yolande ont effetueusement accueilli ...

[(73)]
   a. Ils ont accusé de complot tous les étudiantes de Première.
   b. Ils ont arbitrairement accusé ...

[(74)]
   a. Jeanne s'exprimée devant toutes ses camarades.
   b. Jeanne s'est ardement exprimée ...

One might say that an auxiliary has a close relation to its verb, and for this reason is in a liaison context with it. When an adverb intervenes, the unit is, in some sense, broken, and the Aux does not make liaison with the adverb.[]

Selkirk is correct in her informal explanation of the failure of liaison in the b. examples of (72)-(74), but she is mistaken that there is no purely syntactic explanation. In a purely configurational theory there is none, but in CRG this fact can be given a relationally based explanation, based on the fact that the adverbs are adjuncts. We will assume, not unreasonably, that Style I represents the unmarked case, and state the following universal principle.

(75) Liaison-type rules may not link a phrase head with a following ADJ (in the unmarked case).

Thus, the relational distinction between arguments and adjuncts is relevant to the operation of phonological rules such as liaison.

Internal to LFG and CRG, a strong case for reference to the relational representation (f-structure in LFG) can be made based on phrasal stress in English. The argument is based on that of Bresnan (1971), which aimed to show that a Standard Theory surface structure did not encode enough information for the assignment of sentential stress. Bresnan notes that the Nuclear Stress Rule (NSR) of Chomsky and Halle (1968) stresses the final major category of a sentence (excluding pronouns, which cliticize onto the verb; see Jacobson 1982), but there are sentences that do not
follow this pattern. She notes the following contrasts (using the Chomsky and Halle notation for main stress):

(76) a. George has plans to leave. (=George plans to leave.)
    b. George has plans to leave. (=George has plans which he wants to leave.)

(77) a. Helen left directions for George to follow. (=Helen wants George to follow her.)
    b. Helen left directions for George to follow. (=Helen wants George to follow the directions she left for him.)

(78) a. Mary liked the proposal that George leave.
    b. Mary liked the proposal that George left.

In the a. cases in (76)-(78), the final $S$ is a complement to the N and the NSR applies as expected. In the b. cases, where the $S$ is a relative clause, stress is not where it is expected. Consider also the following.

(79) a. John asked what books Helen had written.
    b. What books has Helen written?
    c. George found some friends he'd like you to meet.

In the b. examples of (76)-(78) and in (79), the word bearing primary stress is the antecedent of a missing element which in current theories is represented by a trace, a configurationally represented phonetically null element. Bresnan, accounting for these facts in pre-trace-theoretic transformational syntax and the Chomsky and Halle (1968) theory of stress (premetrical) proposes that the NSR applies in the syntactic cycle. (79a) is derived as (80).

(80)  
\begin{align*}
\text{Deep structure:} & \quad [\text{John asked} \ [\text{COMP} [\text{Helen had written some books}])] \\
\text{Lexical stress:} & \quad [1 \ 1 \ [\text{COMP} \ [1 \ 1 \ 1]]] \\
\text{First cycle:} & \quad 2 \ 2 \ 1 \\
\text{Second cycle:} & \quad \text{some books} \ 2 \ 2 \ \emptyset \\
\text{Third cycle:} & \quad 2 \ 2 \ 1 \ 3 \ 3
\end{align*}
A derivation of this nature is not available in current theories, but trace theory provides an alternative. First we state the NSR as Liberman and Prince (1977) do, in metrical terms.

(81) In a configuration $[_{C}AB_{C}]$, if C is a phrasal category, B is strong. Assuming that the auxiliary have, which is stressless, cliticizes onto the following verb, the metrical structure of (79a) following the application of NSR is (82).

(82)

\[
\text{John asked what books Helen had written e}
\]

The final output of the phonological component has to look something like (83).

(83)

\[
\text{John asked what books Helen had written e}
\]

To derive the representation in (83) from (82), we propose the following rules, ordered after NSR.
(84)  
  a. \([AB] \rightarrow \text{sw if } A \text{ is the antecedent of a trace dominated by } B\).
  
b. \([AB] \rightarrow \text{sw if } B \text{ is phonetically null.}\)

We can also account for (85).

(85)  
John asked what Helen had written.

The difference between (85) and (79a) is that in the former, unlike the latter, the final constituent is stressed. This is because \(wh\)-pronouns are unstressed. If we account for this as we accounted for \(have\), by assuming that \(what\) cliticizes, NSR yields (86).

(86)

Now note that while (84b) applies as expected, (84a) cannot apply.

The \(A\) of the structural description is the node dominating the metrical constituent [what Helen], which is not the antecedent of the trace, although it contains the antecedent. The resulting metrical structure is (87).

(87)

The problem is (84a). In LFG, fronted phrase - trace relations
are represented in f-structure, not c-structure. For instance, the c-structure and f-structure of (81a) are (88).

(88) a.

\[
\text{S} \\
\text{NP} \\
\text{John} \\
\text{VP} \\
\text{asked} \\
\text{DET} \\
\text{N} \\
\text{NP} \\
\text{AUX} \\
\text{V} \\
\text{NP} \\
\text{what} \text{books} \\
\text{Helen} \\
\text{had} \\
\text{V} \\
\text{NP} \\
\text{written} \\
\text{e}
\]

b. 

\[
\begin{array}{l}
\text{SUBJ} [\text{PRED 'John']} \\
\text{TENSE past} \\
\text{PRED 'ask<(SUBJ)(COMP)>'} \\
\text{COMP} [\text{FOCUS [SPEC 'what']} \\
\text{PRED 'book'} \\
\text{NUM pl} \\
\text{WH +} \\
\text{SUBJ [PRED Helen]} \\
\text{TENSE past} \\
\text{PRED 'perf<(XCOMP)>(SUBJ)>'} \\
\text{XCOMP [SUBJ} \\
\text{PART perf} \\
\text{PRED 'write<(SUBJ)(OBJ)>'} \\
\text{OBJ}]
\end{array}
\]

The link between what books and e is not encoded in the c-structure (88a).
Thus, unless the phonological component has access to the information represented in (88b), (84a) cannot be stated. In LF, then, this rule would have to be global, referring to two distinct levels of representa-
tion. In CRG, too, the fronted phrase - trace relation is relationally represented, as discussed in §3.8. Thus, (84a) is a configurational rule. Clearly, in a framework such as Government/Binding theory this rule would be considered purely configurational, because the coindexing of fronted phrases (operators) and traces (variables) is configurationally represented. However, at least internal to LFG and CRG, we do have another piece of evidence here for configurationality.

It has been pointed out to me by Jane Grimshaw (personal communication) that these cases do not require a single rule to be configurational; i.e. neither the condition on liaison nor the rules for sentential stress in English actually make reference to both grammatical configurations and grammatical relations: rather, it is the phonology as a whole that must have access to both. One problem with looking at phonology and its interaction with syntax is that it is not always clear what type of information is being used. A case in point is what is often referred to as "wanna contraction," the process by which certain sequences of want+to are contracted to wanna. We will follow Jacobson (1982) in considering this to actually be the cliticization of to onto want (or any other verb), followed by an idiosyncratic rule deleting the ts. We will therefore refer to the rule as "want-to cliticization." In CRG, two of the descriptions in the literature are available to account for the following contrast.

(89)  a. Who do you wanna see?
    b. *Who do you wanna win?

Anticipating the discussion in §3.8 on the representation of fronted ("in COMP") phrases and adopting a makeshift notation for coreference in Equi (89a) and Raising-to-Object (89b), the syntactic representations are (90).22
The configurational description, due to Jacobson (1982), is that the empty category, or trace, in (90b) blocks cliticization of *to* onto *want* because cliticization must be to an adjacent category. The relational description is Postal and Pullum's (1977), which claims that the relevant distinction is that in (90a) *want* and the complement verb "share" a *SUBJ* while in (90b) they do not. These two approaches, both of which can be adopted in CRG, make the same predictions for a wide range of facts con-
carning want-to cliticization, but each of the cited articles mentions data which are problematic for the other theory: Jacobson brings in data from other V-to cliticizations which cannot be accounted for by appeal to SUBJ-sharing, and Postal and Pullum show that want-to cliticization is impossible in Heavy NP Shift sentences, which do not involve a trace in the present framework. The result is that it is unclear whether or not relational information is needed for the cliticization. However, if Postal and Pullum are correct, even in part, this rule is a counterexample to the LFG approach.

2.5. Configurelationality in Anaphora

2.5.1. Overview

In this section, we will discuss evidence for configurelationality from anaphora. We include under the term "anaphora" the phenomena of personal pronouns, reflexives, reciprocals, NP Raising, Equi, SuperEqui, and syntactic binding (wh movement constructions in the framework of Chomsky 1977). The justification for this will be deferred until Chapter 3, which presents a detailed account of anaphora within the representation-al system of Configurelational Grammar. Here we will outline the arguments for configurelationality; cross-references to Chapter 3 will be provided where applicable so that the interested reader will be able to find a more detailed discussion, but it is not necessary to refer to Chapter 3 to understand this section.

We will follow a principle of representation that has often been as-sumed in nontransformational approaches (such as Lexical-Functional Grammar and Generalized Phrase Structure Grammar): only elements which have phono-logical (though not necessarily phonetic) realization are configurationally
represented. As not all anaphoric elements are phonologically realized, this means that anaphora in CRG cannot be defined as a relation between nodes. Rather, anaphoric binding is defined between positions in syntactic argument lists. Each anaphor is associated with a set of features which determine the binding properties of the anaphor, just like the features [ipronominal] and [tanaphoric] in GB (Chomsky 1982). This set of features is annotated to the appropriate argument position, and the link between antecedent and anaphor is represented as a dashed line connecting the set of anaphoric features with the argument position that represents the antecedent, as in (91).

(91)  a. The hamster saw himself.

   b. 

\[
\begin{array}{c}
\text{S} \\
\text{VP} \\
\text{NP} \\
\text{DET} \\
\text{NP} \\
\text{V} \\
\text{NP} \\
\text{SPEC} \\
\text{N} \\
\text{NP} \\
\text{OBJECT} \\
\text{NP} \\
\text{SUBJECT} \\
\text{S} \\
\end{array}
\]

The bundle of dots in (91b) represents the anaphoric features associated with reflexives in English. We will propose a set of features in Chapter 3; for the remainder of this chapter some transparent descriptive notation (such as "refl" for reflexives) will be used. In (91b), the dashed line indicates that the SUBJ (the hamster) is the antecedent.

It is very important for anaphora to be configuralational. This is because of the highly important role played by anaphora in the CRG conception of grammar. The central processes of grammar are split into two groups: relation-changing rules, which are lexical and not configuralational,
and anaphora. Aside perhaps from constituent order, all the evidence we have uncovered has been derived from marginal areas of syntax. If it were to turn out that these central areas of syntax were not configurational, an explanation would have to be supplied. In the case of relation-changing rules, there is a reason to exclude configurationality, namely the fact that they are lexical, not syntactic. There is no such reason in the case of anaphora. Dealing as it does with semantics and being argument list based, it might be expected to have a stronger relational component, but both relational and configurational aspects must be present.

2.5.2. Anaphora in General

We begin by discussing some rather superficial, but nevertheless illuminating, evidence that anaphora is configurational. It is generally agreed that anaphora is subject to precedence and c-command restrictions. Precedence is generally considered to be a configurational property. In LFG (at least in its published versions) and in the theory of Marantz (1981) the representation of grammatical relations is unordered, and in Relational Grammar, where grammatical relations are the only kind of syntactic representation available, ordering is defined on grammatical relations at a very late stratum, after the application of relationally based rules (Johnson and Postal 1981). In some recent LFG research, an attempt has been made at defining a notion of f-precedence; as we showed in §§2.1 and 2.3, this notion is inadequate for at least some of the evidence we have presented, and essentially turns LFG into a configurational theory. The command relation is generally thought to be configurationally defined. This is a problem for a theory like LFG in which anaphora is relationally represented. An alternative to c-command, Mohanan's "c-command at
f-structure" (Mohanan 1983b), which is termed *f-command* by Bresnan (1982a), may seem promising, even though it was originally proposed for a condition on control. The most general restriction, the one that applies to all anaphors, is that the anaphor may not c-command its antecedent (Reinhart 1976). We can test to see whether or not this restriction can be said to hold at the level of f-structure. Consider (92). Its c-structure is presented as (93), and its f-structure in two notations as (94): (94a) is conventional f-structure notation, and (94b) is a tree representation in which the c-command relation may be more discernable (see Mohanan 1983b).

(92) The senator's contradicting the president embarrassed him.

(93)

(94) a. [SUBJ [SUBJ [PRED 'senator'] [SPEC 'the'] [CASE GEN] [PRED 'contradict<(SUBJ)(OBJ)>'] [OBJ [PRED 'president'] [SPEC 'the'] [TENSE past] [PRED 'embarrass<(SUBJ)(OBJ)>'] [OBJ [PRED 'him']]]]
b. 

Both the senator and the president can antecede him; neither is c-commanded by him in (93), but in (94) we can see that both are f-commanded (or c-commanded at f-structure) by the pronoun. Thus, c-command at c-structure seems to be required. We will see in §2.5.4 that c-command itself is configurational, but for present purposes we can think of it as essentially configurational. There are thus two essentially configurational restrictions on anaphora. On the other hand, there are languages in which particular anaphors require their antecedents to be SUBJs, or even not to be SUBJs, as discussed by Mohanan (1982b) and in §3.5 of this dissertation. Note that the requirement that an anaphor have (or not have) an antecedent bearing a certain grammatical relation is a relational specification.

It is not clear whether an approach like GB theory or the framework of Marantz (1981), which claim to derive special properties of SUBJs from their "government" properties, can account for this without simple stipulation. However, if they do this, then the claim of explanatoriness of theories in which SUBJs are not governed vanishes with respect to anaphors. It may be explanatory to say that PRO can only appear in certain contexts (SUBJ of nonfinite clauses) because it is a pronominal anaphor and therefore must be ungoverned to satisfy Binding Theory (GB's theory of anaphora); it is not explanatory to say that SUBJ (or some equivalent notation, such as [NP,S]) must antecede certain anaphors. Similarly, we cannot appeal to the GB-internal property of the SUBJ as an "external argument" (an argument whose θ-role is not assigned directly by the verb
but rather compositionally by the entire verb phrase), because the grammatical relations involved here are *surface* grammatical relations, and θ-role assignment, although it occurs at LF (logical form), is to D-structure (deep structure) position. A passive SUBJ receives its θ-role in the position of the trace it binds in the VP. Thus, every theory must use a relational stipulation here, either directly in terms of grammatical relations, as we do, or by positing an abstract level of "lexical structure" or "virtual structure" in which all languages have VPs so the relations SUBJ and OBJ can be represented uniquely. We thus take the view that this is a true relational constraint on some anaphors. More theory-internally, anaphors which have no phonological realization are only represented relationally in both LFG and CRG. Thus, anaphora is both relational and configurational, and must be either represented at both kinds of levels of representation or must be represented at a configural level.

2.5.3. Reflexives and Reciprocals

We will now look at reflexives and reciprocals and see what evidence they offer. They are often thought to be subject to a clause-mate condition. In LFG (Bresnan 1979; class lectures, 1980 and 1981, MIT) this is restated as a clause-nucleus-mate condition, where a clause nucleus is defined either as an f-structure constituent containing both a SUBJ and a PRED (Bresnan 1979, 1980) or an f-structure constituent whose PRED is a lexical form (i.e. has argument structure) (Bresnan 1982a). For our purposes, the latter is identical to CRG argument lists and the former is identical to a clause. Thus, we are left with two alternatives for the restriction on antecedents of reflexives and reciprocals: a clausemate restriction or a coargument restriction, where coarguments are arguments...
in the same syntactic argument list. However, "picture noun" reflexives show that neither of these alternatives is correct. Consider the following, adapted from Jackendoff (1972).

(95)  
a. Max told Bill a story about himself. (himself = Max or Bill)  
b. Tom told Dick Harry's story about himself. (h' = Harry; himself ≠ Tom or Dick)  
c. That the picture of himself in the post office enrages Charlie. (himself = Charlie)

The syntactic representations are (96).

(96)  
a.  
\[ S \rightarrow \langle SUBJ \rangle \rightarrow \langle VP \rangle \rightarrow \langle NP \rangle \rightarrow \langle V \rangle \rightarrow \langle OBJ \rangle \rightarrow \langle OBJ2 \rangle \rightarrow \langle NP \rangle \rightarrow \langle DET \rangle \rightarrow \langle N \rangle \rightarrow \langle OBL \rangle \rightarrow \langle PP \rangle \rightarrow \langle HEAD \rangle \rightarrow \langle refr \rangle \rightarrow \langle about \rangle \rightarrow \langle himself \rangle \]

b.  
\[ S \rightarrow \langle SUBJ \rangle \rightarrow \langle VP \rangle \rightarrow \langle NP \rangle \rightarrow \langle V \rangle \rightarrow \langle OBJ \rangle \rightarrow \langle OBJ2 \rangle \rightarrow \langle NP \rangle \rightarrow \langle POSS \rangle \rightarrow \langle N \rangle \rightarrow \langle OBL \rangle \rightarrow \langle PP \rangle \rightarrow \langle HEAD \rangle \rightarrow \langle refr \rangle \rightarrow \langle about \rangle \rightarrow \langle himself \rangle \]
In (95a), *himself*’s possible antecedents are clause-mates but not coarguments. In (95b), on the other hand, *himself* must be bound to its coargument; in this case it is the clause-mate restriction that fails. In (95c), *himself*’s antecedent is neither a clause-mate not a coargument. The key to understanding (95) is the fact that in (95a,c) the reflexive does not have a coargument, while in (95b) it does. In (95c) the next argument list up contains only a predicative complement (XCOMP) which is not referential and therefore cannot serve as an antecedent for a reflexive. Thus, *Charlie* is the closest possible antecedent. The way we formalize this idea in §3.3 is by treating the argument list containing the anaphor (i.e. the anaphor’s coarguments) as the basic domain which can be extended upwards if there is no possible antecedent. In English, this extension goes one argument list at a time but is not limited. Other languages have different rules for extending the domain. Even in English, it is not quite as simple as it first seems. Note (97).
(97)  
\begin{enumerate}
\item *Max thinks that himself is smart.
\item *Max thinks that himself saw Joe.
\end{enumerate}

In (97a) there is no possible antecedent for the reflexive in the comple-
ment clause and the only potential antecedent in (97b) is c-commanded by
the anaphor. However, reflexives (and reciprocals) in SUBJ position
apparently cannot look out of the clause for their antecedents. Note that
(95c) shows that the extension of the binding domain is not clause-bounded
in general; we have already established that there is no clause-mate
restriction on English reflexives. On the other hand, there are langu-
ages where the clause does seem to be an absolute boundary, such as
Russian.

(98)  
(from Yang 1983)

*Dissidenty znali chto stat'i drug o druge pojavilis'
dissidents knew that articles each about other appeared
v zapodnoj presse.
in western press
'The dissidents knew that articles about each other had appeared
in the western press.'

Let us examine these conditions. The basic idea of extending the binding
domain is relational, both because it is stated in terms of argument lists
and, more significantly, because noncategorial information about NPs is
required: the post office in (95c) should be a possible antecedent except
that it is in an ADJ. Russian, however, clearly has a categorial require-
ment: the domain cannot be extended beyond the S node. Thus, the Russian
reciprocal is subject to a configurational constraint. The status of
the English restriction is less clear; in §3.3 we state it configurational-
ly by imposing a restriction that for the domain to extend, the anaphor
must not be the argument of V or M. If this formulation is correct, then
English, too, has a configurational constraint.
One language in which reflexives clearly are governed by a configurational constraint is Malayalam, described by Mohanan (1982b). The reflexive in question must be bound to a SUBJ. This SUBJ must be a coargument of the reflexive. However, if the reflexive (swa-) is a SUBJ or is inside a SUBJ and separated from the root of the SUBJ only by NP nodes, the entire structure becomes the binding domain. Interestingly, this is not a one-list-at-a-time extension like English. Mohanan states the condition as follows:

(99) **Definition**

\( \alpha \text{ NP-contains } \beta \text{ iff}
\begin{align*}
(\text{i}) & \quad \alpha \text{ is } 3 \text{ or} \\
(\text{ii}) & \quad \alpha \text{ dominates } \beta \text{ with no intervening non-NP nodes.}
\end{align*}
\)

(100) If swa- is not NP-contained by the subject, it must find its antecedent in its minimal clause nucleus.

Mohanann's formulation ("NP-contained by SUBJ") clearly brings out the configurationality.

One final case, although slightly more tentative, comes from Dutch. The Dutch reflexive zich has some puzzling properties which are discussed in §§3.4 and 3.6. In §3.4 we propose that zich, which must be bound to a SUBJ, takes as its domain the minimal \( \tilde{S} \) which dominates it, minus its coarguments. We note in a footnote there that this analysis, which we take from Yang (1983), may not be correct, but if it is it represents another case of configurationality, referring to SUBJ and coarguments (relational) and \( \tilde{S} \) (configurational/categorial).

2.5.4. C-command

The idea that an anaphor may not be configurationally "higher" that its antecedent is one that has pervaded much of the generative work on anaphora, both in transformational approaches and interpretive ones. The
formalization of "higher" that gained favor in the 1970s was Langacker's (1967) "precede-and-command." Reinhart (1976) shows, however, that precede-and-command should be replaced by a different relation, which she calls \textit{c(onstituent)-command}. For instance, in (101) the pronoun in each example is in the matrix clause, so it commands everything. Yet the pronoun can precede its antecedent. (Here and elsewhere in this section, we use underlining to indicate coreference.)

(101)  
\begin{enumerate}
\item The chairman hit \underline{him} on the head before the \underline{lecturer} had a chance to say anything.
\item We finally had to fire \underline{him} since McIntosh's weird habits had finally reached an intolerable stage.
\item Rosa won't like \underline{him} anymore, with Ben's mother hanging around all the time.
\item We'll just have to fire \underline{him}, whether McIntosh likes it or not.
\item Believe it or not, people consider \underline{him} a genius in Ford's home town.
\end{enumerate}

In (101), pronominalization can go in either direction. In other words, switching the anaphor and its antecedent in each case will result in another grammatical sentence. There are, however, other cases where backwards pronominalization is fully grammatical but forward pronominalization is impossible. Compare (102) and (103).

(102)  
\begin{enumerate}
\item Near \underline{him}, Dan saw a snake.
\item In \underline{her} bed, Zelda spent her sweetest hours.
\item For \underline{his} wife, Ben would give his life.
\item How obnoxious to \underline{his} friends Ben is.
\item Fond of \underline{his} wife though Ben is, I like her even more.
\item (I predicted that Rosa would quit her job and) quit \underline{her} job Rosa finally did.
\end{enumerate}
(103) a. *Near Dan, he saw a snake.
   b. *In Zelda's bed, she spent her sweetest hours.
   c. *For Ben's wife, he would give his life.
   d. *How obnoxious to Ben's friends he is.
   e. *Fond of Ben's wife though he is, I like her even more.
   f. *(I predicted that Rosa would quit her job and) quit Rosa's job she finally did.

Thus, the precede-and-command condition is both too strong (101),(102), and too weak (103). Reinhart's proposal, which has been widely adopted, is that anaphors be disallowed from c-commanding their antecedents. She gives two alternative definitions of the c-command relation, a complete one and a simplified one.

(104) Definition

   a. (complete) Node A c(onsituent)-commands node B iff the first branching node $\alpha_1$ dominating A either dominates B or is immediately dominated by a node $\alpha_2$ which dominates B, and $\alpha_2$ is of the same category as $\alpha_1$.

   b. (simplified) Node A c(onsituent)-commands node B iff neither A nor B dominates the other and the first branching node which dominates A dominates B.

The difference between a. and b. relates to certain assumptions about constituent structure that Reinhart makes. That is to say, although she believes that topicalized PPs appear as in (105a(i)) and verb-phrasal PPs as in (105b(i)), she assumes the structures in (ii) for convenience. This allows her to use the simplified definition.
In both a. and b., NP₁ c-commands NP₂. As discussed in §2.1, there is evidence that (105a(ii)) is the correct constituency (although perhaps without the COMP node); as to verb-phrasal PPs, according to the theory of the X-Bar Convention as formalized by Jackendoff (1977), these PPs can appear either under V' or V'', making both (i) and (ii) partially correct. We can provisionally restate Reinhart's definition of c-command as (106).

(106) Node A c(onsituent)-commands node B iff the maximal projection of the node that immediately dominates A dominates B.

If (104) or (106) is the correct definition, then c-command is a purely configurational notion. If it were, it would not be a problem for CRG, as nothing prevents there from being definitions purely in terms of grammatical relations, but it would also provide no evidence for our theory. In this section, we will examine two pieces of evidence that argue against a purely configurational definition of c-command: one rather straightforward, the other somewhat less conclusive.

The first problem we will consider is one noted by Reinhart. Consider the following, from Reinhart (1976).
(107)  

a.  

(i) *It didn't surprise her that Rosa has failed the exam.

(ii) *It didn't occur to her that Rosa has failed the exam.

b.  

(i) *I met him in Ben's office.

(ii) *I spoke to him in Ben's office.

c.  

(i) *Someone should tell her that Rosa's driving is dangerous.

(ii) *Someone should point out to her that Rosa's driving is dangerous.

d.  

(i) *Him, I met in Ben's office.

(ii) *To him, I spoke in Ben's office.

e.  

(i) *Him, Max's mother gave a book.

(ii) *To him, Max's mother gave a book.

f.  

(i) *Him, Don's mother found a gun near.

(ii) Near him, Don's mother found a gun.

(108)  

a. *We told her (the truth) about Rosa's son.

b. *We talked to her about Rosa's son.

c. ??We talked with her about Rosa's son.

d. ??We talked about her with Rosa's son.

In (107), the (i) examples are predicted to be ungrammatical with the underlined NPs coreferential and the (ii) examples are predicted to be grammatical by the configurational definition of the c-command relation. Yet all are ungrammatical except for (107f(ii)). (108) shows that the generalization lurking behind (107) is not linked to some notion of "indirect object," although the judgments recorded by Reinhart (which do not match mine; I think all are equally bad) are less clear with prepositions other than to. Consider the syntactic representation of the sentences in (107c).
b.

\[
S \rightarrow \text{AUX} \rightarrow \text{V} \rightarrow \text{OBL} \rightarrow \text{PRT} \rightarrow \text{PP} \rightarrow \text{HEAd} \rightarrow \text{S}
\]

\[
\text{NP} \rightarrow \text{SUBJ} \rightarrow \text{XCOMP} \rightarrow \text{VP} \rightarrow \text{OUT} \rightarrow \text{OBL} \rightarrow \text{P} \rightarrow \text{HEAd} \rightarrow \text{S}
\]

\[
\text{S} \rightarrow \text{AUX} \rightarrow \text{VP} \rightarrow \text{NP} \rightarrow \text{POSS} \rightarrow \text{AP} \rightarrow \text{XCOMP} \rightarrow \text{NP} \rightarrow \text{SUBJ} \rightarrow \text{XCOMP} \rightarrow \text{S}
\]

\[
\text{should} \rightarrow \text{someone} \rightarrow \text{V} \rightarrow \text{out} \rightarrow \text{that} \rightarrow \text{to} \rightarrow \text{her}
\]

\[
\text{is} \rightarrow \text{driving} \rightarrow \text{Rosa's} \rightarrow \text{dangerous} \rightarrow \text{controllee}
\]
In (109a), the root of the c-command domain of the pronoun is, by our working definition, the V₂, which means that it c-commands Rosa and they cannot be coreferential. However, in (109b) the pronoun is inside a PP, and it is that PP that should serve as the root of the c-command domain. Similarly in the sentences with preposed NPs and PPs: in the (i) examples of (107d-f) the pronoun c-commands the intended antecedent while in the (ii) examples our definition says that it does not. An examination of (107f(ii)), which does match our predictions, supplies the necessary clue. In (107a-e) the PP represents an OBL argument of the verb (or a topicalized PP serving as the antecedent of a trace in V₂ serving the OBL function). *Find*, however, takes an XCOMP, a propositional argument which predicates something of the verb's OBJ. *Near* is a predicate, taking a lexical and a "controlled" SUBJ. The internal structure of the PPs in (107e(ii)) and (107f(ii)) are (110).

(110) a. PP  
   |  <HEAD>  
   |  NP  
   to him

b. PP
   |  <OBJ>  
   |  NP
   near {SUBJ, OBJ} him

Configurationally, there is no difference at all between the two PPs, but it is clear that they are quite distinct relationally. In (110a), the NP following the preposition is an argument of the predicate governing the PP rather than of the preposition itself; or, in the terminology we will introduce in §4.2.2, the PP in (110a) is the maximal functional head of the pronoun. It seems that the c-command relation is sensitive to this relational difference.

The second problem is one that Reinhart does not actually consider a problem. It relates to differences between two groups of PPs, which we
will refer to as *tight PPs* and *loose PPs*. The intuition behind this terminology is that tight PPs are more closely related to the verb than loose PPs: Reinhart sees this difference as being realized configurationally: tight PPs are verb-phrasal (daughters of VP) and loose PPs are sentential (daughters of S).

Like adverbs, certain PP's are inherently sentential or verb-phrasal, i.e., their position in the tree is fixed. Thus, as Williams (1974) points out, *in order to*-phrases can only be attached to S, while infinitive *to*-phrases, as in *We sent Rosa home to please her father*, can be attached either to S or to VP. Other sentential PP's discussed by Williams (1974, 1975) are *although*..., *causal since*..., *whether or not*..., and *with*-phrases of the type illustrated in [(111c)]. (I am following Emonds (1976) in his analysis of clauses etc. as PP's.) PP's which are always verb-phrasal (according to Williams and many others) are instrumental (*with*...) and manner (*by*...) PP's. In certain other cases, the PP itself is not inherently marked as to its position, yet its occurrence in a sentence with a verb which is strictly subcategorized to require a PP determines its obligatory position in the VP. Thus, locatives (e.g., *in NP*) are obligatorily verb-phrasal following verbs like *dwell, reside, or put*. Similarly, the verb *fright* requires a *with*-phrase, which will then be verb-phrasal.

In cases where the position of the PP is clear-cut, we get clear coreference judgments. While the sentences of [(111)] ... are possible in a proper discourse, no discourse will permit coreference in sentences like [(112)].

[(111)]

a. We sent him to West Point in order to please Ben's mother.

b. We'll just have to fire him whether McIntosh likes it or not.

c. Rosa won't like him any more, with Ben's mother hanging around all the time.

[(112)]

a. *Rosa tickled him with Ben's feather.

b. *It's time to put him in the baby's bed.

As we saw, given the c-command definition of domain, sentential PP's are not in the domain of the object. Hence the coreference restriction does not apply to block coreference in [(111)]. Verb-phrasal PP's, on the other hand, are in the domain of the object (c-commanded by the object). In the sentences in [(112)], then, the antecedent is in the domain of the pronoun, which violates the coreference assignment.

However, in many cases, the PP itself is not inherently marked (nor do selectional restrictions force its occurrence in only one position), and, as was first observed by Kuno [(1975)], its position in the sentence seems to depend on subtle semantic considerations. I will illustrate this situation with two pairs of sentences with locatives in [(113)] and [(114)]. [Reinhart (1976, 59f); footnotes 1, 2, 3, and 4 omitted]
(113)  
   a. Rosa \{looks sick \(\text{is riding a horse}\}\) in Ben's picture.
   
   b. Rosa found a scratch in Ben's picture.

(114)  
   
   b. The gangsters killed Hoffa in Detroit.

In (113) and (114), the a. sentences involve loose PPs, which Reinhart analyzes as sentential, and the b. sentences have tight (verb-phrasal) PPs. Note (115).

(115)  
   a. Rosa is riding him in Ben's picture of Mr. Ed.
   
   b. People worship him in Kissinger's native country.
   
   c. *I can't find her in Ben's picture of Rosa.
   

Reinhart also notes the following contrasts involving fronted PPs.

(116)  
   a. In Ben's picture of Rosa, she is riding a horse.
   
   b. In Kissinger's native country, he is worshipped by many people.
   
   c. *In Ben's picture of Rosa, she found a scratch.
   
   d. *In Hoffa's home town, he was killed by gangsters.

This is due, according to Reinhart, to a further configurational contrast: PPs moved from within a VP (including loose PPs from embedded clauses) move into COMP, as in (105a), while PPs moved from directly under S become daughters of the node E (which Reinhart 1981 tentatively identifies with Chomsky's (1977) S node) and are not c-commanded by the subject.

The point of these PP cases is that for Reinhart's account to go through, loose PPs must be out of the c-command domain of the object (and of the subject when fronted). This means that a configurational difference, such as the one adopted by Reinhart, must be shown to exist, or at least be defensible. Evidence concerning constituency is often hard
to come by, and there are probably no tests which are both uncontroversial and fairly framework-independent, but it should be generally possible to survey the arguments and decide that one side or the other is stronger. What we will do is discuss Reinhart's evidence bearing on the (postverbal) positioning of loose and tight PPs, and show that all three tests she uses are rather weak. We will then propose two tests which show that Reinhart is wrong and that all the PPs are within the VP: one of these tests is controversial and the other is framework-dependent. The result is that, at least within the shared assumptions of LFG and CRG, Reinhart's account of the distinction between tight PPs and loose PPs is untenable. We will then discuss a possible relational difference between the two classes of PPs, but one which is even harder to argue for than constituency; we will suggest that tight PPs are arguments of the verbs they modify, while loose PPs are adjuncts. As stated at the beginning of this section, this argument is not as straightforward as the one involving a distinction between OBLs and XCOMP; it is at least a potential argument that relational information is needed. As a reference point, (117) shows Reinhart's constituents.

(117) a. loose PPs

```
S
  NP   VP
    V   NP
```

b. tight PPs

```
S
  NP   VP
    or
    V   NP
    PP
```

Reinhart's first constituency test involves the pseudocleft construction. Reinhart argues that the right-hand side of the pseudocleft should simply be the VP. Thus, we expect the pattern in (118).
The right-hand brackets in each case in (118) should be the VP of (117). These predictions are met.

(119)  a. (i)  *[what ... did] be [V ... PP_X]
        (ii)  [what ... did] be [V ... PP_X]

b. (i)  [what ... did + PP_X] be [V ...]
        (ii)  *[what ... did + PP_X] be [V ...]

The judgments in (119) are crystal clear; those in (120) somewhat shakier. Nevertheless, even (120) is strong enough to make this a very attractive argument. However, Reinhart does not present any evidence that the test is a valid one. Note that Reinhart draws on a parallelism between PPs and ADVPs at the very beginning of the quote on p. 80. The configurational positions of adverbs are less controversial than PPs, making them a good test. In the terminology of Jackendoff (1972), manner adverbs are verb-phrasal and speaker- and subject-oriented adverbs (Emonds 1976 calls these "factive adverbs") are sentential. We will concentrate on verb-phrasal adverbs because we wish to claim that all the PPs in question are verb-phrasal. The question is: how do verb-phrasal adverbs behave in pseudoclefts. Consider (121).
(121) Rosa rode a horse clumsily.

The adverb *clumsily* is generally agreed to be a manner adverb and therefore verb-phrasal. It should therefore pattern with the tight PPs in (118), the (ii) cases.

(122) a. What Rosa did was ride a horse clumsily.

b. What Rosa did clumsily was ride a horse.

(122a) is predicted to be grammatical, but (122b) should not be. Note that there is a difference between the two sentences in (122): in (122b), but not (122a), there is a presupposition that Rosa did something clumsily.

Reinhart (1976) quotes Kuno (1975) as showing that this kind of theme/rheme difference is quite general among the two classes of PPs: that is, loose PPs are always presupposed. Thus, it is possible for loose and tight PPs to have exactly the same configurational position under VP and display the contrast in (119)-(120). If a constituent on the right-hand side of the pseudocleft cannot be interpreted as presupposed, then loose PPs, which must be so interpreted, will never appear there. It seems, then, that the pseudocleft argument for constituency cannot be maintained.

Reinhart's second test concerns VP preposing, as in (123).

(123) They said that Mr. Spock would beam up to the Enterprise, and beam up to the Enterprise he did.

On the assumption that this construction is due to a rule which moves the VP, we expect the judgments schematized in (124).

(124) a. (i) *[V...+PP_P] pronoun did

(iii) [V...+PP_P] pronoun did

b. (i) [V...+PP_P] pronoun did

(ii) *[V...+PP_P] pronoun did

The bracketed string represents the VP. Reinhart claims, however, that
some people accept (124b(ii)), derived by moving only the lower VP in the fuller version (117b), which Reinhart considers to be the correct constituency. With this caveat, Reinhart presents the sentences and judgments in (125) and (126). (The lower-case letter and Roman numeral subdivisions are keyed to (124) and presented in the order they are for typographical convenience.)

(125) (i) I wanted Rosa to ride a horse in Ben's picture,
   a. *and ride a horse in Ben's picture she did.
   b. and ride a horse she did, in Ben's picture.

(ii) They wanted Rosa to find a scratch in Ben's picture
   a. and find a scratch in Ben's picture she did.
   b. *and find a scratch she did in Ben's picture.

(126) (i) It was predicted that people would worship Kissinger in Washington
   a. *and worship Kissinger in Washington they did.
   b. and worship Kissinger they did in Washington.

(ii) It was predicted that gangsters would attack Hoffa in Detroit
   a. and attack Hoffa in Detroit they did.
   b. *and attack Hoffa they did in Detroit.

The judgments are actually not as clear as Reinhart's annotations indicate. In the first place, there is a comma in (125(ib)), which is necessary for the sentence to be grammatical. To my ear, this comma is also necessary in (126(ib)), and with an inserted comma (126(iib)) sounds fine, too. As to (125(iiib)), it needs a slightly longer pause than the others, but it, too, is acceptable. Thus, the predictions of (124b) seem not to be borne out. What this means, according to Reinhart, is that the lower VP of (117b) can move alone. However, the necessity of comma intonation
is unexplained. What about (124a)? (125(ia)) and (126(ia)) do not sound as bad as they ought to under Reinhart's analysis. Significantly, they do not sound worse than (127).

(127) ??I told Miriam to tell Max that he was a real loser, and tell M Max that he was a real loser she did.

The clause in (127) is clearly a complement of tell, as the verb is strictly subcategorized for a clausal complement, and is therefore under the VP. The intuition about (127) is that one would like the preposed string to be as short as possible, and the one in (127) is too long. Perhaps this is a performance constraint, due to the fact that the construction skews the normal order in English, where the subject precedes the verb, and one would like to know what the subject is. In the case of PPs, which are shorter than clauses, if they are tightly associated with the verb they are apparently allowed, but loose PPs tend to be rejected.

Reinhart's final structural test is simply incorrect. She claims that verb-phrasal adverbs may not intervene between the subject NP and the VP, while sentential adverbs may. The example she gives is (128).

(128) John \{probably\} was eating a carrot.

She then presents (129) and (130), with anyway added to the PPs to avoid a reading where the NP and PP form a single constituent.

(129) a. Rosa, in Ben's picture (anyway), looks sick.
   b. *Rosa, in Ben's picture (anyway), found a scratch.

(130) a. People, in Washington (anyway), worship Kissinger.
   b. *The gangsters, in Detroit (anyway), killed Hoffa.

Note, however, that the auxiliary be, which appears under the Modal (or AUX) node when finite (Emonds 1976; Jackendoff 1972, 1977; Falk, to appear) appears in (128) but not (129) and (130). Compare (128) with (131).
(131) Bugs slowly ate the carrot.

Also, (129a) and (130a) sound very odd with anyway omitted. The ordering of sentential adverbs can be analyzed as being due to free ordering of adverbs under S (Keyser 1967; Jackendoff 1972, 1977; Falk 1983a); the addition of anyway should not have such a profound effect on the PPs if they are to be explained the same way. In addition, given (131), the ungrammaticality of (129b) and (130b) is left unexplained. I have no explanation for (129)-(130), but they do not provide evidence for the structural distinction that Reinhart is attempting to establish.

We will now counter Reinhart's constituency tests with two of our own. At least from the point of view of CRG and LFG, these tests are superior to Reinhart's.

The first test is the Right Node Raising construction. This test is somewhat controversial because people disagree on the basic facts (see, e.g., Abbot 1976), but this author's dialect is in line with the more restrictive version which can only RNR constituents. Thus, the following judgments hold in the author's dialect (from Bresnan 1982a, examples (144b) and (145b)).

(132) a. Mary admired, and Catherine disliked, Peter the Great.

b. *Mary gave, but Catherine did not give, Peter the grape.

Reinhart's analysis of loose and tight PPs predicts the judgments schematized in (133).

(133) a. X, conj Y, [...PP₁]

b. *X, conj Y, [...PP₂]

The reason is that tight PPs form a constituent with the verb, etc., viz. a VP, while loose PPs do not. The prediction in (133a) is correct, but (133b) is not.
(134) a. (i) Rosa can, but Zelda can't, find a scratch in Ben's picture.

   (ii) Gangsters did, but denied they would, kill Hoffa in Detroit.

b. (i) Max will, but Miriam won't, ride Mr. Ed in Ben's movie about famous horses.

   (ii) Republicans do, but Democrats don't, worship Kissinger in Washington.

None of the sentences in (134) are perfectly natural, but RNR sentences generally are not. They all sound grammatical, and the contrast of (133) is simply not attested. Of course, it is always possible to deny the relevance of RNR as a test for constituency; it is true that we do not have an analysis of RNR, but the constituents-only restriction is present, and this should serve as adequate justification for the use of this construction as a test.

The second test is based on a prediction about order made by structural statements: a VP constituent following the verb may not follow an S constituent following the verb. Therefore, if Reinhart is correct, in postverbal position loose PPs must always (not just in the semantically unmarked order) follow tight PPs. This is incorrect.

(135) a. (i) Miriam found a scratch [PP<sub>t</sub> in Max's picture] [PP<sub>c</sub> on Thursday].

   (ii) Miriam found a scratch [PP<sub>c</sub> on Thursday] [PP<sub>t</sub> in Max's picture].

b. (i) The muggers killed the old lady [PP<sub>t</sub> near her home] [PP<sub>c</sub> on her birthday].

   (ii) The muggers killed the old lady [PP<sub>c</sub> on her birthday] [PP<sub>t</sub> near her home].

Here, someone wishing to defend Reinhart's analysis of the PPs could take one of two approaches. One possibility would be to claim that the semantically marked orders are derived by "stylistic rules" in the PF-component, as in GB and its theoretical ancestors. Thus, at the level(s) of represen-
tation relevant for anaphora Reinhart's distinction might still hold.
Alternatively, one could assume the existence of "discontinuous constituents" (McCawley 1982) or "virtual structures" (Zubizaretta and Vergnaud 1982), in which case the ordering prediction no longer holds. Both of these alternatives discount the relevance of (135) to the question at hand. Within nontransformational frameworks such as LFG and CRG, stylistic movement rules are generally not recognized, and discontinuous structures have never been considered. Both weaken the predictive value of proposing constituency, making analyses like Reinhart's virtually unfalsifiable. For this reason, it is preferable not to resort to these approaches.

What we have seen is that Reinhart's evidence for claiming that loose PPs are sentential and tight PPs verb-phrasal is very weak. This does not mean that the distinction between the two types of PPs is invalid; after all, the distinction is motivated by empirical differences in their behavior. Rather, it means we must look elsewhere for an explanation.

What makes Reinhart's analysis of the PPs so appealing is the difficulty of constructing an alternative explanation, a difficulty I attribute to our comparative ignorance about syntactic structure over a quarter of a century after the publication of the first book on generative syntax. Our alternative will be even harder to support than Reinhart's, in that direct arguments are lacking. What we would like to suggest is that the root of the difference is a distinction which, until very recently, has had little or no importance in syntactic theory: the distinction between arguments and adjuncts. The closest notion in classical transformational theory is strict subcategorization: an item that is strictly subcategorized must be present in a sentence in order for it to be grammatical, such as the OBL PP argument of put. However, this test is useless for us, for
the tight PPs, which we would like to suggest are arguments, are not always obligatory.31

(136)  
   a.  Max put the hamster in its new Cadillac.  
   b.  *Max put the hamster.  
   c.  Gangsters killed Hoffa in Detroit.  
   d.  Gangsters killed Hoffa.

However, as shown by Bresnan (1980), this is not the only test for argumenthood. In particular, she argues that instrumental PPs are optional arguments of the verbs with which they appear, not adjuncts. The argument is based on the principle of Function-Argument Biuniqueness, which states that each semantic argument may be represented syntactically by only one syntactic argument, and each syntactic argument which represents a semantic argument may represent only one such argument.

[T]he biuniqueness condition admits many types of prepositional phrases as possible grammatical arguments of the verb. An example is the instrumental phrase in [(137)].

[(137)]  John escaped from the prison with dynamite.  [= John used dynamite to escape from prison]

No more than one instrumental phrase can occur with a single verb:

[(138)]  John escaped from prison with dynamite with a machine gun.  
          [= John used dynamite and used a machine gun to escape from prison]

Given the biuniqueness condition, this fact can be explained by analyzing the instrumental phrase as a grammatical argument of the main verb. It is possible to define a lexical rule of Instrumentalization (analogous to lexical rules of Causativization) which converts an n-adic predicate argument structure $P$ to an n+1-adic predicate argument structure $P$-with... For motivation, note that Instrumentalization alters the inherent semantic properties of a predicate as illustrated in [(139)-(140)]....

[(139)]  
   a.  John killed Harry.  
   b.  John killed Harry with dynamite.

[(140)]  
   a.  An explosion killed Harry.  
   b.  #An explosion killed Harry with dynamite.

(pp. 114f (165 in Bresnan, ed. (1982)); bracketed comments in (137) and (138) are Bresnan's)

Unfortunately, biuniqueness cannot be used to argue for an analogous rule of Locativization. Locatives can serve as adjuncts, unlike instrumentals;
note that locatives can be both loose and tight PPs. It is possible to suggest that in a sentence like (141), one PP (presumably the first) is an argument, and the other an adjunct.

(141) I found a hamster in my room under my desk.

In addition, while the semantic effects of Instrumentalization are fairly obvious (the SUBJ has to be able to use an instrument), there is no obvious semantic effect to Locativization, although (142) may be an unanalyzed example of one.

(142) a. In Kissinger's native country, he is widely worshipped.
   b. *In Zeus's temples, he was worshipped daily.

However, the fact that the nature of the semantic difference is not readily characterizable precludes this from being a convincing case. I am not aware of any other direct arguments for argumenthood. However, we will present an indirect argument for considering tight PPs to be arguments.

The source of our indirect argument is a semi-frequently discussed construction dealt with in LFG by L. Levin (forthcoming), which she calls Oblique Inversion. Examples of the construction are:

(143) a. In this pet store can be found some of the most expensive hamsters in the world.
   b. From that nation came some of the most important scientists in history.
   c. Out from under a rock slithered a most disgusting looking politician.

Levin analyzes this construction as involving a PP SUBJ which is not a thematic argument and is treated as a meaningless idiom chunk. This SUBJ is lexically specified by the verb as being identical to a clause-internal PP, and the "logical subject" is the OBJ. For instance, she provides the
following lexical entry for the Locative Inversion form of the verb *dwell*.

(144) *dwell*: V; (+PRED)=dwell<(OBJ)(OBL)>,
(+SUBJ FORM)=d((OBL),
(+SUBJ NUM)=d((OBJ NUM), (+SUBJ PERS)=d((OBJ PERS).

She shows that the class of verbs which can appear in this construction is a subset of those which have been treated in various frameworks as involving an underlying representation with an object but no subject, what Perlmutter (1978) calls *unaccusative verbs* and Burzio (1981) calls *ergative verbs*. Unlike *dwell*, there are cases of inverted PPs which are not obviously arguments.

(145) Between the trees twinkled the lights of cottage candles and far down flared bright windows of the village stores.

Intuitively, the locative PPs which have been "inverted" in (145) are not arguments of *twinkle* and *flare*, but adjuncts. This suggests a lexical entry along the lines of (146).

(146) *twinkle*: V; (+PRED)=`twinkle<(OBJ)>',
(+SUBJ FORM)=c(+ADJ), ...

However, the second line of (146) is problematic. First of all, Bresnan (1982a) states (p. 347 (287f in Bresnan, ed. (1982))) "The subcategorizable functions ... are the only functions to which lexical items can make reference." If this is literally true, then ADJ cannot be mentioned in lexical entry (146). If, on the other hand, we reject this view of the distinction between the Lexical Relations ("subcategorizable") and the Syntactic Relations ("nonsubcategorizable"), it is not clear that we have any distinction between them at all: the lexical form in (146) cannot occur in a clause without an ADJ. In addition, there is a technical problem with (146): according to Kaplan and Bresnan (1982) and many (although not all) of the other articles in Bresnan, ed. (1982), adjuncts are represented as members of a set. Kaplan and Bresnan state the following
(partial) phrase structure rule.

\[(147) \ VP \rightarrow V \ NP \ NP \ PP^*\]
\[\text{(+OBJ)}=+ \ (\text{+OBJ2})=+ \ \epsilon(\text{+ADJUNCTS})\]

This rule allows a sentence to have more than one ADJ without violating Functional Uniqueness. If this is the case, then the equation specifying the FORM feature of the SUBJ would have to make reference not to the clause nucleus's ADJ, but to a member of the ADJ set. The notation of the theory would have to be expanded to allow for this.

A way out of the difficulties with treating the PPs as ADJs would be to consider them arguments, deposited in the lexical entries of the verbs in question by our proposed rule of Locativization, and perhaps another rule or two to deal with nonlocative PPs. We will restrict our attention to locatives. We now have two constructions in which we would like to say that there are locative PPs which look like adjuncts but are really arguments. We now have a way of providing a suggestive argument that tight PPs are arguments: all we have to do is show that there is sufficient overlap between the verbs that take tight PPs and those that allow Oblique Inversion. I say "sufficient overlap" rather than complete overlap because there are extraneous factors involved in Oblique Inversion, such as its limitation to unaccusatives (or ergatives) and certain other restrictions of a semantic nature discussed by Levin. In (148)-(157) we present some of Levin's examples of Oblique Inversion involving locatives, and sentences using the same verbs testing c-command: what we predict is opposite grammaticality judgments for Oblique Inversion and coreference.

\[(148)\]
\[
a. \text{ Between the trees twinkled lights of cottage candles.} \]
\[
b. \text{*In the star's anticipated place in the sky, it twinkled merrily.} \]
(149)  a. In the airshow flew planes of many different models.
      b. *At the plane's testing site, *it flew beautifully.

(150)  a. In these hallowed halls had walked many famous people.
      b. *On Max's porch, he walked nervously

(151)  a. In this factory work people of strong moral character.
      b. *In Miriam's office, she works all day long.

(152)  a. In this yard will play children from all walks of life.
      b. *In Joey's yard, he plays all day.

(153)  a. Under the heat of the midday sun withered even the heartiest desert flowers.
      b. *In the plant's old pot, it would have withered.

(154)  a. In this deceptively calm-looking lake had drowned many unsuspecting tourists.
      b. *In Max's new swimming pool, he might drown.

(155)  a. On open fires boiled large pots of soup and stew.
      b. *In the soup's pot, it boiled for hours.

(156)  a. In war-torn villages exploded hidden mines and booby-trapped buildings.
      b. At the bomb's destination, it exploded.

(157)  a. *Before her disbelieving eyes shattered pieces of expensive china.
      b. In the crystal's box, it shattered.

(148)-(157) illustrate a remarkable relation between the two constructions. The multitude of question marks can probably be explained as a case of confusion between loose and tight PPs, which have slightly different semantics. The correlation is extremely close. Among non-unaccusative verbs, there is no correlation: Oblique Inversion is ungrammatical in all of them, as predicted by Levin.
These examples show that speak is not subject to Locativization, cry and laugh are, and kneel may or may not be, pending an explanation of the fuzzy judgment. None of these may appear in the Oblique Inversion construction because they are not unaccusatives. Note, however, that there is another class of verbs that undergo Oblique Inversion: some passives-

(162) In this forest can be seen some of the world's most vicious bunny rabbits.

Now recall the contrast in (142), repeated here.

(142) a. In Kissinger's native country, he is widely worshipped.
    b. *In Zeus's temples, he was worshipped daily.

This contrast is mirrored in (163).

(163) a. *In this city are worshipped some very prominent politicians.
    b. ??In this country are worshipped some of the world's strang-est deities.

(163b), although not particularly natural, sounds much better than (163a). This is another indication that there is some relationship between Oblique Inversion and the loose/tight PP distinction.

A possible objection that may be raised at this point is that it has not been shown what the difference is in the preposed PPs. Even if it is correct to claim that the (very vague) judgments on postverbal PPs are due
to argument/adjunct status, it is still possible that preposed PPs appear in two distinct positions, as suggested by Reinhart. As it turns out, Reinhart's evidence for a structural distinction in the case of preposed PPs is even weaker than her evidence for a structural distinction postverbally. As she herself points out, there really is no viable position for preposed loose PPs. She suggests that they are daughters of the node E, but notes that E, which has been implicated in left-dislocation structures, is generally considered nonembeddable, while clauses with preposed loose PPs freely embed. Her evidence for preposing loose PPs into a position other than "COMP" comes from contrasts like the following.

(164) a. In Ben's picture of her, how does she look? (loose)
    b. *In Ben's picture of her, what did she find? (tight)

(165) a. In Washington, who do they worship? (Loose)
    b. *In Detroit, who did the gangsters kill? (tight)

The intermediate structures of the sentences in (164) are (166) prior to the fronting of the wh phrase.

(166) a. 
The idea is that in (166a), movement-into-COMP goes through without a hitch, but in (166b) COMP is already filled, so \(\hat{w}\) movement cannot apply. In CRG, there is no single "in COMP" position; as we showed in §2.1, there is evidence that topicalized phrases do not occupy the same position as fronted \(\hat{w}\) phrases. There is, however, an alternative analysis which depends on the argument/adjunct distinction (or, more accurately, Lexical vs. Syntactic Relations) again. In (166b), the PP is a fronted XCOMP, which means that it bears the topicalization relation FOCUS and is syntactically bound to a trace in XCOMP position. As shown by Bresnan (1978), any relation can be preposed, and it is therefore impossible to associate the fronted position with a lexically-linked GR (such as XCOMP) directly. Adjuncts, on the other hand, can occur anywhere in the sentence (with certain limits specific to each individual type), and there is no reason to suppose that fronted adjuncts are FOCUSes instead of just ADJs. We propose in §3.8 that FOCUS, which is also involved in questions, is a Lexical Relation; if this is correct, it is subject to Uniqueness. (164b) and (165b) are ungrammatical because they contain two FOCUSes. Now consider (167).
In Ben's next picture, he promised Rosa that she would look more attractive.

According to Reinhart, although the PP here is sentential (loose) in its clause of origin, the fronting rule treats it as verb-phrasal because it gets fronted out of the matrix VP. In the CRG approach, this fronted PP cannot bear the relation ADJ because it is not an adjunct of the main clause but rather the subordinate clause. Therefore it bears the relation FOCUS (or some other Lexical Relation).

What we have seen in this section is that the definition of the c-command relation has to be sensitive to grammatical relations. Thus c-command, often considered the epitome of a configurational notion, proves to actually be configurational.

2.5.5. Island Constraints

By including syntactic binding as a kind of anaphora, CRG makes the prediction that island constraints can be configurational. Unfortunately, it is difficult to show this without a detailed account of island constraints and the anaphors that are subject to them. For this reason, we will defer the discussion of configurationality and island constraints until §3.9.4. It will transpire that there is further evidence for configurationality from our account of islands.

2.6. Summary

We have shown in this chapter that there is reason to believe that there is a configurational level of syntactic representation, at which the central processes and properties of syntax are determined. The evidence discussed does not exhaust the possibilities. As one further example, con-
sider subject-verb agreement in cases with conjoined subjects. In some
dialects of English, for instance, it is possible for a verb to agree with
the closest conjunct of a disjunctive subject, i.e. a subject with the
conjunction or. As Jespersen (1914) puts it

... extremely often the verb [with a disjunctive subject] is put in
the plural, the idea of plurality prevailing over that of disjunc-
tion...

The pl is inevitable if the word nearest to the verb is in the
pl: [e.g. Arthur Conan] Doyle [in The Adventures of Sherlock Holmes]
the vessel in which the man or men are. [first italics added;
¶6.62, p. 177]

In Biblical Hebrew, VSO sentences illustrate this sort of agreement even
in conjunctive subjects. The examples in (168) come from Gesenius, et al.
(1910, §146f,g, p. 468); underlined stops are spirantized.

(168) a. wayyahô nôâh übânâw
   and-came Noah and-sons-his 'Noah and his sons came'
   ms (Genesis 7:7)

b. wattaśan râhêl wâlê?â
   and-answered Rachel and-Leah 'Rachel and Leah answered'
   fs (Genesis 31:14)

c. wattaâdabbër miryâm wâ?ahârôn
   and-spoke Miriam and-Aaron 'Miriam and Aaron spoke'
   fs (Numbers 12:1)

d. wattaśigga/ gam lê?â wîlâdehä
   and-approached also Leah and-children-her
   fs 'Leah and her children also approached' (Genesis 33:7)

The Biblical Hebrew pattern can be seen in many languages (Hale, personal
communication), including Modern Irish (McCloskey and Hale 1983). Thus
agreement, a relational phenomenon, must be able to make reference to
constituent order, particularly closeness to the verb.

It is also interesting to note that other researchers have, on occa-
sion, noted configurational facts but not appreciated the fact that they
call for either configurational representation or globality. For instance,
Mohanan (1983a) argues that while Chomsky's (1981, 1982) binding condition C (R[everential]-expressions are free) applies at the configurational level, conditions A and B (relating to "anaphors" (reflexives and reciprocals) and pronominals) must apply at a relational level (what Chomsky 1981 calls "lexical structure" in his discussion of nonconfigurational languages). Thus, a version of Government/Binding theory that incorporates such a level (the syntactic structure of Marantz 1981 may be an equivalent) must allow global representation of anaphoric relations.

The survey in this chapter shows both the pervasiveness and the subtlety of configurationality. There is no principled distinction between those processes which are configurational and those which are not. This is what our theory predicts.
Footnotes for Chapter 2

1(3a) is a preferred α-rule because heavy constituents can be shifted over light Ss. I thank Haj Ross for pointing this out to me.

2If (6b) is not completely bad, it is certainly worse than (6a). There may be a marked ADV position under the root S.

3It is possible that people who disagree with these judgments assign a constituency more in line with the standard transformational view, with topicalized phrases out of S. This is quite plausible, as there is little overt evidence in English for the attachment of these phrases, unlike Icelandic (see below) where the "verb second" constraint provides direct evidence.

4Maling and Zaenen have the topicalized phrase Chomsky-adjointed to S. A strict reading of the X-Bar Convention does not allow Chomsky-adjunction (Jackendoff 1977), and having a flat structure allows an α-rule analysis. Maling and Zaenen do not present any evidence for the Chomsky-adjointed structure.

5See §4.1 for a discussion of the relations OBJ and OBJ2.

6The fact that (19a) is not quite as bad as (18b) and can be salvaged by making the PP heavier is accounted for by making the S ordering rule a preferred α-rule. (18b) cannot be so salvaged.

7I am ignoring the question of whether Romance "indirect objects" are objects or obliques. They are realized as PPs, just like their English counterparts, which are OBLs. However, there is some evidence from quantifier floating in Italian (Perlmutter 1979) that "indirect objects" pattern with direct objects and contrast with other α-headed PPs. Furthermore, indirect objects pronominalize as preverbal clitics, just as direct objects do, and Romance languages lack the double object construction. Thus, it is possible that indirect objects in Romance languages bear the relation OBJ2. For discussion, see Simpson (1983, §1.3.1).

8A possible objection to this line of reasoning is that while (30b) is ungrammatical, (i) is fine.

(i) It is thought (by linguists) that Hamsterian syntax is complex. A traditional transformational view would say that (i) is derived from (30b) by an obligatory Extrapo­sition. This view does not explain why Extraposition is obligatory sometimes (as in (i) and with verbs like seem) and optional other times (with any verb that can take a sentential SUBJ). In CRG and LFG, seem never takes a sentential SUBJ, even at an abstract lexical stratum. Instead, it takes a COMP. Similarly, (i) has a dummy SUBJ and a COMP. Bresnan (class lectures, 1981, MIT) has suggested that (i) is an example of an impersonal passive, limited in English to verbs that take sentential complements. The generalization discussed here can be taken as evidence that the complement clause in (i) is not a SUBJ at any level.
But note that in many of the same articles, SCOMP, which looks like another instance of XCOMP, is actually used in the sense of what is now called COMP, a non-(functionally-)controlled clause.

For the remainder of this section, when we present paradigms such as (39), we will place the symbol for the head of the category of the XCOMP in parentheses: N for NP, A for AP, P for PP, and V for VP. We indicate VP as to.

(39a) has an irrelevant reading in which a hamster vendor is an adjunct.

For some unknown reason, Max continued being a hamster vendor sounds pretty good.

Grow can occur with a PP when the preposition is into.

(i) Miriam grew into a competent hamster vendor.
The fact that there is such a strict restriction on the preposition indicates that grow into is a lexical unit, and that this is therefore a different verb grow.

This is grammatical for some people. This is to be expected if we are dealing with an idiosyncratic categorial restriction, but would be strange if the ungrammaticality were due to a semantic restriction.

The preposition worth is inexplicably strange with seem.

(i) a. Hamsters seem worthless. (A)
   b. ??Hamsters seem worth nothing. (P)

For some unknown reason, the intransitive form of prove can only take VPs.

Even in Government/Binding theory, where LF is taken to be the interface between syntax and semantics, argument positions (GF-o) are represented at D-structure.

On the feature "past tense," see §5.1.2.2.

There is an additional problem with examining the effects of bar position on arguments: the only way to tell that a PP argument is a daughter of V″ instead of V′ is if it follows an adjunct. There is no way of knowing if the marked interpretation is due to the structure or the ordering.

"X Y" means that liaison is permitted between X and Y, and "X/ Y" means that it is not.

The relative clauses (the b. cases in (77)-(79) and (80c)) are actually accounted for differently by Bresnan, because they involve deletion. The reader is referred to the article for details.
We are following Postal (1974) in treating *want* as a "Raising-to-Object" verb, despite the well-known problems with this analysis (such as inapplicability of passivization). These problems have led other theories, such as Government/Binding Theory, to treat the NP-to-VP following *want* as a normal infinitive in which *for* has been deleted or omitted. Whether we are correct or not is immaterial; what is important is that there is a configurationally represented empty category at the extraction site.

Some of the problems cited by Postal and Pullum for trace-the descriptions of *want-to* cliticization are inapplicable here, because the traces that have to be invisible to the rule, such as the traces of Raising and of COMP-to-COMP movement, do not exist in CRG.

An anaphor that is represented configurationally even though it has no phonetic content is the trace of syntactic binding constructions, because it has demonstrable phonological effects. See Jacobson (1982) for discussion.

Technically speaking, §2.5.4 does not exactly show that *c-command* is configurational. In one sense of the term *c-command*, this is impossible, as *c-command*, in its original sense, defines a particular configurational situation relating to the first branching node dominating a particular constituent, which may or may not be relevant for various aspects of linguistic theory, such as the height condition on anaphora. What is shown is that, given that there is a height relation X such that an anaphor may not be in relation X to an antecedent, this relation has to have both a configurational component and a relational one. I am following current usage in using the term *c-command* for this relation X, however it is defined. (There is nothing unusual in this; much of the recent literature within the Government/Binding framework is devoted to finding the correct definitions for such notions as *c-command*, *government*, and *proper government*.) It is, of course, possible that X is not a unitary relation, but we will assume that it is as a working hypothesis. Even if it is not, it is not clear that it would split discreetly into configurational and relational components. At any rate, if X is to be split, the argument that anaphora is configurational will not be effected. (Thanks to Jim Higginbotham for a discussion that led to this footnote.)

I do not mean to imply that I agree with the government approach to *PRO*; see Mohanan (1983b) and §3.7 of this dissertation for discussion. The point here is simply that within the GB system, and accepting the premises of the theory, the explanation for the distribution of *PRO* offered by government is intriguing. As it is, I do not accept the premises of the theory.

The crucial assumption here is that picture nouns take arguments realized as *POSS* and *OBL*. This is forced on us by the theory for two reasons. In the first place, by representing anaphora as a relation between arguments, anything that can be at either end of an anaphoric link must be an argument. Secondly, *POSS* and *OBL* are sometimes arguments, as in action nominals (cf. Rappaport 1983). The theory of grammatical relations is stronger if we do not allow a single relation to be both a Lexical Relation and a Syntactic Relation. The EST treatment of possessives as "subjects" is essentially equivalent to treating *POSS* as an argument.
In CRG and LFG, it is not possible to analyze point out to as a complex verb taking an NP OBJ. The crucial test is passivization, and point out to fails.

(1) *Rosa was pointed out to that most people stop at red lights.

I have a good deal of difficulty with Reinhart's backwards anaphora judgments; almost all the cases sound pretty bad to me. This is probably due to some constraint, either grammatical or pragmatic, relating to precedence. Examples such as those in (116), involving forward anaphora, are much clearer. Wherever possible, I will use sentences with preposed constituents.

Technically, we have not explained why tight PPs cannot appear on the left-hand side of pseudoclefts. In this case, we agree with the basic thrust of Reinhart's approach: tight PPs are too tightly bound to the verb to be separated in a pseudocleft. We disagree on the nature of the relation between a tight PP and a verb: Reinhart considers it configurational and we will suggest that tight PPs are arguments.

It is actually not quite this simple, even in TG. For instance, there are verbs which take optional arguments, such as eat, whose object is optional. The PP argument of find, which is one of the tight PPs, is not implausibly an optional argument in the classical TG sense.

Levin actually uses the verb genuflect, about which I have no intuitions.

Levin's sentence is actually one with there-insertion.

(1) *There cried hungry children in the street.

However, the judgment is the same for Oblique Inversion.
CHAPTER THREE: THE THEORY OF ANAPHORA

In the remainder of this dissertation, we develop a more formal and explicit theory of Configurational Grammar. In this chapter, we examine anaphora, building on the base established in the previous chapter. We propose a set of anaphoric features and discuss some areas where languages may differ.

3.1. Introduction

Anaphora plays a highly important role in CRG, essentially covering all processes traditionally considered transformational aside from the purely lexical ones, such as passivization. As explained in §2.5.1, anaphora in CRG is defined at the level of argument lists. This differs from transformational grammar, in which anaphora is taken to be a relation between two NPs, either pronominalization of one NP by another (Langacker 1967 and other transformational studies) or coindexing of an anaphoric NP with some other NP (as in the Extended Standard Theory and Government/Binding theory). It also differs from Relational Grammar, in which coreference is viewed as underlying multiple attachment of a single nominal (Johnson and Postal 1981). It also differs, for the most part, from Lexical-Functional Grammar, in which coreference is generally expressed between positions in f-structure, although Kisala (1981) analyzes certain cases of obligatory Equi in nominals as involving the indexing of argument positions. This difference is forced on us, as the argument list is the only place that phonologically null anaphors, such as PRO, are represented. The fact that this enables us to account elegantly for the same data that Kisala used to
motivate a major addition to LFG (see §3.7.3) is an added bonus.

By including in our account of anaphora some constructions that are not usually included, we are able to capture relations which have gone unnoticed previously. For example, while it is well known that PRO is pronoun-like (Postal 1970; Chomsky 1981, 1982), we show (in §§3.7.3 and 3.9) that it is like the gap of syntactic binding in being subject to a subset of the island constraints. This, in combination with the CRG translation of the analysis of the Fixed Subject Constraint proposed by Falk (1983b), leads us to a revisionist view of the restriction of PRO to nonfinite SUBJs in languages like English (end of §3.8).

Note that our notation, like Jackendoff (1972), distinguishes between antecedent and anaphor, unlike the standard coindexing. Higginbotham (1983) argues on the basis of several formal problems in the theory of anaphora (split antecedents, Chomsky's (1981) Binding Condition C, and others) that this is desirable.

We will begin with the most commonly studied anaphors: personal pronouns, reflexives, and reciprocals. We will not have much to say about personal pronouns per se, but we will use the discussion of them (§3.2) to introduce the formal machinery of the theory of anaphora. We distinguish four types of anaphors that are traditionally called reflexives, and discuss the properties of each class in §§3.3-3.6, with a digression in §3.5.2 on a special type of personal pronoun. We will also discuss (§3.5.3) the Relational Grammar claim that a theory of anaphora must allow reference to nonsurface grammatical relations. We then turn to raising and control phenomena (§3.7) and syntactic binding (§3.8). §3.9 presents a first approximation at the island constraints (and evidence that although Bresnan's (1982a) f-command condition is needed, the LFG definition is inadequate).
3.2. Personal Pronouns

The single most studied class of anaphors is personal pronouns (henceforth *pronouns*). Their properties are summarized in (1).

(1)  

a. Pronouns must agree in certain grammatical features with their antecedents.

b. Pronouns may not be superordinate (in some sense) to their antecedents.

c. Under some circumstances, pronouns may not precede their antecedents.

d. Pronouns may not appear in the same argument list as their antecedents.

e. Pronouns need not have antecedents.¹

Of these properties, a.–c. are general properties of anaphors while d.–e. are particular to pronouns.

Agreement is a common phenomena in many languages. In the case of anaphors, it is likely that some merger of anaphor and antecedent occur during the semantic interpretation process, and the relevant grammatical features may still be accessible. One could, for concreteness, propose a principle requiring agreement, but we will not do so.

Property (1b) is very important. We follow recent work in taking the appropriate sense of superordinate to be Reinhart's (1976, 1981) *c(onsituent)-command*, as altered in §2.5.4 of this dissertation.² Property (1c) is often thought to be related, because the earliest notion of superordinate was Langacker's (1967) *precede and command*. Reinhart, in proposing the notion c-command, argued that precedence plays no role in the theory of anaphora, where anaphora is taken to mean the grammatical determination of coreference. She claimed that apparent cases of prece-
dence restrictions which could not be captured by c-command were due to discourse constraints. However, it seems that precedence must be allowed to play a role even if we replace 'command' with 'c-command,' as we have done. Mohanan (1982b) shows that in the Dravidian language Malayalam pronouns must be preceded by their antecedents but reflexives need not be. A convincing case in English, involving reflexives rather than pronouns, is (2).

(2)  
a. I spoke to Max about himself.  
b. *I spoke about himself to Max.

The inability of pronouns to have antecedents in the same argument list (property (1d)) is an important defining property of pronouns. We use it to propose the first anaphoric feature: [±local]. We define the '−' value of the feature as (3), using the definition (4); we will return to the '+' value in the next section.

(3)  A [-local] anaphor may not have a coargument as an antecedent.

(4) **Definition**

For two arguments α and 3 in a syntactic structure, if α and 3 are in the same argument list, we say they are coarguments.

As an illustration, consider (5).

(5)  
a. Max expects that he will win. (he = Max, optionally)  
b. Max expects him to win. (him ≠ Max)

The anaphors in (5) are pronouns, which means they have the feature [-local]. This feature has to be appended to the argument position associated with each occurrence of the pronoun. To do this, we state the following.

(6) **Annotate the anaphoric features associated with a configurational anaphor to the argument position that the configurational anaphor fills.**
The term *configurational anaphor* refers to the configurational constituent *(he in (5a) and him in (5b)) that indicates which anaphoric features to annotate. It is the feature bundle, not the configurational anaphor itself, that is subject to anaphoric binding. Note that (6) does not actually have to be stated; it is a special case of the Argument Feature Convention ((64) of Chapter 2). If we invoke (6), the syntactic structures of (5) are as follows.

(7)  

a. 

\[
\begin{array}{c}
\text{Max expects that he will win} \\
\end{array}
\]

b. 

\[
\begin{array}{c}
\text{Max expects him to win} \\
\end{array}
\]

(We will discuss the nature of the SUBJ of *win* in §§3.7.2 (7b) and 3.7.3 (7a).) The anaphoric binding of the [-local] anaphor *him* to *Max* in (7b) is ungrammatical because they are coarguments.

The fact that pronouns need not have antecedents (property (5e)) has been taken by Chomsky (1981) to mean that pronouns have a capacity for
"inherent reference" which other anaphors lack. In the present framework it will be assumed that (le) is the norm. An anaphor will have an obligatory antecedent only if required by positive anaphoric features.

We impose an additional condition on anaphora, due to Jackendoff (1972).

(8) **Noncoreferentiality**

An anaphor and an argument position in a single structure may not be interpreted coreferentially if they are not grammatically marked coreferential.

The major purpose of (8) is to rule out structures where A is coreferential with B and B is coreferential with C but A and C cannot be grammatically marked coreferential. Such a case is (9), where $A=\text{he}$, $B=\text{David}$, and $C=\text{him}$.

(9) *The fact that he realized David was sick bothered him. (where $\text{he}$, $\text{David}$, and $\text{him}$ are all coreferential)

Finally, we assume the following rules.

(10) **Free Coreference Assignment** (Optional)

Assign an anaphor one or more antecedents consistent with its features.

(11) **Transitive Coreference Assignment**

For arguments $\alpha$, $\beta$, and $\gamma$, $\beta$ and $\gamma$ anaphors: if $\alpha$ is $\beta$'s antecedent and $\beta$ is $\gamma$'s antecedent, assign $\alpha$ as an antecedent of $\gamma$ iff $\gamma$ does not c-command $\alpha$.

(11) is needed because of Noncoreferentiality (8). Without it, a sentence like (12) could not be grammatical.

(12) Max said that he hates himself. (where $\text{Max}$, $\text{he}$, and $\text{himself}$ are all coreferential)

(10) would not allow $\text{Max}$ to be an antecedent for $\text{himself}$ because they are not coarguments (roughly; we will see in the next section that the condition on reflexives is slightly more complicated), but if they are not marked
coreferential, the sentence is ungrammatical on the desired reading.

Let us consider the ungrammatical (9) in more detail. The structure after the application of Free Coreference Assignment (10) is (13).

(13)

(13) represents the maximal application of Free Coreference Assignment (10). Applying it only once would result in one of the grammatical readings. (See Jackendoff 1972 for a discussion of the preferability of the two readings.) The $\alpha$, $\delta$, and $\gamma$ of Transitive Coreference Assignment (11) are labeled. In order for $he$ and David to be interpreted as coreferential, they have to be marked so in the syntax, but (11) specifies that $\gamma$ must not c-command $\alpha$.

In the coming sections, we will discuss other anaphors. The same principles that apply to pronouns (other than the definition of [-local]) apply to them as well.
3.3. R-anaphors

We turn now to reflexives and reciprocals, which have been given various names in the recent literature, including anaphors (Chomsky 1981) and bound anaphors (Bresnan 1982a). Cross-linguistically, this is a rather disparate group of anaphors, and we will not treat them as a single class. The various types are dealt with and distinguished in this and the following three sections. In this section, we will take a look at reciprocals and those reflexives which, like the ones in English, exhibit the same behavior as reciprocals. We call them r-anaphors.

R-anaphors snare properties (la-c): agreement, c-command, and precedence. They differ on (ld,e): r-anaphors can, and often must, be bound to coarguments, and they always must have antecedents. These two properties are expressed formally by the single feature [+local], which we define provisionally as (14).

(14) A [+local] anaphor must have an antecedent in its Local Binding Domain.

What is a Local Binding Domain? In §2.5.3, we showed that the Local Binding Domain of the anaphor consists of more than just its coarguments. Thus, there is a lack of symmetry between the definitions of [+local] and [-local]. We repeat the sentences in (95) of Chapter 2 here.

(15) a. Max told Bill a story about himself. (himself = Max or Bill)

b. Tom told Dick Harry's story about himself. (himself = Harry; himself ≠ Tom or Dick)

c. That the picture of himself in the post office is ugly enrages Charlie. (himself = Charlie).

Note what happens if we substitute him (a pronoun: [-local]) for himself (an r-anaphor: [+local]) in (15).
(16)  

a. Max told Bill a story about him. \((him = Max \text{ or } Bill, \text{ optionally})\)

b. Tom told Dick Harry's story about him. \((him = Tom \text{ or } Dick, \text{ optionally})\)

c. That the picture of him in the post office is ugly enrages Charlie. \((him = Charlie, \text{ optionally})\)

The pronoun seems to be obviative only within the argument list in which it is located, supporting our definition of \([-\text{local}].\)

The idea that we put forward in §2.5.3 is that an r-anaphor can look for its antecedent outside of its own argument list is there is no possible antecedent in it. The intuition is similar to that behind Chomsky's (1981) notion of "accessible SUBJECT" determining the governing category. One way in which our account differs from Chomsky's is that he uses the SUBJECT-based domain to define both the domain in which r-anaphors are bound and the one in which pronouns are free. As we see in these English examples, however, it is not the case that pronouns and r-anaphors are in complementary distribution. One problem with our noncoargument restriction on the antecedents of pronouns is that it is not quite strong enough. Chomsky cites sentences such as (17).

(17) Max heard a story about him. \((him \neq Max)\)

The contrast between (16a) and (17) is particularly puzzling, and neither Chomsky nor I can account for it.

Although the Local Binding Domain is not identical to the argument list containing the anaphor, the anaphor's coarguments do form an essential part of the Domain. If we examine r-anaphors cross-linguistically, we find that all r-anaphors must be bound somewhere and can be bound to coarguments. Where languages differ is in what in addition to the anaphor's argument list can be included in the Domain. One subtle distinction cross-linguistically in languages that show a pattern very similar to that seen in
English relates to the behavior of anaphors which are SUBJs or arguments of SUBJs. (These examples are from Yang 1983.)

(18) **English**

a. *The boys thought that each other was smart.*

b. *The boys thought that each other's pictures were on sale.*

(19) **Dutch**

??Zij denken dat elkaars verhalen interessant waren.
they think that each other's stories interesting are
'They think that each other's stories are interesting.'

(20) **Italian**

?? Sanno che certi libri l'uno a proposito dell'altro
(they) know that some books one about of-the-other
sono interessanti.
are interesting
'They know that some books about each other are interesting.'

(21) **Russian**

*Dissidenty znali chto stat'! drug o druge pojavilis'*
dissidents knew that articles each about other appeared
v zapodnoj presse.
in western press
'The dissidents knew that articles about each other had appeared
in the western press.'

(22) **Norwegian**

a. *De skryter over at hverandre er intelligente.*
they boast over that each other is intelligent.
'They boast that each other is intelligent.'

b. *Marxistene visste at artikler om hverandre hadde*
Marxists knew that articles about each other had
statt i vestlige auiser.
been in western newspaper
'The Marxists knew that articles about each other had
appeared in western newspapers.'
In English, an r-anaphor may not occupy SUBJ position: this is because SUBJ position ([NP,S]) c-commands the SUBJ's coarguments and, for some reason, if the r-anaphor is a SUBJ the antecedent must be within the clause. We will hypothesize that the inability of the binding domain to extend in this case is a more general prohibition of all arguments of Verbs and Modals; r-anaphors which are arguments of nouns, adjectives, and
prepositions can look out of their argument lists if there is no potential coargument antecedent. (Note that V and M form a natural class in terms of the category distinctive features proposed in the Appendix: [+PGR, +OA].) Russian and Norwegian, on the other hand, require the antecedent of an r-anaphor to be a clause-mate (although not necessarily a coargument); that is, they must be in the same minimal S. The judgments shown in (19) and (20) place Italian and Dutch as intermediate languages, a status that is difficult to formalize. However, according to Rizzi (personal communication), many speakers of Dutch find (19) completely ungrammatical and he personally considers (20) unsalvageable. This would make Dutch and Italian just like Russian and Norwegian. Finally, Korean, Japanese, Chinese, and the Dravidian language Kannada are less restrictive than English: they allow SUBJ r-anaphors to find their antecedents outside the clause.

The way we will account for this cross-linguistic variation is to assume that there is a core notion of Local Binding Domain which is universally identical to the set of the r-anaphor's coarguments. There will then be a language-specific extension of this Domain. We also redefine the realization of [+local] anaphors to make reference to Extended Local Binding Domain; this will be further refined in §3.8. In the definitions of Extended Local Binding Domains, we make use of the notion d-command. Intuitively, an argument is d-commanded by all argument lists which "dominate" the list containing the argument. We also add the definition of the Malayalam Extended Local Binding Domain, discussed in §2.5.3, repeating the definition of NP-contain.
(27) **Definitions**

a. For argument lists $\alpha$ and $\gamma$, $\gamma$ heading a phrase filling an argument position for some lexical item $A$, and $\beta$ an argument position in $\gamma$, $\alpha$ *d-commands* $\beta$ iff $\alpha$ constituent which corresponds to an argument in $\alpha$ dominates $A$.

b. $\alpha$ *NP-contains* $\beta$ iff
   (i) $\alpha$ is $\beta$ or
   (ii) $\alpha$ dominates $\beta$ with no intervening non-NP nodes.

c. For an anaphor $\alpha$, the *Local Binding Domain (LBD)* of $\alpha$ consists of $\alpha$'s coarguments.

d. For an anaphor $\alpha$, the *Extended Local Binding Domain (ELBD)* of $\alpha$ consists of $\alpha$'s LBD plus
   (i) (English) the minimal d-commanding argument list containing a referential argument if there is none in the LBD and $\alpha$ is not an argument of a lexical item of category [+PGR, +OA] (≠V or M);
   (ii) (Russian) the minimal d-commanding argument list heading an S or containing a referential argument if there is none in the LBD and $\alpha$ is not an argument of a lexical item of category [+PGR, +OA];
   (iii) (Korean) the minimal d-commanding argument list containing a referential argument if there is none in the LBD;
   (iv) (Malayalam) all d-commanding argument lists if $\alpha$ is NP-contained by a constituent bearing the grammatical relation SUBJ; otherwise the ELBD is identical to the LBD.

(28) A [+local] anaphor must have an antecedent in its Extended Local Binding Domain.

To determine the universal component and parameters of the definition of ELBD would take us too far afield, but (27d) shows the extent to which variation is permitted.

3.4. Pseudo-Reflexives

All reciprocals seem to conform to the general outline presented in the previous section; however, this is not true of reflexives. In many languages, the binding domain for reflexives is larger than that predicted in the previous section. In all cases considered there, the ELBD extended
beyond the anaphor's coarguments only when necessary to provide a possible antecedent. However, some reflexives have larger binding domains even when not required for an antecedent. For instance, the Icelandic reflexive \( \text{sig} \) can have an antecedent anywhere in the minimal indicative clause containing the anaphor. The Russian reflexive \( \text{sebja} \) can be bound anywhere in the minimal finite clause dominating it. In these languages, unlike English, reciprocals and reflexives behave differently. Consider the following examples in Icelandic, from Thráinsson (1979).

(29)  
\begin{enumerate}
  \item a. Jón skipaði méð að raka sig.
    \textit{ordered me to shave self}
    \textit{\text{'}John ordered me to shave him\text{(self)}\text{'}.
  
  \item b. *Þeir skipuðu méð að raka hvor annan.
    \textit{they ordered me to shave each other}
    \textit{\text{'}They ordered me to shave each other\text{'}.
\end{enumerate}

(30)  
\begin{enumerate}
  \item a. Jón telur mig hafa skivið sig.
    \textit{believes me (to-)have betrayed self}
    \textit{\text{'}John believes me to have betrayed him\text{(self)}\text{'}.
  
  \item b. *Menninir töldu mig hata hvor annan.
    \textit{the-men believe me (to-)hate each other}
    \textit{\text{'}The men believe me to hate each other\text{'}.
\end{enumerate}

What contrasts such as these show is that these reflexives do not simply involve additional definitions of ELBD, because if they did then reciprocals would exhibit the same properties as reflexives. In addition, there are languages with two types of reflexives: one that is bound in the ELBD and one that behaves like Icelandic and Russian reflexives: Dutch (r-anaphor \( \text{zieh} \) \text{\textit{self}}, reflexive \( \text{zieh} \)), Italian (r-anaphor \( \text{stesso} \), reflexive \( \text{esso} \)), Norwegian (r-anaphors \( \text{selv} \) and pronoun + \( \text{selv} \), reflexive \( \text{selv} \)), German (r-anaphor \( \text{selbst} \), reflexive \( \text{selbst} \)), and Swedish (r-anaphor \( \text{själv} \), reflexive \( \text{själv} \)) are the ones listed by \( \text{Ian\text{"}{\textsc{g}}} \) (1983), who analyzes these data from the point of view of Government/Binding theory, and whose data, collected from various sources, will be used
extensively here.

Yang documents four different definitions of the domain for the reflexives in question: minimal d-commanding argument list associated with a finite verb (Norwegian, Russian, Hindi, Gothic); minimal d-commanding argument list associated with a finite verb or deverbal noun (Swedish, Polish); minimal d-commanding argument list associated with an indicative verb (Icelandic, Italian); and, in the terminology to be introduced in §4.2.2, the minimal d-commanding argument list associated with the functional head of $\bar{s}$. The first three of these are self-explanatory, and are illustrated by Yang by the following examples.

(31) **Finite Verb**

a. **Russian:**

(i)  
Vanja znaet chto Volodja ochen’ ljubit sebja.  
knows that verymuch loves self  
'Vanja knows that Volodja loves self very much.'  
(sebja = Volodja; sebja ≠ Vanja)

(ii)  
Professor proposil assistenta chitat’ svoj doklad.  
professor asked assistant (to-)read self’s report  
'The professor asked his assistant to read self’s report.'  
(svoj = assistenta or professor)

b. **Hindi:**

(i)  
Ashok ne kəha kii Lalita ṭpne liye cha kəreegi.  
said that self for tea make  
'Ashok said that Lalita would make tea for self.'  
(ṭpne = Lalita; ṭpne ≠ Ashok)

(ii)  
Ashok ne Lalita se ṭpne liye cay b nane ko kəha.  
with self for tea to make asked  
'Ashok asked Lalita to make some tea for self.'  
(ṭpne = Ashok or Lalita)
Finite Verb or Deverbal Noun

a. Polish:

(i) Jan kazat Marie napisać artykuł o sobie.
    told (to-)write article about self
    'Jan told Marie to write an article about self.'
    (sobie = Jan or Marie)

(ii) Jan czyta jej książkę o sobie.
     read her book about self
    'Jan is reading her book about self.'
    (sobie = Jan or jej)

(iii) Maria nie rozumie jego naglej niechęci do siebie.
     not understand his sudden resentment to self
    'Maria does not understand his sudden resentment to self.'
    (siebie = jego; siebie ≠ Maria)

b. Swedish:

(i) Han bad honom klippa sig.
    he asked him (to-)cut(hair) self
    'He asked him to cut self’s hair.'
    (sig = honom or han)

(ii) Han tillät hans offentliggörande av sina privatbrev.
     he allowed his publication of self’s private letters
    'He allowed his publication of self’s private letters.'
    (sina = hans; sina ≠ han)

Indicative Verb

a. Icelandic:

(i) Jón skipaði Haroldi að raka sig
    ordered to shave self
    'John ordered Harold to shave himself.'
    (sig = Haroldi or Jón)

(ii) Jón segir að María viti að Haroldur
     says indic that knows subjunct that
     vilji að Billi meðaði sig.
     wants subjunct that hurts subjunct self
    'John says that Mary knows that Harold wants Bill to hurt self.'
    (sig = Billi, Haroldur, María, or Jón)
The case of Dutch is a little less straightforward. We base our analysis on Yang (1983), who cites the following pair.4

(34)  a. Jan liet Karel over zich praten.
    let about self (to-)talk
    'Jan let Karel talk about him.' (zich = Jan)

     b. *Jan vroeg Karel over zich te praten.
       asked about self to talk
       'Jan asked Karel to talk about him.' (zich ≠ Jan)

For reasons we will return to, zich cannot be coreferential with Karel in either sentence. The key point is the contrast in whether or not Jan is within the domain of zich. The key point is apparently the presence of the infinitival marker te apparently blocking coreference in (34b).

Yang, on the basis of these and some other examples, has this to say about Dutch zich:

"the conditions for ... zich are quite complex, but as the contrast between [(34a)] and [(34b)] clearly shows the COMP plays a major role in defining the binding domain for zich." [p. 19]

It is unclear how a GB-based analysis could consider te an element of the complementizer. As van Riemsdijk (1978) shows in surveying certain aspects of Dutch word order, the te+verb combination can never be broken up, not even by particles ("separable prefixes") which can be incorporated
into the verb. Thus, the most straightforward analysis of *te* would be to treat it as a part of the verb itself, a kind of inflectional element. However, the theory of categories proposed in Chapter 4 can treat infinitives as the heads of *S*, as we show in §4.2.2. Assuming the analysis there, our characterization of the domain of *zich* stands.

How can we characterize these reflexives formally? The fact that they are not r-anaphors means that we cannot use the same feature [*±local*] that we used to distinguish r-anaphors from pronouns. Nevertheless, the condition on these reflexives is also a locality condition of a sort: the anaphor has to be bound in the minimal d-commanding argument list associated with a certain type of predicate or phrase. We call this feature [*±local*] to emphasize the fact that intuitively this is a second locality feature; formally, however, it is a distinct feature. Because each definition of the domain for these reflexives describes some idea of a clause in the informal sense, we will refer to the binding domain as a pseudo-clause; anaphors bearing the feature [+local'] will be called pseudo-reflexives. As with the feature [-local], there is no evidence that [-local'] is obviative in the same domain as [+local'] anaphors are bound. That is to say, there is no evidence of "pseudo-pronouns" which are free in the minimal pseudo-clause. We will assume that [-local] and [-local'] have the same interpretation.

\[(35)\]
\[\begin{align*}
a. & \quad \text{A [+local']} \text{ anaphor must have an antecedent in its minimal pseudo-clause.} \\
b. & \quad \text{A [-local']} \text{ anaphor may not have a coargument as an antecedent.}
\end{align*}\]

As separate features, it may be asked how [+local] and [-local'] interact. The definitions of the '+' and '-' values for the feature [*±local*] are repeated as (36).
A [+local] anaphor must have an antecedent in its Extended Local Binding Domain.

b. A [-local] anaphor may not have a coargument as an antecedent.

The combinations in (37) are thus predicted

(37) a. \([-\text{local}, -\text{local'}\)]: like either of these features individually, an anaphor bearing these features together may not have a coargument antecedent.

b. \([+\text{local}, +\text{local'}\)]: an anaphor bearing these features must be bound in the more restrictive of its ELBD and its minimal pseudo-clause. This will almost always be the ELBD.

c. \([-\text{local}, +\text{local'}\)]: an anaphor with these features is bound to an antecedent in its minimal pseudo-clause, but not to a coargument.

d. \([+\text{local}, -\text{local'}\)]: an anaphor with these features has an antecedent in its ELBD, but not a coargument.

In addition, either of these features may be left unspecified (i.e. ambiguous between '+' and '-'). As we can see from (35), (36), and (37), \([-\text{local}, -\text{local'}\) and \([-\text{local}, -\text{local'}\) are equivalent, and describe an anaphor which is obviative within its argument list. Similarly, \([+\text{local}]\) and \([+\text{local}, +\text{local'}\) are both specifications for an anaphor which is bound in the ELBD. Thus, there are five possible combinations of features. (37a) and its synonymous single-feature bundles identify ordinary pronouns, (37b) and its synonym identify r-anaphors, and \([+\text{local'}\) identifies pseudo-reflexives. The crucial question is whether or not (37c,d) are realized. (37d) is not, and it is not hard to see why. Such an anaphor would have a very limited use, essentially appearing only as arguments to single-argument nominals. Whether this combination is ruled out by some principle of UG or just never occurs is immaterial.
(37c), on the other hand, is realized. Yang cites (38) in Norwegian.

(38)   *Ola korrigerer seg
        corrects self 'Ola corrects himself.'

Seg may not be bound to a coargument. Dutch zich also seems to be
[-local], although the data are not as clear. Consider the following
from Everaert (1980).

(39)   a. Jan houdt van zichzelf/*zich.
       loves of 'Jan loves himself.'

       b. Jan vindt zichzelf/*zich het geloofwaardigst.
          thinks the most-reliable
          'Jan thinks himself the most reliable.'

       c. Jan zag zichzelf/*zich Karel wassen.
          saw (to-)wash
          'Jan saw himself wash Karel.'

(40)   a. Jan wast zichzelf/zich.
       washes 'Jan washes himself.'

       b. Jan droogde zichzelf/zich af
          rubs down 'Jan rubs himself down.'

The data in (39) suggest that zich is [-local], while those in (40)
suggest that it is not. We claim that zich is [-local], but the zich
in (40) is not the anaphor; instead it is a signal that lexical reflexivi-
ization (Grimshaw 1982b) has taken place. We will discuss lexical
reflexivization briefly in §3.6, and answer an objection raised by
Everaert to an analysis of this type. Note that the existence of [-local]
pseudo-reflexives provides evidence for treating pseudo-reflexives as
distinct from r-anaphors, and for a feature analysis.
3.5. Subjecthood of Antecedents

3.5.1. Subjective Anaphors

Many languages specify that certain anaphors must or must not have antecedents which are SUBJs. These anaphors are often translated into English as reflexives, even though they may not be subject to any sort of locality principle. The theoretical importance of the existence of such anaphors was discussed in §2.5.2; here we deal with specific analyses.

The clearest case of anaphors making reference to the subjecthood of their antecedents comes from Yoruba. As described by Mohanan (1982b), Yoruba has two anaphors: òun and ò. The former must be bound to a SUBJ, while the latter must either be bound to a nonSUBJ or not bound at all.

\[(41)\]
\[
\begin{align*}
\text{(a)} & \quad \text{Tolú sòfùn Sègun pè òun sanra.} \\
& \quad \text{told that is-fat} \\
& \quad '\text{Tolu told Segun that he is fat.' (òun \# Sègun; òun = Tolú, obligatorily)} \\
\text{(b)} & \quad \text{Tolú sòfùn Sègun pè ó sanra.} '\text{Tolu told Segun that he is fat.' (ó \# Tolú; ó = Sègun, optionally)
\end{align*}
\]

\[(42)\]
\[
\begin{align*}
\text{(a)} & \quad \text{Adé rò pè Tolú sòfùn Sègun pè òun sanra.} \\
& \quad \text{thought that} \\
& \quad '\text{Ade thought that Tolu told Segun that he is fat.'} \\
& \quad (òun \# Sègun; òun = either Adé or Tolú, obligatorily) \\
\text{(b)} & \quad \text{Adé rò pè Tolú sòfùn Sègun pè ó sanra.} \\
& \quad '\text{Ade thought that Tolu told Segun that he is fat.'} \\
& \quad (ó \# Adé or Tolú; ó = Sègun, optionally)
\end{align*}
\]

We posit a feature [+subj].

\[(43)\]
\[
\begin{align*}
\text{(a)} & \quad \text{A [+subj] anaphor must have an antecedent which is a SUBJ.} \\
\text{(b)} & \quad \text{A [-subj] anaphor must not have an antecedent which is a SUBJ.}
\end{align*}
\]

We will call an anaphor with the value '+' for this feature subjective and one with the value '-' unsujective. It should be clear that the
Yoruba facts are adequately covered by this single feature. Another language with a clear distinction between subjective and nonsubjective anaphors is Norwegian. Norwegian has two r-anaphors: seg selv and pronoun + selv, the former subjective and the latter nonsubjective, as illustrated by the following data from Fang (1983).

```
(44) a. Han fortalte Knut om seg selv.
    told about himself
    'He told Knut about himself.' (seg selv = han; seg selv ≠ Knut)

b. Han ba Knut snakke med Ola om ham selv.
    he asked (to-)talk with about himself
    'He asked Knut to talk with Ola about himself.'
    (ham selv = Ola; ham selv ≠ Knut)
```

Many languages have anaphors which are subjective but none which are exclusively nonsubjective. For instance, Malayalam has three types of anaphors: pronouns (awan 'he', etc.), subjective pronouns (taan, etc.) and subjective r-anaphors (swa-) (Mohanan 1982b). In the present theory, the anaphors are analyzed into features as follows.

```
(45) awan: [-local]

    taan: [-local] [+subj]

    swa-: [+local] [+subj]
```

```
(46) a. Moohan awane aafaadhir'k'unnu.
    him worships
    'Mohan worships him.' (awan ≠ Moohan)

b. *Moohan tanne aafaadhir'k'unnu.
    'Mohan worships him.' (tanne ≠ Moohan)

c. Moohan swayam aafaadhir'k'unnu.
    'Mohan worships himself.' (swayam = Moohan)
```
(47) a. Kutti tante aniyattiye nulli
    child nom its sister acc pinched
    'The child pinched its sister.' (tante = kutti)

b. *Kuttiye tante aniyatti nulli.
    child acc sister nom
    'Its sister pinched the child.' (tante ≠ kuttiye)

(48) _taan aanaye nulli enu3 raajaaw3 mantriyoott3 parannu.
    elephant pinched that king nom minister dat said
    'The king told the minister that he had pinched the elephant.'
    (taan = raajaaw3; taan ≠ mantriyoott3)

In many languages, the pseudo-reflexives are subjective. We propose
without argument that there is a universal convention (49).

(49) If an anaphor is [+local'], in the unmarked case it will also
be [+subj].

Languages need not incorporate (49), but those that do are more highly·
valued by the evaluation metric.

3.5.2. Nonsubjective Pronouns

Let us now consider the interaction of the locality features with
the subjectivity feature. By simply taking our present definitions of
these features, the various combinations result in the following properties.

(50) a. [+local()] [+subj] : bound to a SUBJ in the ELBD or minimal
    pseudo-clause.

b. [+local()] [-subj] : bound to a nonSUBJ in the ELBD or minimal
    pseudo-clause.

c. [-local()] [+subj] : bound to a noncoargument SUBJ.

d. [-local()] [-subj] : not bound to a coargument and not bound to
    any SUBJ.

(50a) is very common, especially with the feature [+local']. (50b) is
rare; one realization of it is the Norwegian r-anaphor 'pronoun + selv'.
Presumably, the explanation for this is to be found in a theory of markedness. I am unaware of any realizations of (50c), but one would presume it to be as rare as (50b), so the lack of any examples can be taken as an accidental gap. What is disturbing is (50d), for two reasons. In the first place, if (50c) is as rare as (50b), then (50d) should be roughly as common as (50a). In the second place, (50d) has a very unlikely ring to it. What we will propose here is that (50d) is very common, but that it translates into a special property rather than the one derived from combining the meanings of the individual features.

In many languages, the pronouns (i.e. anaphors with the feature [-local] and/or [-local']) are SUBJ-obviative in some domain. Consider the following examples from Yang (1983).

(51) **Norwegian**

a. De liker deres bøker.
   they like their books 'They like their books.' (deres ≠ de )

b. De leste mine klager mot dem.
   they read my complaints against them
   'They read my complaints against them.' (dem = de)

c. De leste klager mot dem.
   they read complaints against them
   'They read complaints against them.' (dem ≠ de)

(52) **Swedish**

a. Han gav honom hans motocylkel.
   he gave him his motorcycle
   'He gave him his motorcycle.' (hans = honom; hans ≠ han)

b. Hon bad henne klippa henne.
   she asked her (to)cut(hair) her
   'She asked her to cut her hair.' (second henne = hon;
   second 'hennie ≠ first henne)

c. hans beskrivning av honom
   his description of him
   'his description of him' (honom ≠ hans)
In these languages, the pronouns may not be coreferential with the closest SUBJ (or POSS). A different pattern is displayed by the following.

(53) **Polish**

a. Maria rozmawia\_\_ z Ani\_\_ o niej.  
   talked to about her  
   'Mary talked to Anna about her.'  (*niej = Ania; niej ≠ Maria*)

b. Maria martwi sie ich stosunkiem do niej.  
   worries over their attitude to her  
   'Mary worries about their attitude toward her.'  (*niej = Maria*)

(54) **Icelandic**

a. J\_\_n retti Haroldi hans f\_\_t.  
   handed Harold his clothes  
   'John handed Harold his clothes.'  (*hans ≠ Haroldi; hans ≠ J\_\_)  

b. J\_\_n skipa\_\_i mer a\_\_ raka hann.  
   ordered me to shave him  
   'John ordered me to shave him.'  (*hann ≠ J\_\*)

c. J\_\_n segir a\_\_ Maria elski hann.  
   says that loves him  
   'John says that Mary loves him.'  (*hann ≠ J\_\*)

(55) **Hindi**

a. Ashok ne k\_\_ha kii Lalita unke liye cha k\_\_reegi.  
   said that for tea would-make  
   'Ashok said that Lalita would make tea for him.'  (*unke ≠ Ashok; unke ≠ Lalita*)

b. Sita raste bh\_\_r unke g\_\_hne girati g\_\_i.  
   way all her jewelry dropping go  
   'Sita kept dropping her jewelry all the way.'  (*unke ≠ Sita*)

(56) **Latin**

qui noluerunt me regnare super e\_\_  
who not-wanted me (to-)rule over them  
'who did not want me to rule over them'  (*eos ≠ qui*)

In these languages, the pronoun may not be bound to any SUBJ or POSS up to the closest SUBJ of a finite verb. These pronouns do not have a clear feature decomposition in terms of already defined features. We propose that the combination \([-\text{local}'\), \(-\text{subj}\) receives a special inter-
pretation to cover these cases.

(57) An anaphor bearing the features \([-\text{local}(\text{'})], -\text{subj}\)] may not have as its antecedent a \text{SUBJ} or \text{POSS} up to the closest \text{SUBJ} or \text{POSS <of a finite verb>.

The phrase in angle brackets may or may not be present in a particular language.

3.5.3. Nonsurface Subjects

One of the central claims of a theory like LF geometrical (syntactic and semantic) processes are dependent on surface grammatical configurations and surface grammatical relations, as well as semantic information (such as thematic relations). Another claim is that there are no other types of information (aside from morphological information, such as case marking) which are necessary for the smooth operation of the syntax and the semantics. Much work in LF, as well as the bulk of this dissertation, is aimed at the first of these claims. If the conclusions of these structures are correct, then a theory which does not recognize both types of representation cannot be the "correct" theory of the mental representation of language. However, the second claim needs to be defended as well, for if it can be shown that there are other types of representation that are equally necessary, then CRG must adopt them or be abandoned.

Such claims have been made for nonsurface grammatical relations in the literature of Relational Grammar. Perlmutter (1982), concentrating on the notion "subject," surveys the RG literature and summarizes the arguments. From the point of view of CRG, the constructions discussed by Perlmutter and his colleagues fall into two groups: lexical processes and anaphoric processes. Of these, lexical processes are not directly
relevant here. The reason for this is that, although a reanalysis of the data which does not involve nonsurface grammatical relations may be possible for some of these constructions, it does not violate the basic architecture of CRG to concede to RG on this point. Suppose, for instance, that Lawler's (1977) claim that Achenese verbs agree with the "underlying subjects" is correct. The affixation of verbal agreement morphemes is, under a lexicalist theory of morphology (Lieber 1980, Kiparsky 1982, Bresnan 1982b) a lexical process. Under a level-ordered or stratified theory of the lexicon, lexical processes are partially ordered by being assigned one or more "levels" or "strata" as domains of application (Kiparsky 1982, Mohanan 1982a, Falk 1983c). It is certainly possible that in Achenese verbal agreement morphology is associated with an earlier level than relation-changing rules.

On the other hand, if it can be shown conclusively that there is an important generalization relating to anaphora which can only be captured by reference to nonsurface grammatical relations, that would constitute a powerful argument against the CRG approach. Thus, from the perspective of this dissertation, it is vitally important to address the issues raised in the RG literature. To ignore them would be to concede defeat.

In the realm of reflexives, facts in Cebuano, Georgian, Russian, and Japanese are claimed to support the existence of multiple strata of subjects in the syntax. In the case of Cebuano and Georgian, the antecedent is an "initial 1" (i.e. underlying subject); in the case of Russian, a nominal heading a 1 (subject) arc at any level (henceforth a "global 1"), and in the case of Japanese a "working 1" (a nominal heading a 1 arc at any level which also bears a Pure GR at the surface). There is an additional language mentioned, Choctaw, which also refers to
the notion "global 1," but the reflexive in question is a verbal affix, and could be a case of lexical reflexivization.

The alternative to a description of the facts of the languages in question in terms of nonsurface grammatical relations is a semantic description, or one involving both (surface) syntax and semantics. It is here that the weakness of the RG literature on the issue is apparent, because the RG analysis is usually contrasted with a nonspecific unformalized semantic theory. This leaves open the possibility that the alleged impossibility of accounting for the facts on the basis of semantics is due to the lack of an appropriate semantic formalism. A notable exception in this regard is Davies (1981); appearing as it does in a volume on Dik's theory of Functional Grammar, Davies's paper attacks the Functional Grammar account of semantic relations. Perlmutter (1981), in the same volume, shares this virtue of Davies's paper to a point, but at one point he states (p. 330)

FG's failure to account for Achenese Verb Agreement is not due solely to the typology of the four types of predication and the definitions of their associated semantic functions. Even if one were to adopt a different class of semantic functions..., a formulation in terms of these semantic notions would still fail to capture the generalization. The correct generalization requires reference to the notion 'initial 1', and that requires a theory that recognizes grammatical relations at more than one level.

But, of course, asserting that no other semantic theory could possibly capture the equivalent of RG's "initial 1" does not make it so. We will assume here a version of the cognitive semantic theory of Jackendoff (1983), henceforth CST.

The Cebuano facts are reviewed by Bell (1982). There are two sometimes distinct nominals which are traditionally called "subject" in Cebuano and other Phillipine languages. Bell refers to these as the actor and the nominative nominal, and argues that the actor is the initial 1 and the
nominative nominal the final 1. In sentences where the verb is inflected for active voice, the actor and the nominative nominal are the same. In the various other voices, the actor is marked with genitive case and one of the other nominals is the final 1, SUBJ in CRG terms. The crucial point is that the actor is always the antecedent of a reflexive (r-anaphor) and it cannot itself be one.

(In 58), the final 1 (or SUBJ) is bracketed. ACT and OBJ are the active and objective voices, respectively.) It is easy to see that it is not surface grammatical relations that determine the distribution of grammaticality judgments. What evidence is there that the actor cannot be described in semantic terms? Without making reference to a specific theory of semantic roles, Bell simply states that the actor need not be an Agent on the basis of (59a), and that it can even be nonconcrete, as in (59b).

The thematic relation Agent, which is the usual relation borne by the
Cebuano actor, is the highest relation on Jackendoff's (1972) Thematic Hierarchy. This suggests the following account of Cebuano r-anaphors.

(60) In Cebuano, the antecedent of an r-anaphor must be the highest argument on the Thematic Hierarchy available in the clause.

Unfortunately, (59a) seems not to support (60). In (59a), the arguments are a Goal (Fred), a Theme (libro 'book'), and a Source (Tomas). The Goal and Source appear at the same level in Jackendoff's hierarchy, and both outrank Theme; yet here Goal seems to outrank Source. This may be a necessary change in the hierarchy, anyway, because the English verb receive can be passivized, and according to Jackendoff (1972), the NP in the by phrase has to outrank the SUBJ in a passive. (59b) is more problematic, as the Theme, which is the lowest position on the hierarchy, is the actor. There is an alternative, involving a slight extension of CST. Under this analysis, the correct restriction is (61).

(61) In Cebuano, the leftmost argument in the semantic representation is the antecedent of an r-anaphor.

The representation of an agentive sentence is (62), according to CST.

(62) a. $\text{[Event \text{CAUSE } ([\text{Thing } X], [\text{Event } Y])]}

b. $\text{[Event } X_1, [\text{Action } \text{CAUSE } (i, [\text{Event } Y])]}

These are two expressions in the mental code which are linked by the equivalence rule (63).

(63) $\text{[Event } i(X_1, X_2, X_3, \ldots, X_n) \leftrightarrow [\text{Event } X_1, [\text{Action } F(i, X_2, X_3, \ldots, X_n)]]

What we would like to suggest is that there is a similar "conceptual reanalysis" in the case of possessives.

(64) $F_{\text{Poss}}(X_i, [G(Y_j)]) \leftrightarrow [\text{POSSSESSOR } Y_j], [F_{\text{Poss}}(X_i, [G(Y_j)])]$
That is to say, possessional Locations, Sources, and Goals can be "topicalized" in semantic representation. The current view of Possessor as simply a Location (or Source or Goal) does not explain why Possessors are often realized grammatically as SUBJs. If this sort of representation is correct, then (61) can be adopted as the rule for Cebuano. We will not discuss this further, as the motivation for this alteration of CST, and for CST itself, lie outside the scope of this dissertation. The point is that an appropriately formalized semantic theory can account for facts that seem not to be describable in terms of informal semantics. Furthermore, referring to the appropriate collection of semantic roles as "initial 1" does not explain anything; it simply restates the problem. In much of the RG literature, it is supposed that there is a universal mapping between semantic roles and initial grammatical relations (e.g. Perlmutter 1978). If this is true, then the rules effecting this mapping must make reference to exactly the same set of semantic roles as a rule making direct reference to semantic roles for, say, anaphora, would. By allowing initial grammatical relations to enter the picture, a route is provided for evading the question of the representation of semantic roles and the trivialization of semantic rules. This is a rather strange position for some of the pioneers of the theory of Generative Semantics to find themselves in. Note that even if there is not a completely universal mapping (Perlmutter 1982), there has to be an unmarked mapping that basically matches the proposed universal mapping.

Perlmutter (1982) argues that Russian pseudo-reflexives must have "global 1" antecedents. He first presents evidence that SUBJs and no other grammatical relations can antecede pseudo-reflexives.
(65)  a. Ja rasskazal Borisu anekdot o sebe. I told Boris joke about self 'I told Boris a joke about myself.' (sebe = ja, obligatorily)

b. Poetomu ja otošel ot temy iskusstva i stal therefore I left from topic art gen and began rasspraviw' xudožnika o nem/*sebe samom (to-)question artist acc about him/*self EMPH 'Therefore I left the topic of art and began to question the artist about himself.'

c. Mne bylo skazano obo mne/*o sebe. me dat was told about me /*about self 'To me it was told about myself.'

(66)  a. Anna otpravila rebenka k svoim roditeljam. sent child acc to self's parents. 'Anna sent the child to her parents.'

b. Rebenok byl otpravljen k svoim roditeljam. child nom was sent to self's parents. 'The child was sent to his parents.'

(65a) shows that only the SUBJ can be the antecedent of the pseudo-reflexive. (65b,c) are examples where English uses an r-anaphor but Russian uses a pronoun because the antecedents are not SUBJs: note that (65c) has no overt SUBJ at all. (66) illustrates a contrast due to passivization in the b. sentence 'child' can antecede the pseudo-reflexive but it cannot in a. This may lead us to assign the Russian pseudo-reflexive the feature decomposition [+Local , +subj]. However, (67) shows that this is not correct.

(67)  a. Èta kniga byla kuplena Borisom dlja sebja. this book was bought Boris instr for self 'This book was bought by Boris for himself.'

b. Mne nužno bolee udobnaja kvartira ne dlja sebja me dat needs more comfortable apartment not for self a dlja svoej sem'i. but for self's family 'I need a more comfortable apartment, not for my self, but for my family.' (sebja/svoej = mne)
In (67a), the "logical subject" (the Agent), which surfaces as an oblique argument bearing instrumental case, antecedes the reflexive. (67b) represents a construction known in the Generative Semantics literature as "Flip" or "Psych Movement" and in the RG literature as "Inversion." It involves a verb which usually has some sort of mental image or representation as part of its meaning. The possessor of the mental image (often referred to as the "experiencer" or "cognizer") is realized in some languages as the SUBJ, and in others as something else (an oblique or "indirect object"). The verb need, for instance, realizes its possessor as SUBJ in English but "indirect object" in Russian. On the other hand, the possessor of the verb seem is recognized as an OBL (a 3 (indirect object) in RG terms) in English. Regardless of the surface realization of this argument, it often displays "subject properties," such as anteceding reflexives. In RG, Inversion constructions are treated as involving a change of grammatical relations from a 1 to a 3. Thus, both cases in (67) involve initial 1s. Russian pseudo-reflexives can be antecedded, on the RG view, by initial 1s and final 1s. Lacking evidence to the contrary, Perlmutter suggests that the notion "global 1" is actually the correct one here; this would predict that arguments which advance to 1 and then demote should also antecede pseudo-reflexives, but there are no such constructions.

We have already seen that CRG does not need the notion "initial 1"; the generalizations which are alleged to refer to it can be stated at the semantic/cognitive level. Using our alteration of CST, the facts about Russian pseudo-reflexives can be stated as follows:

(68) Russian pseudo-reflexives are ambiguous between [+subj] and [-subj]. [-Subj] pseudo-reflexives must have as their antecedents the leftmost argument in the semantic representation.
Perlmutter (1979) argues that the antecedent of the Japanese anaphor \textit{zibun} is a "working 1," an argument that bears the relation 1 (SUBJ) at some stratum and a Pure GR at the surface. The argument is based on data such as the following.

(69)  
\begin{itemize}
  \item a. Dono hito ga Ando-san ni zibun no koto ni tuite hanasimasita ka?
      \textit{which person} NOM \textit{DAT GEN thing about talked}
    'Which person talked to Mr. Ando about himself?'
    \textit{(zibun = dono hito; zibun \neq Ando-san)}
  \item b. Yamada-san wa Ando-san o zibun no ie de korosita.
      \textit{TOP ACC GEN house LOC killed}
    'Mr. Yamada killed Mr. Ando in self's house.'
    \textit{(zibun = Yamada-san, zibun \neq Ando-san)}
  \item c. Ando-san wa Yamada-san ni zibun no ie de korosareta.
      \textit{TOP DAT GEN house LOC killed passive}
    'Mr. Ando was killed by Mr. Yamada in self's house.'
    \textit{(zibun = Ando-san, zibun \neq Yamada-san)}
  \item d. Kimura-san ni zibun no koto sika wakaranai.
      \textit{DAT GEN thing only understand}
    'Mr. Kimura understands only things pertaining to himself.'
    \textit{(zibun = Kimura-san)}
\end{itemize}

(69a,b) show that there is a SUBJ-hood requirement on the antecedent of \textit{zibun}. (69c) shows that in passives it is the surface SUBJ, not the "underlying subject" (initial 1) that counts. The crucial case is (69d), where the antecedent is marked with dative case. Perlmutter claims that this is an Inversion structure, and the antecedent here is an initial 1 and a final 3. Unlike the demoted underlying subject of the passive, this former 1 still bears a Pure GR and thus counts as a working 1.

The weakness in Perlmutter's argument is the evidence he gives for surface 3-hood of the "experiencer." He gives two arguments. The first is case marking, but irregular ("quirky") case marking is well attested (see, e.g., Andrews 1982 on Icelandic) so this is not a persuasive argument.
(70) a. [Sannin no kodomotati wa] iti do ni nakhazimeta.
   three GEN children TOP at-once cry-began
   'The three children began to cry at once.'

   b. [Kodomotati wa] sannin iti do ni nakhazimeta.

(71) a. Kimuri-san wa [sanbiki no inu ni] mizu o yatta.
   TOP three GEN dog DAT water ACC gave
   'Mr. Kimuri gave water to three dogs.'


We can see from (71) that quantifiers cannot float from indirect objects.

Consider now a sentence with an Inversion verb.

(72) a. [Korera no sannin no kodomotati ni] eigo ga wakaru.
   these GEN three GEN children DAT English NOM understand
   'These three children understand English.'

   b. *[Korera no kodomotati ni] sannin eigo ga wakaru.

Perlmutter takes this as evidence that the "experiencer" is not a surface
SUBJ.

Considered abstractly, it is clear that Perlmutter's conclusion is
only one of two alternatives. There is a cluster of properties X (includ-
ing antecedence of zibun and also the system of SUBJ Honorifics) and another
cluster Y (case marking and quantifier floating), each of which is general-
ly thought to be associated with surface SUBJs, but which do not occur in
the same environment. The logical conclusion is that one of these clusters
is not really associated with surface SUBJs. Perlmutter assumes that
Y represents the test for surface SUBJ-hood and that X must represent some-
thing else. However, the opposite conclusion, that X is the test for surface
SUBJ-hood in Japanese, is equally possible, and has been argued for convinc-
ingly by Shibatani (1977). Shibatani argues that quantifier floating is
governed by morphological case, not grammatical relation, and that the
antecedent of zibun is always a SUBJ. We will summarize his arguments
here; the reader is referred to the article for a more complete treatment.
First there is a rule in Japanese which raises a genitive NP out of a larger NP, making it a sister of the latter NP and marking it nominative. This new nominative can float a quantifier but cannot be the antecedent of *zibun*.

(73) a. \[[\text{Sorerano } \text{sannin no sensei no}] \text{ okusan ga}] \text{ waka-i} \text{ three GEN teacher GEN wife NOM young-PRES}  
'These three teachers' wives are young.'

b. *[[\text{Sorerano sensei no}] \text{sannin okusan ga}] \text{ waka-i.}  
(Also ungrammatical if \text{sannin} follows \text{okusan ga} under desired reading, OK if 'three' modifies 'wife'.)

c. \[[\text{Sorerano } \text{sannin no sensei ga}] \text{ [okusan ga] waka-i.} 

(d. [[\text{Sorerano sensei ga}] \text{sannin [okusan ga] waka-i.} 

(74) a. \text{Yamada sensei no musuko ga zibun ni unzarisite i-ru.}  
teacher GEN son NOM DAT disgusted be-PRES  
'Prof. Yamada's son is disgusted with himself.'  
\((zibun = \text{musuko}, zibun \neq \text{Yamada sensei})\)

b. \text{Yamada sensei ga musuko ga zibun ni unzarisite i-ru.}  
\((zibun = \text{musuko}, zibun \neq \text{Yamada sensei})\)

These data are very problematic for Perlmutter. (73) indicates that the nominativized NP has to be considered a surface SUBJ by Perlmutter, making it a working 1. We know from passives that "derived" SUBJs can antecede *zibun*. So why can *Yamada sensei* not be the antecedent in (74b)? Furthermore, in the case of "Inversion" verbs themselves, the nominative cannot antecede *zibun*, and cannot trigger SUBJ Honorification.

(75) \text{Sensei ni (wa) okusan ga zibun no kokyoo ni oarini raru.}  
teacher DAT TOP wife NOM GEN native-town DAT have (honorific)  
'The teacher has a wife in self's native town.'  
\((zibun = \text{sensei}, zibun \neq \text{okusan})\)

In addition, the "experiencer," which Shibatani claims is the SUBJ, can be nominativized with a change in meaning (which Shibatani claims is like topicalization), and when this happens it can float a quantifier.
(76) a. [Amerika no taasu no hyakusyoo ni (wa)] okane ga ar-u
America GEN many GEN farmers DAT TOP money NOM have
'Many American farmers have money.'

b. *[Amerika no hyakusyoo ni (wa)] taasu okane ga ar-u.

c. [Amerika no taasu no hyakusyoo ga] okane ga ar-u.

d. [Amerika no hyakusyoo ga] taasu okane ga ar-u.

All the facts follow if, as Shibatani suggests, the condition on quantifier floating is based on case rather than grammatical relations. We conclude that *zibun* is [+subj].

There are some facts in Japanese, not considered by Perlmutter, which indicate that our theory needs to be weakened slightly. In sentences involving the causative affix *sase* and "adversity passives" both the SUBJ and the "logical subject" of the root verb can antecede *zibun*, as discussed by N. McCawley (1976) and Inoue (1976). One possible explanation is that the affixes *sase* and adversity *rare* have the lexical property that the nonSUBJ arguments in question are possible antecedents of *zibun*. Such marking would be similar to the marking of "raised" arguments (§3.7.2). Alternatively, there may be a fourth Pure GR, perhaps SUBJ2, which comes into play here. The key point is that this is limited to two affixes; it is not a general property of *zibun*, and should not lead us to abandon the analysis of *zibun* as [+subj].

Thus, there is no evidence from anaphora that nonsurface grammatical relations are needed in syntactic theory. What is needed is an explicit semantic theory, such as that of Jackendoff (1983), and the correct assignment of surface grammatical relations.
3.6. Lexical Reflexivization

The study of reflexives of various sorts is complicated by the existence of a lexical rule of nonanaphoric reflexivization in some languages. It is most obvious in languages such as the Semitic languages, in which reflexives are formed morphologically. For instance, in Hebrew reflexives are marked with the prefix "hit". Observe the following examples, drawn from Halkin (1970).

(77) a. hilel 'to praise'  hithalel 'to praise oneself'
    b. geel 'to defile'  hitgael 'to defile oneself'
    c. kisa 'to cover'  hitkase 'to cover oneself'
    d. nifek 'to kiss'  hitnafek 'to kiss each other'
    e. raxac/ rexec 'to wash'  hitraxec 'to wash oneself'

(78) a. hexele 'to make NP sick'  hitxale 'to make oneself sick'
    b. hilbil 'to dress NP'  hitlabef 'to dress oneself'

(77) illustrates several cases of Hebrew reflexives and their nonreflexive counterparts. The examples in (78) are similar, except that the nonreflexive forms are causatives, characterized by the prefix "hi" or "he".

Assuming a lexicalist theory of morphology, as we are, Hebrew reflexivization must be a lexical rule, and reflexives must be semi-autonomous lexical entries. If so, we would expect to find some cases of lexical drift, and perhaps some subregularities. We do.

(79) a. men 'to train'          hitamen 'to train oneself, practice'
    b. xizek 'to strengthen'  hitxazek 'to strengthen oneself, exert oneself'
    c. maxar 'to sell'        hitmaker 'to sell oneself, devote oneself'
    d. nise 'to exalt'        hitnase 'to boast'
    e. katav 'to write'       hitkatev 'to correspond'

(80) a. iver 'to air'         hitaver 'to become aired'
    b. beer 'to explain, elucidate'  hitbaer 'to become clear, be elucidated'
    c. giyer 'to make NP a proselyte'  hitgayer 'to become a proselyte'
    d. yibef 'to dry'           hityabef 'to become dry'
    e. piteax 'to develop NP'   hitpateax 'to develop (intrans.)'
The examples in (79) have undergone some rather random semantic drift. (80) exemplifies a common subregularity: the use of lexical reflexives for inchoatives.

Hebrew is not a problem from our point of view. It is clear from the general assumptions underlying the theory that Hebrew reflexivization is lexical. The problem is that there are languages in which reflexivization looks superficially like an anaphoric process when it is, in fact, lexical. This is discussed with respect to French (and, by extension, other Romance languages) by Grimshaw (1982b). She points out that although French reflexive clitics resemble object clitics, and are identical in non-third-person forms, reflexive verbs in French display some properties of lexical reflexives. For instance, she cites the verb s'évanouir 'to disappear', where s' is the reflexive clitic. There is no verb évanouir of which s'évanouir is the reflexive. Se is also used for inchoatives in French, just as in Hebrew. Thus, at least some cases of se must be taken as meaningless morphemes. Grimshaw then goes on to show that even in cases of simple reflexivization, a lexical analysis is well justified. She states a rule which binds the semantic argument corresponding to the OBJ to the one corresponding to the SUBJ. Baker (1983a) provides a similar analysis for Italian, although, for reasons extraneous to present concerns, he states the rule to eliminate the SUBJ instead, followed by an independently motivated OBJ $\rightarrow$ SUBJ rule.

It is pointed out by Yang (1983) that many of the forms in various languages that we have identified as pseudo-reflexives also appear in "inherent" reflexives, i.e. verbs that are reflexive in form but not in meaning, like the Hebrew hitnase or the French s'évanouir. One very puzzling case is the Dutch zich, discussed by Everaert (1980). We
saw in §3.4 that *zich* is usually noncoreferential with its coarguments; hence we assigned it the features [-local, +local']. However, as Everaert points out, there are counterexamples to our analysis of *zich*, such as (81).

(81) Jan *wast* zich/zichzelf.
    washed himself

In (81), it seems that both the r-anaphor *zichzelf* and the pseudo-reflexive *zich* can be coreferential with the SUBJ. Everaert goes on to show that *zich*, but not *zichzelf*, can be used in inherent reflexives.

(82) a. Rini vergist zich. 'Rini is mistaken.'
    b. Jan schaamt zich 'Jan is ashamed.'

However, he argues against an analysis whereby the use of *zich* in (81) and (82) is attributed to an idiosyncratic lexical property of each verb, an analysis which states that *zich* is usually not coreferent with a coargument. Although the exact analysis he argues against is not one that considers *zich* a marker of lexical reflexivization, it is similar enough for us to examine the arguments. Our analysis would state that there are two lexical items *zich* in Dutch (presumably, their phonetic identity is due to historical circumstance): one a pseudo-reflexive, as discussed earlier, and the other not an anaphor at all but rather a marker of the lexical rule of reflexivization. The analysis that Everaert argues against considers *zich* a "free anaphor" (i.e. a pronoun) which can, under lexical government, be coreferential with a coargument, but lacking an appropriate governing verb cannot. He states (pp. 6-8)

There is a group of verbs that confirms the opposite assumption, that inherent reflexive verbs are not idiomtic [i.e. idiosyncratic] expressions. These verbs can have both "zich" and "zichzelf" in
object position, while other verbs again can have only "zichzelf" in object position. The opposition between these groups of verbs in general cooccurs with the (im)possibility to keep the object position unlexicalized.

[(83)]  \( \text{Jan}_1 \) droogde - af
\[ \text{zich}_1 \]
\[ \text{zichzelf}_1 \]
(John rubs -/himself down)

[(84)]  \( \text{Zij}_1 \) nam -* waar
\[ \text{zich*} \]
\[ \text{zichzelf} \]
(She observed -/[her]self)

If "zich" were a free anaphor, the verbs of [(83)] would have to allow the analysis of an idiomatic expression, and therefore double subcategorization.

[(85)]  Jan wast Karel  (transitive verb)
\[ \text{zichzelf} \]

[(86)]  Jan wast Karel*  (inherent reflexive verb)
\[ \text{zich} \]

Such a distinction could be motivated by a difference in meaning between the two subcategorizations. But there is no difference in meaning between [(85)] and [(86)], hence the assumption of two subcategorizations seems unmotivated. In the following two sentences a change in meaning can be discerned:

[(87)]  Hij verzekert het huis  (transitive verb)
(He assures the house)

[(88)]  Hij verzeerde zich van zijn medewerking  (inherent reflexive verb)
(He secures his cooperation)

In this case it is justified to speak of two subcategorizations: inherent reflexive [(88)] and transitive [(87)].

The same objections mentioned in connection with [(87), (88)] apply to the following inchoative verbs...

[(89)]  a.  \( \text{Hij}_1 \) buigt de tak
\[ \text{zich}_1 \]
\[ \text{zichzelf}_1* \]
(He bends the branch / the branch bends)

[(90)]  a.  \( \text{Hij}_1 \) verbreedde de weg
\[ \text{zich}_1 \]
\[ \text{zichzelf}_1* \]
(He broadened the road / the road broadened)

[(91)]  a.  \( \text{Hij}_1 \) verspreidde het gerucht
\[ \text{zich}_1 \]
\[ \text{zichzelf}_1* \]
(He spread the rumour / the rumour spread)

[(92)]  a.  \( \text{Hij}_1 \) beweegt de boom
\[ \text{zich}_1 \]
\[ \text{zichzelf}_1* \]
(He moves the tree / the tree moves)

These verbs can leave their object position unlexicalized [or] have
"zich" in object position, while there is no difference in meaning in the b-sentences of [(89)-(92)].

If one assumes inherent reflexive verbs to be idiomatic expressions, the appearance of inherent reflexive verbs would become completely arbitrary. [(83)-(92)] shows this would be an unfortunate position to hold.

The problem, then, is that lexical reflexivization, where it is regular, means the same thing as anaphoric reflexivization, as well as the existence of subregularities such as the one relating to inchoatives. As to the synonymy with anaphoric reflexivization, this would not be a problem in a language like Hebrew or French where this is the only way to reflexivize an object; while one may wish to argue against a lexical analysis of reflexivization in French, or even in Hebrew, it would not be based on the fact that some of these lexically derived reflexives have simple reflexive meaning. The problem is that Dutch possesses an r-anaphor in its catalog of anaphors. What we would expect under a lexical analysis of *zich* is that all coargument reflexives would be able to use *zichzelf*, while the use of *zich* would be lexically governed. This seems to be the case. In fact, two additional pieces of information provided by Everaert in the section just cited provide further evidence in favor of a lexical analysis: the correlation between a null "object" and *zich*, and (despite Everaert's comments) the use of *zich* in inchoatives. We have already seen that lexical reflexive verbs often have inchoative meaning; it is unclear what the subscripts in the inchoative sentences mean, but (80) surely does not really mean 'The branch bends itself'? This is why *zichzelf* is ungrammatical here. As to the omissibility of the "object," even if one wishes to claim that *zich* fills an argument position in sentences like (83), the fact that the argument need not be realized is itself an idiosyncratic lexical fact about the verb, one to which the use of *zich* could be
linked by lexical rules. In our analysis, all that needs to be said is that the use of *sich* to mark lexical reflexives is optional, unlike Hebrew *hit-* or French *se*.

The point of mentioning lexical reflexivization here is to show that apparent counterexamples to our analyses are not necessarily real counterexamples. Chomsky often states that raw data cannot disconfirm a theory; this is surely true of the distribution of Dutch *sich*.

3.7. Raising and Control Phenomena

We turn now to the phenomena that deal with the interpretation of understood subjects of nonfinite verbs — the phenomena referred to in the transformational literature as Raising-to-Subject, Raising-to-Object, Equi, and SuperEqui.

3.7.1. Division of the Phenomena

The area of raising and control phenomena is a very controversial one. It is not even agreed how the phenomena should be split up. In transformational grammar, because of the need to represent predicate-argument relations in deep structure (or D-structure), Raising-to-Subject and (in some analyses) Raising-to-Object were treated as movement transformations, while Equi and Super Equi were deletions or interpretive rules. However, in some current frameworks, including CRG, this particular division is not necessary.

We will basically follow the view put forth in LFG (Bresnan 1982a): raising and control phenomena split up into two subphenomena, what Bresnan calls *functional control* and *anaphoric control*. As we are treating all forms of control as anaphora, the terms that Bresnan uses are inaccurate.
from our point of view. We will therefore refer to the types of control by the names of the anaphors. For the anaphor in anaphoric control we use the familiar designation PRO. It covers the same ground as Government/Binding theory's PRO (although Mohanan 1983b suggests that Chomsky's 1982 pro is more appropriate for anaphoric control) and LFG's 'PRO'. For functional control we will resurrect the symbol Δ. Both PRO and Δ are anaphors with no configurational realization. They do not block contraction or participate in any other phonological process.

In anticipation of our analysis and to clarify the differing terminology of various theories, we present in Figure 3.1 a summary of the different ways in which raising and control phenomena are organized in classical transformational grammar (Rosenbaum 1967, Postal 1974, Jackendoff 1972, Chomsky 1973), GB (Chomsky 1981, 1982), LFG (Bresnan 1982a), and CRG.

Two notes are in order about our terminology. In the first place, we will use the word "control" in a nontheoretical observationally descriptive way to refer to any case of a superficially subjectless infinitive, thus subsuming all the phenomena to be discussed in the remainder of §3.7. We attach absolutely no theoretical significance to the term. The second note is that we use the term "obligatory Equi" rather than the more fashionable "obligatory control." This is because the former refers to the obligatory absence of a surface subject in the complement while the latter correctly only requires that if the subject position is absent it be interpreted as coreferential with a particular NP. (This is discussed briefly, although without the terminological distinction, by Mohanan 1983b, fn 1.)
<table>
<thead>
<tr>
<th></th>
<th><strong>Raising</strong></th>
<th><strong>Obligatory Equi</strong></th>
<th><strong>Nonobligatory Equi</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TG</strong></td>
<td>to SUBJ: (Extraposition)</td>
<td>Equi*</td>
<td>Equi/SuperEqui*</td>
</tr>
<tr>
<td></td>
<td>Raising</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>to OBJ: (Raising)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GB</strong></td>
<td>$ deletion, plus: Control (PRO)**</td>
<td>Control (PRO)</td>
<td>Control (PRO)</td>
</tr>
<tr>
<td></td>
<td>to SUBJ:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NP movement</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>to OBJ:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exceptional Case</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marking</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LFG</strong></td>
<td>functional control in verbs:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>functional control</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>in nominals:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>anaphoric control ('PRO')</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CRG</strong></td>
<td>$\Delta$-control</td>
<td>$\Delta$-control</td>
<td>PRO-control</td>
</tr>
</tbody>
</table>

Transformations in parentheses under Raising for TG not accepted as operable in these constructions by some transformational theorists.

*either a deletion transformation or a semantic interpretive rule

**according to Mohanan (1983b), this role should be taken by pro instead of PRO.

Fig. 3.1. Raising and control phenomena
3.7.2. Raising

The phenomena of Raising-to-Subject and Raising-to-Object (Rosenbaum 1967, Postal 1974) can be described in very simple terms: an infinitive has as an understood subject the non-logical subject or object of the governing clause. The LFG analysis is straightforward: the c-structures of (93) are (94), with no empty nodes or traces. The f-structures are roughly as in (95), where the argument outside the angle brackets is not linked to a semantic argument.

(93)  
a. The hamster seems to hate Max.

b. I expect the hamster to hate Max.

(94)  
a.

\[
\begin{array}{c}
\text{S} \\
\text{NP} \quad \text{VP} \\
\text{the hamster} \quad \text{seems to} \quad \text{VP} \\
\text{V} \quad \text{NP} \\
\text{hate Max}
\end{array}
\]

b.

\[
\begin{array}{c}
\text{S} \\
\text{NP} \quad \text{VP} \\
\text{I} \quad \text{V} \quad \text{NP} \\
\text{expect} \quad \text{to} \quad \text{VP} \\
\text{the hamster} \quad \text{V} \quad \text{NP} \\
\text{hate Max}
\end{array}
\]
The grammatical relation XCOMP is an "open function" — its SUBJ is controlled by a SUBJ, OBJ, or OBJ2 in the governing structure: this control is a formal identity of all features. Control is lexically determined by the governing predicate by means of control equations in the lexical entry. For instance, (96a) is in the lexical entry of seem and (96b) in
that of expect. (The up-arrows are metavariables referring to the structures headed by seem and expect, as discussed in Kaplan and Bresnan 1982.)

(96)   a. (+SUBJ) = (+XCOMP SUBJ)

       b. (+OBJ) = (+XCOMP SUBJ)

In order to adapt the LFG analysis of Raising we must consider two issues: the source of the anaphor \( \Delta \) and the determination of the antecedent. We will take them in the given order.

The presence of \( \Delta \) is dependent on the grammatical relation that the phrase of which it is the SUBJ bears: it must be open, in the sense discussed above. Furthermore, at least for many languages, \( \Delta \) is only annotated to SUBJs. We stipulate these facts, as does LFG, as being a part of UG.

(97) Annotate to the SUBJ of an open complement (XCOMP) the anaphor \( \Delta \).

In CRG, as in LFG, the limitation of \( \Delta \) to SUBJs of nonfinite clauses (and nonclauses, such as APs) must be stipulated; it does not follow from any independent principles. Thus, it contrasts with the approach taken in Government/Binding theory, in which Raising(-to-Subject) and Exceptional Case Marking (=Raising-to-Object) can only apply to the complement SUBJ, and not if the complement is a finite clause. This may be seen as a disadvantage of the present theory, but it is only if the generalization that only SUBJs of nonfinites can be \( \Delta \) is true. It is certainly true that in most of the languages that have been placed under scrutiny by generative linguists only the SUBJ can raise. However, Seiter (1983) presents evidence from the Polynesian language Niuean that OBJs can raise as well. In addition, Simpson (1983) suggests that "clause union" constructions in Warlpiri and other languages be treated as involving total functional control (i.e. the matrix SUBJ controls the complement SUBJ, the matrix OBJ
controls the complement OBJ, etc.). It is possible that the universal constraints of GB that limit what can raise will prove not to stand up to scrutiny against the right languages.

What are the features of \( \Delta \)? We begin by observing that Raising is obviously governed by some locality principle. This follows automatically in the LFG account because a string of grammatical functions, as on the right-hand side of the equations in (96), only name two functions. In our framework there are no functional equations, and therefore no universal constraints on functional equations. We cannot assign \( \Delta \) the feature \([+\text{local}]\) without drastically rewriting our definitions of Extended Local Binding Domain. The definitions in §3.3 state that in English and other languages the arguments of verbs cannot look outside their argument list to find an antecedent, yet \( \Delta \) always finds its antecedent one clause up. The other locality feature that we have proposed is \([+\text{local'}]\). By any definition of pseudo-clause that we have proposed except for the one for Dutch, the antecedent of \( \Delta \) is located in the minimal dominating pseudo-clause. We will propose tentatively that this is the correct feature. In many cases, "minimal pseudo-clause" may actually not be restrictive enough. If the dominating clause is nonfinite (or subjunctive) one might expect \( \Delta \)'s antecedent to be even higher. The nature of the open grammatical relation XCOMP and the concept of a "consequent" to be discussed presently may be involved. It is even possible that it will turn out that no locality feature is needed. We leave this question unresolved.

In order to understand \( \Delta \), it is important to realize that its relationship with its antecedent is different from the antecedent-anaphor relations we have seen thus far. There is a very real sense in which
Δ's presence is mandated by the existence of its antecedent. The reason is that the antecedent is a syntactic argument which is not linked to a semantic argument list position. Thus, it needs to be linked to the argument structure of the sentence as a whole, and this is done through Δ-control. The intuition here is similar to the one behind GB's θ Criterion (Chomsky 1981) and one version of LFG's Coherence Condition (Bresnan 1982b; in Bresnan 1982a and Kaplan and Bresnan 1982 the Coherence Condition merely requires that subcategorization be met).

(98) a. θ Criterion (informal version)

Each argument [i.e. pronoun, reflexive, PRO, referential lexical NP, clause] bears one and only one θ-role and each θ-role is assigned to one and only one argument.

b. Coherence Condition (Bresnan 1982b version)

Every meaningful functional argument of the type specifiable by a lexical form [= every item bearing a Lexical Relation] must be bound to a predicate argument.

We need a term to distinguish Δ from pronouns, r-anaphors, pseudo-reflexives, and subjective anaphors. We will use the term consequent (borrowed, with apologies, from Koster 1978a, where it is used the way we are using "anaphor"), and add [+consequent] to our list of features. The following lexical redundancy rule, the CRG equivalent of (98), can now be stated.

(99) Mark a nonlogical argument to require a consequent.

Since only SUBJ, OBJ, and OBJ2 can be nonlogical arguments, they are (under this analysis) the only possible consequent-requiring grammatical relations. We will discuss this further in §3.7.3. In (100), we present the CRG representations of (93), using c to indicate an argument that requires a consequent.
Finally, we define the '+' and '-' values of the consequent feature.

(101)  a. A [+consequent] anaphor must have a consequent-requiring argument as its freely assigned antecedent.

       b. A [-consequent] anaphor must not have a consequent-requiring argument freely assigned as an antecedent.

The anaphoric bind between $\Delta$ and its antecedent is a very close one.
Grammatical requirements on one (such as case, in some languages) are required of both (see Bresnan 1982a and references there). This is due to the fact that $\Delta$ is a consequent.
3.7.3. Obligatory Equi

By hypothesis, Δ is also involved in obligatory Equi constructions. Although Δ-control is roughly equivalent to what is called functional control in LFG, some cases of obligatory Equi, which we will treat as Δ-control, have to be treated as anaphoric control in the Lexical-Functional theory. The result is that, at least for English, it turns out to be extremely difficult to determine which kind of control is involved. We will defend the following hypothesis: Control in an environment in which a full clause with a lexical subject may not appear involves the anaphor Δ and the relation XCOMP.

Functional control in Lexical-Functional Grammar is induced by the following universal lexical rule (Bresnan 1982a).

(102) Let L be a lexical form and $F_L$ its grammatical function assignment (=syntactic argument list). If $XCOMP \in F_L$, add to the lexical entry of L:

\[ (+OBJ2) = (+XCOMP \, SUBJ) \text{ if } OBJ2 \in F_L; \]
\[ \text{otherwise:} \]
\[ (+OBJ) = (+XCOMP \, SUBJ) \text{ if } OBJ \in F_L; \]
\[ \text{otherwise:} \]
\[ (+SUBJ) = (+XCOMP \, SUBJ). \]

Thus, in the unmarked case, the controller is determined by a hierarchy of the Pure GRs. Cases of SUBJ control by a transitive verb (promise, strike, etc.) are assigned an idiosyncratic control equation that overrides (102). Since only SUBJs, OBJs, and OBJ2s are allowed to control, these verbs cannot passivize ("Visser's generalization").

Kisala (1981) examines control phenomena in nominals within the LFG framework. She points out that under the assumption that Pure GRs do not appear as arguments of nominals (Rappaport 1983) it is impossible to treat control in nominals as functional control of XCOMP. even in cases where
lexical SUBJs can never appear. In fact, null (unexpressed) arguments can even control. Consider the cases in (103), in the notation of §1.2, with the controller underlined.

(103) a. The general ordered the troops to retreat.
   <Agent/Source, Goal, Propositional Theme>
   \{SUBJ, OBJ, XCOMP\}

b. the general's order to the troops to retreat
   <Agent/Source, Goal, Propositional Theme>
   \{POSS, OBL, (X)COMP\}

c. the general's order to retreat
   <Agent/Source, Goal, Propositional Theme>
   \{POSS/\emptyset, \emptyset, (X)COMP\}

Therefore, Kisala suggests that even in cases like these, where no lexical SUBJ can appear, anaphoric control (involving the anaphor PRO and the closed argument COMP) is a possible analysis. The controller is determined on the basis of a hierarchy of thematic relations: a Goal or Theme controls if there is one, otherwise the Agent controls.

Brame (1976) rejects a transformational derivation of Equi constructions. One of his arguments concerns the power of subcategorization. He argues that the following two possibilities exist for obligatory Equi verbs (like order, try, etc.).

(104) a. These verbs are subcategorized for $\bar{S}$. They are also lexically marked to trigger the Equi transformation (or have a PRO subject in their complements) obligatorily.

b. These verbs are subcategorized for $\bar{VP}$, which is a category distinct from $\bar{S}$. Nothing more need be stated.

(104a) assigns the lexicon much more power over the syntax than (104b).

Brame's specific argument was based on a traditional system of categorial
subcategorization. In LFG and CRG, however, subcategorization is (basically) relational in nature. The basic form of the argument still holds if we substitute COMP for S and XCOMP for VP. Because we are concerned at present with nominals, we substitute the word "nominals" for "verbs."

(105) a. These nominals are subcategorized for COMP. They are also lexically marked to have a PRO subject in their complements.

b. These nominals are subcategorized for XCOMP. Nothing else need be stated.

In both of the a. specifications, the power of the lexicon is increased. Thus, a more restrictive theory results if we adopt the b. approach. For CRG, this means treating obligatory Equi uniformly as Λ-control.

The above should not be taken to mean that we reject Kisala's analysis in toto. We disagree with her decision, forced upon her by the formalism of LFG and the analysis of obligatory Equi as being relationally based, of treating obligatory Equi in nominals as PRO-control, but we agree that control of Λ is determined by a hierarchy of thematic relations. What we need to do is determine what the hierarchy is and whether or not it is independently motivated. In fact, an independently motivated thematic hierarchy does exist: the one proposed by Jackendoff (1972). (In §3.5.3 it was suggested that this hierarchy be replaced by a condition based on the left-to-right order in an explicit semantic representation. We leave the application of this suggestion to control for future work.) Jackendoff's hierarchy is (106).

(106) The Thematic Hierarchy

a) Agent
b) Location, Source, Goal
c) Theme

It would be reassuring if Kisala's hierarchy could be stated in terms of this one. It can. Kisala could find no nominals with both a Theme and
one of the (b) relations, so her findings can be restated as (107).

(107) In a nominal lexical entry with an XCOMP in its argument list, mark the argument (other than the XCOMP itself) lowest on the Thematic Hierarchy to take a consequent.

In discussing the extension of her analysis to non-nominals, Kisala claims to find evidence for separating Goal and Theme. The resulting hierarchy is Goal-Theme-Agent, which is not compatible with (106). However, the evidence that she uses, purposive clauses, is irrelevant to the question at hand. Purposives can have lexical subjects, and are not even arguments. Therefore, we can maintain (107).

Can (107) be extended to verbs? This question separates into the following:

- Can the controller in verbal argument lists be thematically determined?
- Should control in verbal argument lists be limited to Pure GRs?

The first question can be answered in the affirmative. Note that verbs and their nominalizations share control properties. The most economical lexicon would capture this fact. In fact, some verbs that are exceptional under the theory of functional (relational) control are perfectly normal under thematic control. For instance, in LFG it is accidental that strike and regard have opposite control properties: both are transitive, yet complements to strike are controlled by the SUBJ, while complements to regard are controlled, as expected, by the OBJ. Jackendoff (1972) shows that the SUBJ of strike is a Theme and its OBJ a Goal, and that these are reversed with regard. Under a thematically based theory of control, therefore, we expect the control properties to be exactly what they are. (Promise, on the other hand, is still irregular. However, the acquisitional study in C. Chomsky 1969 suggests that the control properties of promise are not related to its semantics.) So we can restate (107) as (107').
In an argument list with an XCOMP, mark the argument (other than the XCOMP itself) lowest on the Thematic Hierarchy to take a consequent.

(107') subsumes Raising. Jackendoff (1972), on the basis of thematic constraints on English r-anaphors, proposes that nonlogical arguments should count as lowest on the hierarchy. This renders (99) redundant. We will return to it in §3.8.

We turn now to the second question: the restriction of the class of controllers in verbal argument list to the set of Pure GRs. Bresnan's (1982a) evidence for this restriction seems compelling in English and other Indo-European languages. We state it using the categorial feature [+PGR], i.e. taking arguments bearing Pure GRs. This feature is discussed in the Appendix.

(108) In the argument list of a [+PGR] lexical item, only arguments bearing Pure GRs may be marked to take a consequent.

Whether or not a specific language makes use of (108) is a parameter of Universal Grammar. Bell (1982) shows that the controller of obligatory Equi construction in the Phillipine language Cebuano is the initial subject or object, a concept that we have already seen (§3.5.3) can be reduced to thematic relations. Note (109).

(109) Gisa'aran nako' si Lus [sa pagluto' ug panihapon].

promise-LOC I-GEN NOM OBL cooking OBL dinner 'Lus was promised by me to cook dinner.'

In (109) the verb is in a voice (locative) which indicates that something other than the "logical subject" is the SUBJ. The logical subject is 'I' and it controls the SUBJ of the subordinate "clause" (actually an NP). 'I', which is marked with genitive case, is neither a SUBJ nor an OBJ(2). In RG terms, it is a chômeur, a demoted argument. CRG does not recognize the chômeur relation, but it is clear that whatever relation 'I' does
hold is not one of the Pure GRs. The case marking suggests POSS, but it is unimportant here. The crucial point is that, unlike the English translation, which is ungrammatical, it is possible for an argument not bearing a Pure GR to control $\Delta$ in Cebuano. (The reader is referred to Bell 1982 for evidence that this construction shares the properties we have attributed to $\Delta$.)

The moral of this section is that we now have evidence for representing anaphora at the level of argument structure, as forced on us by the theory. It allows for unified treatment of obligatory Equi. Thus, it provides indirect evidence for CRG.$^8$

3.7.4. The Anaphor PRO

We turn now to the anaphor PRO. By PRO, we mean a subset of the way the term is used in GB theory; roughly those uses of PRO which Mohanan (1983b) argues should be replaced by pro. As discussed by Bresnan (1982a), PRO also covers cases of "pro-drop," which Chomsky (1982) calls pro. In English, PRO is introduced into lexical entries by the following lexical rule.

(110) Annotate as an optional value of the SUBJ of a [-finite] verb the anaphoric features of PRO.

What are the anaphoric features of PRO? We begin with the locality features. As is well known, PRO cannot have a coargument as an antecedent, and its antecedent may be as far away as memory permits. Furthermore, it need not have any antecedent at all. Therefore, it must be [-local] and [-local']. It might be supposed that PRO is [+consequent].
It has frequently been observed that the semantic or thematic structure of the matrix predicate can induce control of a complement 'PRO'. This appears to account for the examples in [(111)].

[(111)]

a. Louise signalled to Ted to shave himself.
b. *Louise signalled to Ted to shave herself.
c. *Louise signalled to Ted to shave oneself.

[Bresnan (1982a, 385; 333 in Bresnan, ed.)]

However, this does not seem to be correct. First of all, as already noted, PRO does not require an antecedent, in violation of our definition of [+consequent]. More to the point, oblique arguments can control PRO, even in VPs (see (111)) in violation of (109). In addition, the control properties of verbs like signal seem to be partially dependent on the thematic structure of the complement!

(112) Joey signalled to the teacher to be excused.

In (112), the SUBJ controls, while in (111) it is the OBJ. Therefore, a unique argument of signal cannot be marked to require a consequent. A plausible explanation is provided by Jackendoff (1972), who makes the following observation regarding (113).

(113) a. *I screamed to go.
b. I screamed to Bill to go.
c. I screamed to Bill for Harry to go.
d. I screamed to be allowed to go.
e. I screamed to Bill to be allowed to go.
f. I screamed to Bill for Harry to be allowed to go.

Note first that in [(113c,f)] I am asking Bill to bring it about that Harry go or be allowed to go; it is only by virtue of action on Bill's part that anything will happen to Harry. Bill thus meets our intuitive criteria for being an Agent of the complement clause, suggesting that ... the NP in the main clause ... is marked as Agent [of the complement clause, by virtue of the semantics of the matrix verb]...

Sentences like [(113c)] also show the dependence of the ability to take an Agent on the meaning of the sentence. Compare the following:

[(114)]
a. *I shouted to Bill to be tall.
b. *I shouted to Bill for Harry to be tall.
c. I shouted to Bill for the next recruit to be tall.
Be tall does not normally allow an Agent... Why is \((114c)\) good? The meaning of \((114c)\) is that Bill should exercise choice in selecting the next recruit, and that the one he selects should be tall. [pp. 222£]

Jackendoff then goes on to explain that in \((113e)\), the complement SUBJ cannot be the Agent so the matrix OBL cannot control. Similarly, all the control facts in \((111)-(114)\) can be explained.

The resulting feature complex \([-\text{local}, -\text{local}', -\text{consequent}]\) is identical to that of pronouns. Is there, in fact, any reason to distinguish PRO from pronouns, or can we consider PRO to be merely a kind of pronoun, as does Postal (1970)? The answer seems to be that they are to be distinguished, but it is important to scrutinize claims to this effect in the literature very carefully. We will examine three such proposed differences, one due to Chomsky (1981) and two to Bresnan (1982a). We will determine that only one is justified.

Chomsky (1981) claims that PRO is a pronominal which is unique in that it lacks capacity for inherent reference. Thus, it is a pronominal \((r-)\) anaphor. This property represents the explanation for the generic (or "arbitrary") sense of PRO. In the present system, inherent reference is a property of anaphors, and lack of it is due to a positive value for an anaphoric feature. Yet PRO has the value \('-'\) for all the features we have considered (except that it is neutral with respect to \([±\text{sub}])\). Therefore, it should have capacity for inherent reference; i.e. reference to some extra-sentential entity. In fact, Bresnan (1982a) shows that there are discourse contexts in which PRO can do just that. Furthermore, it is not surprising that PRO has a generic interpretation. There is a pronoun with a generic interpretation: one. Why shouldn't the unexpressed pronoun be able to stand for this particular pronoun?9
Bresnan (1982a) proposes a principle of SUBJ-SUBJ obviation. In its least general, English version, it is stated as (115).

(115) **Obviation Principle: English**

If P is the pronominal SUBJ of an obviative clause C, and A is a potential antecedent of P and is the SUBJ of the minimal clause nucleus that properly contains C, P is or is not bound to A according to whether P is + U[^expressed], respectively.

This gets the right results in (116), if for clauses are obviative in English.

(116) a. Sara wished to vote. (PRO = Sara, obligatorily)

b. Sara wished for her to vote. (her ≠ Sara)

Bresnan herself points out that (116) does not always apply to the [+U] pronoun (PRO) (e.g. in (111)). Note, however, that there is no reason to suppose that the subjectless clause in (116a) is an obviative clause. There is no overt complementizer, and lexicalist theories such as the ones under discussion do not recognize deletion rules. Thus, the obviation principle should never apply to sentences with PRO SUBJ$s because the obviation complementizer for is absent.

A proposed distinction that is valid has been discussed by Bresnan (1982a). It pertains to a structural constraint on the antecedent-anaphor relation when the anaphor is PRO, but not when it is a pronoun.

(117) a. People who know Max often discuss working too hard.  
(PRO = people..., optionally; PRO ≠ Max)

b. People who know Max often discuss his working too hard.  
(PRO = Max, optionally)

(118) a. *Contradicting himself will demonstrate that Mr. Jones is a liar. (PRO ≠ Mr. Jones)

b. His contradicting himself will demonstrate that Mr. Jones is a liar. (his = Mr. Jones)

c. Contradicting himself will discredit Mr. Jones. (PRO = Mr. Jones)
There seems to be some sort of command condition on the controller-PRO relation. Bresnan, basing herself on Mohanan (1983b), states it as a condition of c-command at f-structure, calling it f-command. We will define f-command in CRG terms in §3.9. For now, we note that the f-command condition is not strong enough.

(119)  
   a. Max mentioned that working too hard is a topic of conversation at the plant. (PRO ≠ Max)  
   b. Max resents the stories about working too hard that I always tell. (PRO ≠ Max)

(119) seems to involve a constraint of the same sort as (117) and (118). Once again, it is limited to PRO; if we substitute the pronoun his as the subjects of the gerunds in (119), Max is a possible antecedent. In (119a), PRO is an argument of the SUBJ of an embedded clause, while in (119b) it is in a complex NP. These are both "islands" in the sense of Ross (1967). An interesting hypothesis is that PRO, like the anaphor in syntactic binding dependencies, which we will call gap, is subject to island constraints. It is trivially true that the fronted phrase - gap bind obeys the f-command condition, so we will consider it one of the island constraints.

We will discuss islands in §3.9. We will propose here, however, that there is an additional anaphoric feature which identifies those anaphors which are constrained by island conditions; this will be the feature that distinguishes PRO from pronouns. We will call the island conditions that both PRO and trace are subject to the Configurational Constraints and call the feature [+categorial].

(120)  
   a. The link between a [+categorial] anaphor and its antecedent is constrained by the Configurational Constraints.  
   b. The link between a [-categorial] anaphor and its antecedent is not constrained by the Configurational Constraints.
The intuition behind the terminology is that the constraints define illicit configurations for the antecedent-anaphor relation, which is defined for the most part in terms of categories and configurations. This essentially the view of Ross (1967). A more recent view of islands is that there is some sort of locality (or bounding) constraint on syntactic binding, and apparent long-distance cases can be reduced to chains of local dependencies by successive cyclic movement-to-COMP (Chomsky 1973, 1977, 1981), clause-by-clause indexing of base-generated traces in COMP or of the COMP itself (Koster 1978a, Bresnan and Grimshaw 1978), or, in LFG, equations associated with the COMP linking domains (Kaplan and Bresnan 1982). This view, although still controversial, is not implausible for syntactic bonding constructions. The extension of this subjacency account to the controller-PRO relation seems harder to defend. If PRO is truly subject to island constraints, a unified account of fronted phrase-gap relations and controller-PRO relations is in order. We will therefore not consider a locality account of islands.

3.8. Syntactic Binding

There has been much discussion in the generative literature concerning the analysis of "long distance" dependencies (Ross 1967; Chomsky 1973, 1977, 1981; Postal 1970; Bresnan 1976, 1977; Gazdar 1981; and others). Following Chomsky (1977), we will consider wh-preposing, topicalization, "tough movement," comparativization, and other unbounded relations to be reflexes of the same phenomenon, what is often called syntactic binding and has also been called wh-movement (Chomsky 1977), operator binding (Brame 1978), Ā-binding (Chomsky 1981), overlay relations (Johnson and Postal
1980), and constituent control (Kaplan and Bresnan 1982). A major issue is whether or not these dependencies are local (or bounded). The earliest work in the literature assumed unboundedness (Ross 1967, Postal 1970). This view has been adopted in more recent work by Bresnan (1976, 1977) in an EST framework, Gazdar (1981) in Generalized Phrase Structure Grammar, Johnson and Postal (1980) in Arc Pair (Relational) Grammar, and others. The bounded view has its origins in Chomsky's work (see, e.g. Chomsky 1973, 1977) and has been adopted in Government/Binding theory (Chomsky 1981), in related frameworks (such as Koster 1978a), and in Lexical-Functional Grammar (Kaplan and Bresnan 1982). Locality is supposed to account for the island conditions of Ross (1967). However, as we have seen, PRO seems to be subject to island conditions as well, redering the local account harder to justify.

To begin the discussion of locality, we will summarize a proposal made by Falk (1983b) concerning syntactic binding. It is proposed there that syntactic binding constructions binding a SUBJ do not involve configurationally represented traces. Part of the justification for excluding trace from [NP, S] position is that the dummy auxiliary do need not be present in main clause wh questions in which the subject is questioned. In classical transformational grammar, this was accounted for by ordering the rule of Do Support after Affix Hopping. If the subject NP is fronted, the tense morpheme in AUX remains adjacent to the verb, allowing Affix Hopping to apply. The structural description for Do Support is therefore not met. In trace-theoretic transformational theory, however, a trace remains in [NP, S] position intervening between INFL and the verb after Subject-Auxiliary Inversion. According to Chomsky (1981), Case-marked traces are visible to the rules of the P(honetic)F(orm) component, and
Affix Hopping is a rule of that component. Since [NP,S] is assigned nominative Case in finite clauses, this trace should block Affix Hopping. In the nontransformational approach taken to "subject-auxiliary inversion" in Falk (1983a), direct question interpretation is dependent on an AUX preceding the subject NP. The only exception to this is subject questions. The exception is unmotivated if there are [NP,S] traces, but if there are none, inserting do would not help get question interpretation because there would still not be an AUX preceding an NP. (This argument is given in more detail in Falk 1983b.) Another argument against [NP,S] traces, not given in Falk (1983b), is the fact that it is possible to contract an auxiliary onto a fronted wh word (or the complementizer that) over the putative trace.

(121)  
a. The man who (e) is going to sell me a hamster $\Rightarrow$  
The man who's going to sell me a hamster.

b. A study that (e) has been widely quoted. $\Rightarrow$  
A study that's been widely quoted.

The fact that this sort of contraction is possible suggests very strongly that the traces in parentheses in (121) do not really exist. Falk (1983b) proposes, in the LFG formalism, that the metavariable which actually marks the syntactic bindee (Kaplan and Bresnan 1982) is not introduced by the equation annotated to trace (122a) but by a special equation (122b) optionally annotated to the S node.

(122)  
a. $\uparrow$  
b. $(\downarrow$SUBJ)$\Rightarrow$

(122a) designates the node to which it is annotated as a syntactic bindee; (122b) states that the SUBJ of the clause is a bindee. It is also suggested there that there is a fundamental distinction between the metavariable...
ables in the two equations: the one in (122a) is termed \textit{configurationally induced} while the one in (122b) is \textit{functionally induced}. Functionally induced bindlee metavariables are subject to a special condition called Boundedness.

(123) \textbf{Boundedness}

A functionally induced \textit{c-controllee} metavariable $\mathbb{I}$ must be instantiated with the same variable as a \textit{c-controller} metavariable within the immediately dominating $S$, if there is one.

Boundedness, combined with the nonexistence of $[\text{NP}, S]$ traces, and therefore the impossibility of configurationally induced subject binddees, achieves the results of the Fixed Subject Constraint (Bresnan 1972) or the "\textit{that-trace}" filter (Chomsky and L\'esnik 1976). This analysis can be translated into CRG by positing both a configurational but phonetically null anaphor which we call \textit{trace} (and symbolize \(e\)), and a nonconfigurationally represented anaphor \textit{nontrace}. Nontrace will be introduced by the following lexical rule.

(124) Annotate as an \textit{optional} value of the \textit{SUBJ} of a $[+\text{finite}]$ Verb or Modal the anaphoric features of \textit{nontrace}.

Nontrace will then be subject to the CRG version of Boundedness. Boundedness is a kind of locality condition, and we will account for it by assigning nontrace the feature $ [+\text{local}]$.\textsuperscript{11} The positing of nontrace bears on the question of the value of the locality feature for trace and therefore of the treatment of island constraints. If both trace and nontrace had the value $ [+\text{local}]$, with this feature receiving different interpretation in the two cases (Boundedness for nontrace and island constraints for both), the conditions on anaphora would have to make reference to whether or not an anaphor is configurationally realized. It would also require an interpretation for $ [+\text{local}]$ that has no relation whatsoever to the
already established one. (We will see shortly that Boundedness is related to the notion Local Binding Domain defined in §3.3.) Thus, the existence of nontrace provides another argument against assigning trace the value [+local] in order to account for islands. Note that neither Boundedness nor islands involve the notion pseudo-clause, so we cannot appeal to the feature [+local'].

What is the nature of the antecedent (the fronted phrase) in syntactic binding dependencies? The traditional transformational view is that the fronted position of wh phrases and topicalized phrases, which is not a deep structure position, does not bear a grammatical relation to the clause. This view is mirrored in the treatment of syntactic binding in work on Augmented Transition Networks (see Wanner and Maratsos 1978) as well as in GB theory, where it is important to the theory of anaphora (Binding Theory) that the "in COMP" position be an $\hat{A}$ (nonargument) position. However, we will take the position of Kaplan and Bresnan (1982) and Bresnan (1982a) that the relations FOCUS and TOPIC are assigned to fronted phrases. We will treat FOCUS and TOPIC as a special class of Lexical Relations called the Focusing GRs. (The reader is referred to Figure 4.1c for details of our taxonomy of grammatical relations.) They differ from the other Lexical Relations in that they are not assigned to semantic arguments. They are inserted into syntactic argument lists by the following lexical redundancy rule.

(125) Add to the argument list of a V or M an optional argument bearing a Focusing GR.

The categories V and M form a natural class ([+PGR, +OA]) in the feature system proposed in the Appendix.

We represent syntactic binding dependencies just like any other
type of anaphora.

(126) a. 

b.
There are two points to note about the antecedent-anaphor relationship here. First of all, trace and nontrace are bound to nonlogical arguments. Second of all, as (127) (from Bresnan 1978) shows, trace and nontrace in-
herit all grammatical features from their antecedents.

\[ \text{(127) } \]
\begin{align*}
\text{a. Which problem } & \text{ did your professor say she thought was/were unsolvable?} \\
\text{b. Which problems } & \text{ did your professor say she thought were/was unsolvable?}
\end{align*}

These observations lead us to the following conclusion: trace and nontrace are consequents.

Consider now (99) and (107'), repeated here as (128) and (129).

\[ \text{(128) } \]
Mark a nonlogical argument to require a consequent.

\[ \text{(129) } \]
In an argument list with an XCOMP, mark the argument (other than the XCOMP itself) lowest on the Thematic Hierarchy to take a consequent.

After stating (129), we left the status of (128) unclear. Note that although (128) seemed redundant at the time, this is no longer the case.

The FOCUSes and SUBJ in (126) (as well as TOPICS in relative clauses) are not covered by (129) because there is no XCOMP in any of the relevant argument lists. However, (128) applies, as they are all nonlogical arguments. We can restate (128) and (129) to specify the categoriality of the consequent.\(^{12}\)

\[ \text{(130) } \]
\begin{align*}
\text{a. Mark a nonlogical argument in an argument list lacking an XCOMP to require a [+categorial] consequent.} \\
\text{b. In an argument list with an XCOMP, mark the argument (other than the XCOMP itself) lowest on the Thematic Hierarchy to take a [-categorial] consequent.}
\end{align*}

In trees, we will mark consequent-requiring arguments with the notation "\(\alpha \)\( \alpha \)\), where \(\alpha\) is '+' or '-' and stands for the categoriality feature of the consequent.\(^{13}\)

We turn next to Boundedness, the condition governing the binding of nontrace. We can restate it as (131).

\[ \text{(131) } \]
Nontrace must be bound to a coargument if its argument list is associated with a lexical item that heads an \(S\).
The "if" clause aside, this condition matches the core unextended notion of Local Binding Domain, as discussed in §3.3. Furthermore, the "if" clause refers to categorial information, the sort one would expect a categorial anaphor to make reference to. We thus define a third kind of Local Binding Domain and restate the definition of [+local] anaphors.

(132) **Definition**

For an anaphor α, the *Categorial Local Binding Domain* of α is α's LBD if it is associated with an argument in an argument list associated with a lexical item that heads an S; otherwise it is the entire structure in which α appears.

(133) A [+local] anaphor must have an antecedent in its φ Local Binding Domain, where φ = 'Extended' if the anaphor is [-categorial] and 'Categorial' if [+categorial].

Consider (110), the rule introducing PRO, and (124), the rule introducing nontrace, repeated below.

(110) Annotate as an optional value of the SUBJ of a [-finite] verb the anaphoric features of PRO.

(124) Annotate as an optional value of the SUBJ of a [+finite] Verb or Modal the anaphoric features of nontrace.

Rewritten using features, these rules look like (134).

(134) a. Annotate as an optional value of the SUBJ of a [-finite] verb the features [-local, -consequent, +categorial].

b. Annotate as an optional value of the SUBJ of a [+finite] Verb or Modal the features [+local, +consequent, +categorial].

Keep in mind that Modals are always finite, and that in the distinctive feature system discussed in the Appendix, Verbs and Modals form a natural class [+PGR, +OA]. Now note that (134a,b) can be collapsed as (135).

(135) Annotate as an optional value of the SUBJ of an [afinite]element of category [+PGR, +OA] the features [a-local, a-consequent, +categorial].

The significance of (135) is that, if correct, it shows that the restriction of PRO to nonfinite clauses, although observationally true, is some-
what of an illusion. Rather, the SUBJs of all Verbs/Modals are assigned optional categorial anaphors. For some unknown reason, the rule that accomplishes this is an 'a' rule. Determining why this is an 'a' rule would require serious study of syntactic binding and PRO-control cross-linguistically.\(^\text{14}\) (135) casts doubt on the "government" approach to the distribution of PRO (Chomsky 1981).

We conclude with a chart comparing Chomsky's (1982) taxonomy of empty categories with their LFG and CRG equivalents in order to clear up confusion over terminology.

<table>
<thead>
<tr>
<th>GB</th>
<th>LFG</th>
<th>CRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>anaphoric features</td>
<td>description</td>
<td></td>
</tr>
<tr>
<td>[-pronominal]</td>
<td>A-bound trace e</td>
<td>gap:</td>
</tr>
<tr>
<td>[-anaphoric]</td>
<td></td>
<td>-trace</td>
</tr>
<tr>
<td>[-pronominal]</td>
<td>A-bound trace</td>
<td>-nontrace*</td>
</tr>
<tr>
<td>[+anaphoric]</td>
<td>-passive functional control*</td>
<td>Δ*</td>
</tr>
<tr>
<td>[-pronominal]</td>
<td>[-raising]</td>
<td></td>
</tr>
<tr>
<td>[+pronominal]</td>
<td>PRO</td>
<td>Δ*</td>
</tr>
<tr>
<td>[+anaphoric]</td>
<td>PRO*</td>
<td></td>
</tr>
<tr>
<td>[+pronominal]</td>
<td>pro</td>
<td>'PRO'*</td>
</tr>
<tr>
<td>[-anaphoric]</td>
<td></td>
<td>PRO*</td>
</tr>
</tbody>
</table>

*not represented configurationally

Fig. 3.2. Empty (nonphonetic) categories
3.9. Island Constraints

3.9.1. Overview

We turn now to island constraints. It was hypothesized in §3.7.4 that PRO is subject to the same island constraints as are generally associated with gap. A superficial survey may suggest that the match is imperfect at best. (For typographical convenience, PRO, which is not represented configurationally, is placed in parentheses before the verb here. Coreference is indicated by underlining.) The (i) examples are PRO, and (ii) are gap.

(136)  a. Complex NPs

(i)   a. *Max resents the stories about (PRO) working too hard that I always tell.

       β. Max told the story about (PRO) working too hard that I hate.

(ii)  *Who do you resent stories about my seeing e?

b. Coordinate Structures

(i)   *Max discussed both (PRO) reading the book and my getting into trouble.

(ii)  *What do you think I saw e and told Max to get lost?

c. Relative Clauses

(i)   I gave Max a book (PRO) to lend to his sister.

(ii)  *His sister, I gave Max a book that he could lend to e.

d. Indirect Questions

(i)   I asked Max who (PRO) to give the hamster to.

(ii)  *Who did you ask Max what to give to e?
e. Arguments of embedded SUBJs

(i) a. *Max mentioned that (PRO) working too hard is a topic of conversation at the plant.

β. I said that (PRO) working too hard is Max's strong point.

(ii) Who did a picture of e arrive?

We will first dispose of (136a(iβ)) and (136e(iβ)). In the former, the noun stories probably has the argument list {∅, OBL}, with a null Agent. This null argument is bound to Max, and it is this position which serves as the freely assigned antecedent for PRO. On null "specifiers" in NPs, see the brief discussion in Jackendoff (1972, 166). (136e(iβ)) shows that in order for the island constraint to hold, the anaphor must be embedded relative to the intended antecedent: a fact which would not be obvious from simply examining syntactic binding, where the antecedent is always superordinate to the anaphor. We will assume that this constraint uses the notion "c-command." 15

The remaining lack of parallelism between PRO and gap lies in the area of relative clauses and indirect questions. It is interesting to note that it is just these cases which, when examined cross-linguistically, have led to notions of parametric cyclic (bounding) nodes (Rizzi 1978). What we take this to mean is that these "wh islands" are not among the Configurational Constraints. We will first discuss these cases, and then proceed to the Configurational Constraints.

3.9.2. The Role of the Focusing GRs

Erteschik-Shir (1973) (henceforth ES) suggests that island constraints are due to the following constraint on extraction.
(137) Extraction can only occur out of clauses and phrases which can be considered dominant in some context.

The notion of dominance is a semantic/pragmatic one, meaning roughly "not presupposed." (Erteschik-Shir 1981 shows that presupposition is not exactly the right notion here, but it suffices for our purposes.) We will refer to this notion as semantic dominance.

It can easily be shown that the controller-PRO relation is not constrained by (137) (if we replace "extraction ... out of" with "control ... into"). For instance, ES shows that gerunds are usually presupposed and therefore not semantically dominant. Nevertheless, the PRO SUBJ of "subjectless" gerunds can be controlled. Therefore, (137) is not relevant to determining the characteristics of anaphors bearing the feature [+categorial].

ES's evidence for the condition (137) is very convincing. A plausible way to incorporate it into CRG is as a condition on the Focusing GRs. Intuitively, these relations place emphasis on an argument of some clause or phrase. If the "source" is embedded, it makes sense to impose a restriction that it be semantically dominant. We can thus restate (137) as (138b).

(138) a. Definition
For an argument bearing a Focusing GR, the scope of consists of the argument list in which appears, the argument list in which 's consequent appears, and all intervening argument lists.

b. Semantic Dominance Condition
The matrix argument list in the scope of must be interpretable as semantically subordinate to the rest of the scope of .

The notion of scope of Focusing GRs can be used to account for islands. As ES points out, semantic dominance alone would permit extraction out of relative clauses and indirect questions. In fact, there are
languages in which such extractions are grammatical. In these languages (including the Scandinavian languages), the translation of (139) is grammatical.

(139) This is the ice cream that I wonder who likes e.

Note that the argument list of *likes* is in the scope of both the non-overt relative pronoun and the question word. Suppose that UG allows languages to rule out intersecting scope. This would account for *wh* islands in languages that rule them out, as well as for languages which allow extraction out of these structures. There actually seems to be a subparameter in this condition. Rizzi (1978) states that in Italian it is possible to relativize out of a question but not to question out of one. We claim that this is based on the fact that, as proposed by Kaplan and Bresnan (1982), questions involve the relation FOCUS and relative clauses the relation TOPIC. We state the condition as (140), with the part in angle brackets holding in Italian but not in English.

(140) **Intersecting Scope Condition**

For two arguments \( \alpha \) and \( \beta \), the scope of \( \alpha \) may not intersect with the scope of \( \beta \) <if \( \alpha \) and \( \beta \) bear the same grammatical relation>.

(140) is very similar in form and intent to the Accessible Scope Condition of Brame (1978). Note from (126d) that we do not analyze *tough-movement* constructions as involving a TOPIC, unlike Kaplan and Bresnan. Thus, extraction out of such a construction does not violate the Intersecting Scope Condition. This is because only Focusing GRs have scope by our definition (138a). It is well known that at least some extractions out of these constructions are grammatical.
3.1.9.3. Configurational Constraints: A First Approximation

Given the Semantic Dominance Condition and the Intersecting Scope Condition, is there any reason to posit Configurational Constraints? If, as ES claims, semantic dominance can account for all island constraints, there is no need for Configurational Constraints. However, without Configurational Constraints, the putative relationship between gap and PRO remains uncaptured.

There is, in fact, evidence that the conditions in the preceding section are insufficient by themselves.

(141)  
  a. Dave spoke to me about Rebecca yesterday.  
  b. Who did Dave speak to me about e yesterday?  
  c. Dave spoke to me yesterday about Rebecca.  
  d. *Who did Dave speak to me yesterday about e?  

(142)  
      Miriam: Dave spoke to me yesterday about Rebecca.  
      Joe: That's not true. It was about Karen.  

(141) illustrates a paradigm that has been discussed in the literature recently (see, e.g., Hornstein and Weinberg 1980). The problem to be accounted for is the ungrammaticality of (141d). (142) illustrates one of the tests for semantic dominance. The discourse shown here is well-formed; it is possible to negate the PP, so by ES's criterion it can be semantically dominant. Thus, the impossibility of extraction cannot be explained by the Semantic Dominance Condition.

(136) shows that gap and PRO share the following properties:

(143)  
  a. The anaphor may not be in a "complex NP" not containing the antecedent.  
  b. The anaphor may not be in a single conjunct of a coordinate structure if that conjunct does not contain the antecedent.  
  c. *(136e) The anaphor may not be an argument of an embedded SUBJ if c-commanded by the antecedent.
There are two additional properties that apply only to trace, not to PRO or nontrace.

(144) a. Prepositional "objects" may not be trace if the PP is a daughter of a node other than X'.

b. Trace may not appear in [NP,S] position.

(144a) is illustrated by (141); we follow Jackendoff (1977) in considering the postadverbial PP to be a daughter of V". (144b) is discussed by Falk (1983b) and in §3.8 of this dissertation, where it is argued that apparent extraction of subjects does not involve a configurational anaphor (trace). Although the properties in (143) and (144) are English-specific, we will concentrate on them here with the acknowledgement that the resulting Configurational Constraints will be only a first approximation.

How many Configurational Constraints are there? Recent transformational literature (since Chomsky 1977) suggest that there is only one: subjec-
cency. There is, of course, no logical necessity for this. Subjacency is not a very plausible approach to island constraints if the binding of PRO is subject to them, as we have hypothesized. A movement analysis of PRO would make no sense in terms of the assumptions of transformational theory, standard or government/binding, and it is unclear how a subjacent indexing procedure such as that discussed by Bresnan and Grimshaw (1978) could be extended to PRO. We will assume here a group of Configurational Constraints, or CCs. One of these, the Coordinate Structure Constraint of Ross (1967), (143b), will be omitted, as the analysis of coordination within CRG has not yet been worked out. We will be able to collapse (144a,b) into a single constraint, but the other two will remain distinct. Finally, as discussed in §3.7.4, we will add Bresnan's (1982a) f-command condition. Thus, the general structure of the CCs is as in (145).
We will take them in order.

We begin with CC (1). Note that it refers only to trace. Trace differs from PRO and nontrace in being represented configurationally. We begin by considering (144a). The difference between (141b) and (141d) is that the PP headed by about is a daughter of V in the former and V in the latter. V' is what we might call the minimal projection of V — the lowest projection (in the sense of the X-Bar Convention) of V which has daughters which are arguments of V. According to the analysis of Falk (to appear) and §4.2 of this dissertation, the minimal projection in English is 1 for all categories except M, for which it is 3. The first approximation to CC (1) is (146).

(146) Configurational Constraint (first approximation)
A constituent containing a configurational anaphor may not be the daughter of a node higher in the tree than the minimal projection of the category of which it is an argument.

Can we combine (146) with (144b)? Note that [NP,S] is a daughter of the minimal projection of M, which means that even if we were to revise (146) to refer to the anaphor itself instead of just a constituent containing it, (144b) would still not be covered. Furthermore, we do not want to bar all daughters of S from containing configurational anaphors, because Vr is an immediate constituent of S. However, it is the case that whether a sentence has an auxiliary verb or not, [NP,S] always serves as the SUBJ of the highest (nonauxiliary) Verb in the sentence, either directly
or by being the antecedent of a \( \Delta \) SUBJ. In order to capture this fact, we introduce the following definition.

**(147) Definition**

For a syntactic argument \( X \) and a lexical item \( l \), \( X \) is a linked argument of \( l \) if, for some argument \( Y \) of \( l \), \( Y \) is in the transitive closure of the consequent relation with respect to \( X \), and \( X \) and \( Y \) bear the same grammatical relation.

(The requirement that they bear the same grammatical relation is to prevent the upstairs OBJ of "Raising-to-Object" constructions from counting as a linked argument of the downstairs verb; this OBJ can certainly be a trace. The "transitive closure of the consequent relation with respect to \( X \)" means roughly "a consequent (of a consequent \( \ldots \)) of \( X \).") In sentences with a trace in [NP,S] position, although the trace is a daughter of the minimal projection of M, it is higher than the minimal projection of a category of which it is a linked argument (V). Let us restate CC 1.

**(148) Configurational Constraint 1 (second approximation)**

A constituent containing or constituting a configurational anaphor may not be the daughter of a node higher in the tree than the minimal projection of a category of which it is a (linked) argument.

As stated, (148) applies in English. However, the facts in (144) differ in other languages. For instance, there are several Germanic languages which seem to allow trace in [NP,S] position — they violate the "that-trace" filter (Maling and Zaenen 1978) and there is no evidence in these languages for a SUBJ in [NP,VP] position, as one may argue there is in Spanish or Italian. Suppose that in these Germanic languages finite and nonfinite verbs belong to distinct categories, the equivalents of English M and V respectively (as discussed in the Appendix, where the former is called F). This is not unreasonable as they exhibit different syntactic behavior, such as positioning. SUBJs are daughters of S (=F") as in
English. Thus, the minimal projection for $F$ is 3. Now suppose that what matters for CC(1) is not just any category, but rather what we can call a major category, where this is taken not in the traditional sense of N, V, and A (and maybe P or ADV) but rather as a parametric notion defined as (149).

(149) **Definition**

The *major categories* of a language $L$ are those categories to which nominals and finite verbs belong in $L$.

We can state a parametric version of CC(1).

(150) **Configurational Constraint (1) (final version)**

A constituent containing or constituting a configurational anaphor may not be the daughter of a node higher in the tree than the minimal projection of the highest major category of which it is a (linked) argument.

We leave the question of how preposition stranding is to be ruled out in languages which do not permit it at all for future work.16

We turn now to CC(2), the Complex NP Constraint. ES claims that this constraint can be reduced to the Semantic Dominance Condition. She cites the relative acceptability of the sentences in (151), and notes that it is only in cases involving V-N combinations from (152) (we have added *have* and *hopes* to ES's list) that the complement to the noun can be semantically dominant.

(151) a. ?The money which I am making the claim that the company squandered amounts to $400,000.

b. The money which I have hopes that the company will squander amounts to $400,000.

(152) $\begin{array}{ll}
V & N \\
make & claim \\
lodge & complaint \\
advance & hypothesis \\
put forth & motion \\
have & belief \\
hopes & 
\end{array}$
She summarizes:

In each case, there is a derivative verb which expresses the same thing. Thus, *make the claim = claim, advance a hypothesis = hypothesize, put forth the motion = move*, etc. Thus, these verbs differ from combinations such as *discuss the claim* in that they are semantic units for which paraphrases exist...

There is idiosyncratic variation which suggests that all is not as simple and semantic as ES believes. This author's judgments are given in (153).

(153)  

a. ?Who did Max make the claim that I like e?  

b. *Who did Max make a claim that I like e?*  
c. ??Who did Max make claims that I like e?  
d. *What did Max lodge a/the complaint that I gave him e?*  
e. *What did Copernicus advance a/the hypothesis that the Earth revolves around e?*  
f. Who do you have hopes that I will hire e?  
g. *Who do you have a/the hope that I will hire?*  
h. ?What did the ancients have a belief that the planets circle e?  
i. *What did the ancients have the belief that the planets circle e?*  
j. *What did the ancients have beliefs that the planets circle e?*

Our proposal is that cases of apparent extraction from complex NPs actually involve lexical complex verbs *make-the-claim, have-hopes, etc.* The variation in determiners, in particular, points to a lexical treatment. It should be noted that we do not consider picture NPs to be complex NPs. ES's analysis of these in terms of semantic dominance seems correct. Thus, we are interested not in arguments of nouns, but arguments of arguments (of arguments (...)) of nouns. By talking about arguments we are, of course, eliminating relative clauses from the purview of this constraint. They are, of course, already covered by the Intersecting Scope Condition. To formalize CC ②, we make use of the notion *d-command*, introduced in
§3.3. We repeat the definition as (154).

(154) Definition

For argument lists $\alpha$ and $\gamma$, $\gamma$ heading a phrase filling an argument position for some lexical item $A$, and $\beta$ an argument position in $\gamma$, $\alpha \text{ d-commands } \beta$ iff a constituent which corresponds to an argument in $\alpha$ dominates $A$.

We can now define the notion $d$-argument and state the Complex NP Constraint in terms of it.

(155) Definition

$\alpha$ is a $d$-argument of a category $X$ if the argument list of $X$ d-commands the argument list which includes $\alpha$.

(156) Configurational Constraint ②

A $d$-argument of a noun may not be bound to an antecedent outside of the NP that the noun heads.

Configurational Constraint ③ relates to the unextractability of arguments of SUBJs. In some languages, as pointed out by Woolford (1982), this constraint only applies to arguments of sentential SUBJs. The section in angle brackets applies in these languages but not English. There are additional complications with this constraint, but we will ignore them here.

(157) Configurational Constraint ③

An anaphor within a constituent <of category $S$> bearing the relation SUBJ may not be c-commanded by an antecedent outside of that SUBJ.

Finally, we come to CC ④, the f-command condition. This condition was originally proposed by Mohanan (1983b) and Bresnan (1982a) for control of PRO, but it does hold for syntactic binding as well. In LFG it is defined as "c-command at f-structure," a definition that is not open to us because we do not recognize a level of f-structure. However, it translates straightforwardly into CRG. (The maximal functional projec-
tion of a category is the maximal constituent of which it is the logical head. In LFG, it is the maximal constituent corresponding to the clause nucleus of which the category in question is the source of the PRED feature. See §4.2.2 for definition.

(158) **Configurational Constraint**

An antecedent must f-command its anaphor.

(159) **Definition**

For two arguments α and β, we say that α f-commands β iff:

(i) α and β are coarguments; or

(ii) the maximal functional projection of the lexical item of which α is an argument dominates the lexical item of which β’s argument list is a (linked) argument.

Part (i) covers syntactic binding where the antecedent and trace are co-arguments. To illustrate part (ii), consider Bresnan’s (1982a) example (160), which have the syntactic representations (161).

(160) a. People who know Max often discuss working too hard. (Max does not f-command PRO; people who know Max does f-command PRO)

b. Contradicting himself will demonstrate that Mr. Jones is a liar. (Mr. Jones does not f-command PRO)

c. Contradicting himself will discredit Mr. Jones. (Mr. Jones does f-command PRO)
(161) a.

b. on next page

c.
b. Mr. Jones will demonstrate that he is a liar contradicting himself.
(161a) has two conceivable readings: one where the antecedent of PRO is Max (the argument subscripted \( a_j \)) and the other where the antecedent is people who know Max (the argument subscripted \( a_j \)). \( a_j \) is an argument of the verb discuss: its maximal functional projection is the root S, which means that it dominates everything in the sentence, including the lexical item of which \( \beta \)'s argument list is an argument (which happens to be discuss itself). However, \( a_j \) does not meet the requirements of the definition; it is an argument of know, whose maximal functional projection is the \( \bar{S} \) which bears the ADJ relation to the SUBJ NP. This \( \bar{S} \) does not dominate anything associated with \( \beta \). Therefore, it is not a possible antecedent for PRO. In (161b), \( \alpha \) is an argument of the auxiliary is; the maximal functional projection is the \( \bar{S} \). \( \beta \)'s argument list is an argument of the matrix predicate will and a linked argument of the verb demonstrate; neither of these is dominated by the \( \bar{S} \). In (161c), on the other hand, the maximal functional projection of \( \alpha \), the VP, does dominate the lexical item (discredit) of which \( \beta \)'s argument list is a linked argument. In cases like this, LFG achieves f-command by treating the controller and the controllee as formally identical, the LFG equivalent of what is sometimes called multiple attachment; in f-structure, the gerund is literally in both positions. In CRG, however, that route is not open to us, so we make use of the notion linked argument. This may appear to put the CRG definition of f-command at a disadvantage, because it has to be stipulated, whereas in LFG it is automatically the case that "functionally controlled" positions will behave as if they were in two clauses. It turns out, however, that the two definitions can be distinguished empirically, and the CRG version is superior. An OBJ which serves as the antecedent of a Δ SUBJ is not a linked argument by our definition. This
definition was designed ad hoc for CC (1), but it is available to us here as a test of f-command. Consider (162).

(162)  a. Contradicting himself seemed to discredit Mr. Jones.

       b. *I expect contradicting himself to discredit Mr. Jones.

This contrast is predicted by the CRG definition of f-command. In LFG, on the other hand, it is totally mysterious, and, in fact, inexplicable. Consider the f-structures of these two sentences; for expository purposes we place the content shared by the matrix argument and the complement SUBJ in the complement.

(163)  a. 

```plaintext
| SUBJ | [SUBJ_β [PRED 'PRO']
| TENSE | past
| PRED | 'seem<(XCOMP)>(SUBJ)'
| XCOMP | [SUBJ [PRED 'PRO'
| PART | ing
| PRED | 'contradict<(SUBJ)(OBJ)>'
| OBJ | [PRED 'PRO'
| REFL | +
| U | -
| INF | TO
| PRED | 'discredit<(SUBJ)(OBJ)>'
| OBJ_g | [PRED 'Mr. Jones']
```
If "f-command" is "c-command at f-structure" it is obvious that a f-commands β in both structures. As the features of controller and controllee are identical rather than being in an anaphoric relation, it would be impossible in principle to make the appropriate distinction between a. and b. The CRG definition of f-command, although more complicated, has been vindicated.

Further research on other languages is required to determine the adequacy of the statements of the Configurational Constraints, as well as determining which ones are mandated by UG and which ones are optional. For instance, there are languages in which the Complex NP Constraint
(CC 2) does not hold. The versions here are intended merely as first approximations.16

3.9.4. Configurelationality Revisited

As promised in §2.5.5, we will now discuss evidence for a configurational theory based on island constraints, or, more precisely, the Configurational Constraints.

First note that in postulating a class of categorial anaphors we have included as anaphors one class, gaps, which are not considered anaphors in some other frameworks. In GB, A-bound trace, the equivalent of our gap, is treated as an R(eferential)-expression by the theory of anaphora (Binding Theory), while in LFG consequent-binding phenomena (syntactic binding and A-control) are treated as movement-like phenomena distinct from anaphora. One consequence of including gaps as anaphors is the discovery of a hitherto unnoticed generalization involving trace and PRO (135). If this generalization can hold up in some parametric form under cross-linguistic analysis, it will provide an interesting piece of evidence for our approach to anaphora. The interesting consequence, from the point of view of configurelationality, is that PRO-control, which is and has to be treated relationally in LFG because PRO is not represented configurationally, must generalize with syntactic binding, which is generally thought of as the epitome of a configurational phenomenon. In fact, we reject Bresnan's (1978) distinction between function-dependent and structure-dependent rules: while there are purely function- (or relation-) dependent rules, viz. the lexical ones, rules that are apparently structure- (or configuration-) dependent will, on closer examination, reveal a relational side as well.
The Configurational Constraints themselves are configurational. This fact by itself is not incompatible with Lexical-Functional Grammar. In LFG, the link between fronted phrases and gaps is achieved as part of the mapping between c-structure and f-structure, rather than at f-structure itself. Island constraints are accounted for by means of boxing categories in the various phrase structure rules (Kaplan and Bresnan 1982). The boxed categories represent the roots of islands (called bounding nodes as in Chomsky 1977, 1981), and a syntactic binding path is not allowed to cross a boxed node. (As in Chomsky's work, this can be overcome by linking syntactic binding domains in a formalism strongly reminiscent of COMP-to-COMP movement.) However, some aspects of the CCs cannot be handled by the box notation, and it is unclear how the formalism could be extended to capture the relevant facts and be configurational. In addition, the susceptibility of PRO to island constraints is a mystery in LFG.

We begin by repeating the Configurational Constraints.

(164) Configurational Constraints

1. A constituent containing or consisting of a configurational anaphor may not be the daughter of a node higher in the tree than the minimal projection of the highest major category of which it is a (linked) argument.

2. A d-argument of a noun may not be bound to an antecedent outside of the NP that the noun heads.

3. An anaphor within a constituent <of category S> bearing the relation SUBJ may not be c-commanded by an antecedent outside of that SUBJ.

4. An antecedent must f-command its anaphor

CC 1 is a complicated condition, which unifies illicit "extraction" from PPs in languages which do not rule out preposition stranding completely and the impossibility of [NP,S] traces. As stated, it refers to a
configurational notion (minimal projection, which is based on bar-level)
and a relational notion (linked argument, which refers not only to argu-
menthood, but to sameness of grammatical relations). Thus, it is configure-
relational. For justification of this statement, see discussion in §3.9.3.
LFG can handle the configurelationality with respect to PPs, but it cannot
even state the condition on [NP,S]. For PPs, LFG could state phrase
structure rules of the form (165).

(165) $XP \rightarrow X' \ldots [PP] \ldots$

A phrase structure redundancy rule (metarule) can be stated.

(166) Box any PP in the phrase structure rule expanding a projection
of any category higher than its minimal projection.

However, in the case of [NP,S], boxing the NP node would not prevent
extraction of the entire node.

CC $\textcircled{2}$, the Complex NP Constraint, is also configurelational. It
requires reference to category (N) and argumenthood. Argumenthood is
necessary if we want to exclude relative clauses, which is desirable because
there are languages in which complex NPs in our sense are islands but
relative clauses are not. The need for reference to the category N
should be self-evident. Here we see clearly that LFG’s formalism is too
restrictive with respect to islands and configuralationality. We can
mimic CC $\textcircled{2}$ with the phrase structure rule (167).

(167) $N' \rightarrow N \quad PP \quad [XP] \quad [S]$

$(+PCASE)\Rightarrow (+XCOMP)\Rightarrow (+COMP)\Rightarrow$

(167) seems to accomplish the required results, but there is one case which
is covered by CC $\textcircled{2}$ but not (167).
(168)  
a. *Who did Max read a story about helping?

b. Max read a story about helping hamsters. (PRO ≠ Max)

(We include (168b) to show that this is due to a Configurational Constraint, not the Semantic Dominance Condition.) The problem in (168a) is that trace is in the PCASE (or OBL) PP, which is not usually an island. The "object" NP in PP is also not usually an island.

(169) Who did Max talk about helping?

In CRG terms, however, the trace in (168a) is a d-argument of the noun story. The LFG formalism is too local; it only allows reference to mother and daughter nodes. Thus even in the Configurational Constraints, where LFG does allow some configurelationality, it does not allow enough.

We have not attempted to justify the formulation of CC 3, so the fact that it refers both to SUBJ and, in some languages, S may or may not be relevant. CC 4 is not configurelational.

In summary, CC 1, CC 2, and possibly CC 3 provide us with more evidence that a configurelational theory of grammar is desirable, as does the notion of a class of categorial anaphors that includes not only gap but also PRO.
Footnotes for Chapter 3

1 Haj Ross has pointed out to me (personal communication) that sentences like (i) and (ii) are prima facie counterexamples to the claim that personal pronouns need not have antecedents.

(i) Max took his hamster with him.
(ii) I took him at his word.

However, the obligatory coreference in these cases is probably due to the semantics of the constructions rather than to grammatical requirements. Note that Max and him cannot be coreferential in (iii).

(iii) Max sent his hamster with him.

2 In some cases the antecedent must c-command the anaphor.

3 Our approach differs from that of Lasnik (1976), who argues for a rule of Disjoint Reference rather than rules marking anaphors coreferential with antecedents. The approach we have adopted is easier to formalize in our notation. It is not clear that there is any real difference between the two approaches.

4 A cautionary note is in order here: we are following Yang's description of the Dutch reflexive zich here although it may not be correct. Luigi Rizzi has pointed out to me (personal communication) that bare infinitives in Dutch are associated with properties pointing to a "verb raising" or "clause union" analysis, and that the use of zich may be related to this. We will retain Yang's description here, partially because we do not have an analysis of clause union constructions. For instance, it is possible that the presence of an S node is relevant for other aspects of clause union. The notion of pseudo-clause, which we will introduce later in this section, could very well be a part of the definition of clause union constructions. Just calling bare infinitives clause union constructions does not really explain the use of zich as opposed to zichzelf.

5 Simpson and Bresnan (1983) have proposed a feature [\texttt{[obj]}] to account for the system of obviative complementizers in Warlpiri. I leave as an open question the issue of its necessity.

6 Whether or not case is "transmitted" by consequents seems to be an area of parametric variation. Consider the following data from a very common American dialect, including this author's speech.

(i) a. Who did you see?
   b. *Whom did you see?
(ii) a. Who was Miriam speaking with?
   b. *Whom was Miriam speaking with?
(iii) a. *With who was Miriam speaking?
   b. With whom was Miriam speaking?

The b. examples in (i) and (ii), which are considered "correct" by prescriptivist standards, are very odd in this dialect. It is only when the whom word itself is in an accusative case position that it is marked with this case. Lexical-Functional Grammar cannot account for these facts, and neither can Government/Binding theory if there is any relation between
abstract Case and morphological case. On the assumption that the trace of syntactic binding constructions is a consequent, which we will argue it is, this dialect of English provides evidence against universally requiring consequents to transmit case.

The fact that subcategorization is actually configurational (see §2.2) does not really change the argument. Even in CRG, the basic subcategorization is relational; categorial information is secondary. In fact the only evidence we found for categorial subcategorization was among XCOMP's, which may mean that one is dealing with an XCOMP before allowing categorial selection. Whether this is universally true or not I do not know. I know of no theory that treats VP as a category distinct from both S and VP. Since we saw in §2.2 that we need to distinguish VP from categorically in order to achieve categorial selection, and we also saw in §2.1 that VP patterns with S in terms of positioning, we conclude that our decision to treat VP as a type of S is correct. If this is the case, then to-infinitives cannot be distinguished categorially from full complement clauses.

We are glossing over the question of "functional control" in non-arguments (XADJs). Mohanan (1983b) shows that choice of controller in these cases is determined by grammatical relations, not thematic relations, an unsurprising fact given our lexical thematic analysis.

It is unclear how seriously this argument should be taken because Malayalam has no generic pronouns and yet PRO has a generic interpretation, as in English (Mohanan 1983b).

I am grateful to Joan Bresnan (personal communication) for this argument.

One disadvantage to the CRG translation of Falk's (1983b) LFG account is that the latter extends very elegantly to certain across-the-board extraction facts but the former does not.

The mention of XCOMP in both (130a) and (130b) is a disadvantage of this proposal. I think that it is really not a problem. In order for the formulations to be cleaned up, a more complete account of the organization of the lexical component is needed, presumably based on the level-ordered or stratified lexicon and Elsewhere Condition of Kiparsky (1982) and Mohanan (1982). Such a theory could place the two rules in an order such that both do not need to mention XCOMP.

Specifying the categoriality of the consequent may be too simplistic. Zaenen, Engdahl, and Maling (1981) show that in Scandinavian languages resumptive pronouns can serve as the anaphors in syntactic binding but that, unlike traces, they are not subject to island constraints.

An out-in-left-field conjecture: If it is true that "nonconfigurational" languages do not exhibit syntactic binding, as has been claimed on occasion, it is possible to construct an explanation on the basis of (135). Most nonconfigurational languages allow free "pro-drop" with all grammatical relations. In LFG and CRG, pro-drop is represented as having a
PRO argument: in these languages, (135) would specify PRO (i.e. [-local, -consequent, +categorial]) for both finite and nonfinite clauses and for both SUBJs and nonSUBJs. Configurational traces are probably ruled out in these languages for processing reasons: it would be much harder to posit them when processing a sentence than it is in English; in English, if a transitive verb is not followed by an NP we posit a trace, but in languages with very free order this would be much harder. Thus, there is no source for a categorial consequent in such a language, and therefore no syntactic binding.

Haj Ross has pointed out (personal communication) that control into SUBJs is not always prohibited.

(i) Max said that getting arrested was like being enslaved by a hamster.

Note that Max controls both the PRO in the SUBJ and in the OBJ of like. The sentence in (ii), with control only into the SUBJ, is ungrammatical.

(ii) *Max said that getting arrested was like Miriam's studying linguistics.

This is what we predict, unlike the unexpected (i). Now note (iii), where what is being attempted is control into the OBJ of like only.

(iii) *Max said that Miriam's getting arrested was like being enslaved by a hamster.

It seems that there is an across-the-board condition of some sort operative here. Across-the-board conditions are usually associated with syntactic binding. With that in mind, note the ungrammaticality of (iv).

(iv) *What did Max say that my picture was like a story about?

I have no explanation for these facts, but it is significant that PRO-control and syntactic binding both seem to behave unexpectedly. This confirms our decision to treat PRO and trace as more similar than they have generally been considered.

It is actually possible that the preposition stranding constraint is not a CC at all. For instance, it is possible that languages that do not allow preposition stranding have a phonological condition that prepositions must cliticize onto the following NP.

Note that in Government/Binding theory, the S-structures of (162a,b) are (i).

(i) a. 

```
S
   /\  
 INFL
    /\  
 NP  VP
     /\  
 NP  VP
      /\  
 PRO V NP
       /\  
 (contradicting himself)  seemed NP
       \   S
        \   /\  
         \ e to V NP
          /\  
           discredited Mr. Jones
```
b.

The gerund is a constituent of the lower clause in (ib). Although GB has thus far not addressed the f-command condition, it would seem to predict that if there is an asymmetry between Raising-to-Subject and Raising-to-Object that it would go the other way. The contrast in (162) seems to me to be a real problem for both LFG and GB.

Another possible similarity between syntactic binding and PRO-control that could be formalized as a Configurational Constraint is a requirement to be bound to the closest possible antecedent. It is well known that one may not pass over a possible antecedent for PRO and pick one in a higher clause. Similarly, one may not pass over one consequent-requiring element and pick a higher one in finding an antecedent for trace, saving the lower one for a later trace. On this latter case, often referred to as a constraint against crossing dependencies, see Bach (1977), Kaplan and Bresnan (1982), and Maling and Zaenen (1982).

We call them Configurational Constraints despite this because they have a strong configurational component and they define structural (or, informally, configurational) prohibitions.
CHAPTER FOUR: AN EXPLICIT THEORY OF CONFIGURELATIONAL GRAMMAR

In this chapter, we make explicit the theoretical foundations of the theory. We will discuss the theories of grammatical relations (§4.1) and grammatical configurations (§4.2), as well as the lexicon, which was outlined briefly in Chapter 1 (§4.3).

The framework to be presented in this chapter is a nonderivational one. There is a single level of configurelational structure. There are no rules mapping constituent structures into constituent structures, as in transformational grammar. There are not multiple strata of grammatical relations, as in Relational Grammar. We do not recognize syntactic meta-rules which map phrase structure rules into phrase structure rules, as in Generalized Phrase Structure Grammar and sometimes in Lexical-Functional Grammar, either,

4.1. Grammatical Relations

At the heart of any relationally-based theory is an inventory of grammatical relations. While Government/Binding theory only provides a handful of "Cases" and Marantz (1981) only recognizes four grammatical relations (SUB, OBJ, ARG, MOD), most theories recognize a larger set.

What we will do is present the primitive grammatical relations being assumed in this dissertation, and compare them with the relations that have been proposed in Relational Grammar (Perlmutter and Postal 1983b) and Lexical-Functional Grammar (Bresnan 1982a). We will then discuss the similarities and differences. The three inventories of grammatical relations are presented in Figure 4.1. To emphasize the similarities, Figure 4.2 compares the terminology used in the three theories.
a. *Relational Grammar* (Perlmutter and Postal 1983b)

```
R-signs
Nominal R-signs ...
```

```
Central R-signs Overlay R-signs
```

```
Core R-signs Retirement R-signs Q'
```

```
Term R-signs Oblique R-signs CHO[meur] TOP
```

```
Nuclear Term Object BEN
R-signs R-signs INST TEMPE
1 3 LOC
```

```
[sub]ject][2][indirect object] ...
[direct object]
```

b. *Lexical-Functional Grammar* (Bresnan 1982a)

```
Grammatical Functions
```

```
Subcategorizable Nonsubcategorizable
```

```
semantically unrestricted semantically restricted ADJ[unct]
```

```
SUBJ OBJ OBJ2
```

```
OBL COMP XCOMP XADJ[unct]
```

```
FOCUS* TOPIC*
```

*The subcategorizability of these functions is a parameter that distinguishes "subject-oriented" from "topic-oriented" languages.

c. *Configurelational Grammar*

```
Grammatical Relations
```

```
Lexical Relations Syntactic Relations
```

```
Core GRs Focusing GRs ADJ[unct]
```

```
Pure GRs Impure GRs TOPIC XADJ[unct]
```

```
SUBJ Oblique GRs Propositional GRs EXCL[amative]
```

```
OBJ OBJ2 OBL COMP POSS XCOMP
```

Fig. 4.1. Primitive grammatical relations in various frameworks
The general role of grammatical relations is to mediate between configurational positions and semantic roles. The distinction between what we are calling Lexical Relations and Syntactic Relations represents a difference in the way this role is realized. With the Syntactic Relations, there is a direct interpretation for each relation, or for each combination of relation and configurational position (see §2.3): ADJs and XADJs are modifiers (XADJs are subject to A-control), and SPECs are grammatical markers such as determiners. The Lexical Relations, on the other hand, can only be linked to semantic roles through the mediation of
lexical entries of predicates. For instance, a sentence may have any number of ADJs, but subcategorized oblique arguments are unique (see Bresnan's 1980 discussion of instrumentals, overviewed in §2.5.4 of this dissertation). As relation-changing rules are lexical in LFG and CRG, and therefore involve the Lexical Relations, it is not surprising that RG has not dealt with the nature of Syntactic Relations.¹

Within the general category of Lexical Relations, there are several subcategories. We will restrict our attention here to the Core GRs, ignoring the Focusing GRs, which were discussed in §3.8. The primary distinction is between Pure and Impure GRs.² This is a very important division: the Impure GRs may not be assigned to just any semantic argument. They are restricted to semantic types: the obliques by thematic relations and the Propositional GRs by being limited to propositional and predicative arguments. The relation POSS is argued for convincingly by Rappaport (1980), who shows that possessive NPs are not SUBJs, partially on the basis of observations by Amritavalli (1980). COMP is the relation assigned to "object" clauses which, at least in some languages, do not exhibit any OBJ behavior. XCOMPs are Α-controlled; see §3.7 and, in the formalism of LFG, Bresnan 1982a.

There is a slight difference between our representation of oblique PP arguments and that of LFG and RG. The latter two mark oblique arguments with the thematic relations that they bear (represented by the 'θ' of OBLθ in Figure 4.1b); we will not. It is unclear how to understand the representation of thematic relations in the syntax in a framework which holds to the idea of autonomous levels of syntax and semantics. Furthermore, we will not be able to avail ourselves of Bresnan's (1982a) and Kaplan and Bresnan's (1982) notation "(↑(+PCASE))=↑" which allows a lex-
ical feature of the preposition to percolate up and become the name of
the grammatical relation borne by the PP. Simply labeling all obliques
OBL poses a potential problem for the notion of relational uniqueness
(Kaplan and Bresnan's 1982 Uniqueness (or Consistency) Condition;
Perlmutter and Postal's 1983b Stratal Uniqueness Law) because predicates
can take more than one argument bearing the relation OBL. We will ac­
count for this problem formally by subscripting each occurrence of OBL
with a positive integer; each rule introducing an OBL can subscript a
Greek letter as a variable ranging over the set of positive integers.
In a sense, what we are claiming is that there is an infinite number of
OBL relations.

The Pure GRs are SUBJ, OBJ, and OBJ2. These are semantically un­
restricted in that they may be linked to arguments bearing any thematic
relation, or even no thematic relation at all. There are several ways
to identify Pure GRs. For instance, in English and many other langu­
ages, only SUBJs, OBJs, and OBJ2s can be \$-controllers in verb-headed
structures (see §3.7.2 and Bresnan 1982a). In addition, agreement in
many languages is determined on the basis of grammatical relations. It
is unclear at present whether universal tests for subjecthood and object­
hood exist, but, as Rappaport (1983) shows, it is possible to argue for
or against Pure-GR-hood language-specifically.

With regard to the Pure GRs, it should be noted that we are taking
the approach of Bresnan (1982a) as opposed to Perlmutter and Postal (1983b)
with regard to objects. According to Perlmutter and Postal, the NP im­
mediately following a verb is the (surface) direct object (2), the NP of
a to-headed PP is the indirect object (3), and the second NP in a double
object construction (the "logical direct object") is a 2-chômeur (2).
According to Bresnan, these three NPs are OBJ, OBL<sub>Goal</sub>, and OBJ2, respectively; i.e. the NP "object" of to does not bear a Pure GR in the LFG view while the second NP of double object constructions does. The categorial syntax (in English) favors the LFG view, with PPs dominated by VP never representing Pure GRs. There is also some evidence for the LFG view in the RG literature. For instance, Perlmutter and Postal (in press) point out that many languages mark both the 2 and 2 with accusative case: they create the term acting 2 to cover both relations. As Perlmutter (1982) points out, although the existence of a derived relation acting 2 would seem to predict the existence of acting 1, the evidence for the latter is much weaker. This is strange, considering the pervasiveness of acting 2 cross-linguistically. Assuming the LFG view of object relations, the generalization is that objects (OBJs and OBJ2s) are accusative. On the other hand, Perlmutter (1979) shows that 2s and 3s pattern together in Italian with respect to the floating of quantifiers from cliticized arguments: these two relations do float quantifiers while obliques do not. Of course, Italian lacks the double object construction, so it is possible that the RG view is correct for Italian.

Marantz (1982b) states that it is preferable to have a theory which recognizes only one grammatical relation in the absence of contrary evidence, or at least as few as one can get away with. This claim is interesting inasmuch as it echoes a familiar argument from Generative Semantics concerning grammatical categories, an argument that led to the reduction of the categories of natural language to nouns, verbs, NPs, and Ss. This argument was never accepted by Extended Standard Theorists, because it led to inelegance in the treatment of items traditionally thought to belong to other categories. Similarly, the reduction of grammatical
relations to one or two is not a priori preferable; if a more elegant theory can be obtained, this reduction is motivated, but otherwise it is not.

To the extent that work in RG, LFG, and CRG supports the existence of other relations, Marantz's argument is unconvincing. Much of the evidence that has been amassed bearing on grammatical relations is in the area of relation-changing rules like passivization. These are lexical rules, so we have not been able to examine them in this dissertation (although we did summarize the argument from passivization for distinguishing OBJ and COMP in §2.2. However, configurelationally-stated ordering rules (§2.1), the inclusion of a reference to two elements bearing the same Focusing GR (§3.9.1), and other rules and conditions, can be taken as evidence that a rich set of grammatical relations is desirable.

Unfortunately, most of the evidence bearing on the identification of grammatical relations comes from lexical rules. Therefore, the adequacy of the details of Figure 4.1c has to be left unproven. In particular,

until relation-changing rules are studied in detail from a lexicalist perspective (as in Baker 1983a, 1983b) we will leave the existence of "retirement" GRs (especially CHO) open.

4.2. Grammatical Configurations

4.2.1. Grammatical Categories

The theory of grammatical configurations is a version of the X-Bar Convention (Chomsky 1970; Jackendoff 1977; Bresnan 1982a; Falk 1983a; etc.). According to the Convention, syntactic categories are projected from lexical categories, which are realizations of bundles of distinctive features. Following Jackendoff (1977), we assume a "uniform three-level" system in
which the maximal projection of every category is 3. We will use the "prime" notation, in which the $n$th projection of a category $X$ is represented as $X^{n}$ superscripted with $n$ primes.

For most categories, the identification of lexical and maximal syntactic categories is fairly uncontroversial. For instance, $NP$ is the maximal projection of $N$ ($N''$), $AP$ is $A''$, $ADVP$ is $ADV''$, and $PP$ is $P''$, in all standard views within the X-Bar tradition. We will also follow Jackendoff (1977) in considering "minor" categories like $DET$, $Q[uantifier]$, $P[a]RT[icle]$, and others to have full projection provided by Universal Grammar. In many cases, this results in superfluous levels of structure, but as Jackendoff shows there are sufficiently interesting generalizations of phrase structure which can only be expressed by assuming this extra structure. If we assume, as suggested by Chomsky (1981), that no rule is required to specify that $X^n$ will have a daughter $X^{n-1}$, there is no cost to the grammar in the way of additional rules associated with this superfluous structure.

For the verbal and sentential system, we have adopted the following analysis: in English, $VP$ is $V''$, $S$ is $M[odal]''$, and $\tilde{S}$ is $C[omplementizer]''$. The former two choices are defended by Falk (1979, to appear), the former in a standard transformational framework and the latter in Lexical-Functional Grammar. One apparent problem with taking $VP$ to be the maximal projection of $V$ and $M$ to be the head of $S$ is that $VP$s and $Ms$ do not appear in all languages. As to the former, it is possible that UG provides a hierarchy based on category features for determining what counts as an $S$, and that what is commonly referred to as a Verb in language $X$ may not have exactly the same feature decomposition as a Verb in Language $\tilde{X}$. As to the latter, it is suggested in the Appendix that $Ms$, or $M$-like categories,
may appear in other languages, comprising finite verbs in other Germanic languages, for instance. The identification of Complementizer as the head of \( \tilde{S} \) is not supported by an explicit analysis here, but there are some indications that it is correct. First of all, it has been suggested by other researchers, such as Fassi Fehri (1980) in a lexicalist analysis of Arabic and Stowell (1981) in Government/Binding theory. Secondly, there is evidence that \( S \) and \( \tilde{S} \) are different categories. There are, for instance, syntactic positions in which one but not the other can appear. The definition of Categorial Local Binding Domain (§3.8) crucially depends on \( S \) and \( \tilde{S} \) being distinct in order to derive the Fixed Subject Constraint ("that-trace" filter) effects. Finally, it allows \( \tilde{S}s \) without constituent \( Ss \), which in turn permits a flat for-NP-VP analysis for for-infinitives and allows \( \overline{VP} \) (to-infinitives) to be analyzed as being \( \tilde{S} \), as we have done in this dissertation. The flat structure for for-infinitives has been argued for by Emonds (1976) (although he derives it transformationally from a for-S deep structure) and is necessary if \( S \) is a projection of \( S \). Phrase structurally, \( \overline{VPs} \) act like \( \tilde{S} \); for instance, they appear constituent finally (see §2.1). Treating \( \overline{VP} \) as \( S \)-less \( \tilde{S} \) enables us to capture the best of both the lexicalist VP analysis (Brame 1976, Bresnan 1978) and the transformational \( \tilde{S} \) analysis (Koster and May 1982).

In terms of general questions of constituency, we have adopted the general conclusions of Jackendoff (1977). except that the presence of an extra level in VP allows for more possibilities, as discussed in Falk (1983a) with respect to adverbs. In addition, following the evidence in §2.1, we adopt the following positions for fronted \( wh \) phrases and topicalized phrases.
(1)

In addition to the evidence in §2.1, note that this accounts for the fact that in Middle English, where \textit{wh} phrases could cooccur with overt complementizers, they preceded the complementizer, while even in Modern English, topicalized phrases can cooccur with complementizers and follow them.

Integral to the X-Bar Convention is the notion of categorial distinctive features. We have made use of a new feature system at various points in Chapters 2 and 3. Because of the length of the arguments for this system, discussion is deferred until the Appendix.

4.2.2. Exocentricity

A major problem for the theory of grammatical configurations is the status of apparently exocentric constructions, i.e. constructions of some category X which are headed by an item of category Y. The first problem is to determine whether they should be accepted at face value or explained away by reanalysis. Ultimately, in terms of the X-Bar Convention, this boils down to whether the head of a synaactic constituent is introduced by an explicit rule of grammar (as in standard transformational grammar as well as Jackendoff 1977, Bresnan 1982a, Falk 1983a) or by a convention of Universal Grammar (Chomsky 1981, Stowell 1981). If it is a rule, there is no reason for the head not to be as optional as any other element as
long as a coherent semantic interpretation can be constructed. Under this view, the apparent obligatoriness of heads is due to the fact that the head is usually central to the interpretation of the phrase. This view is adopted, for example, by Falk (to appear), who argues that headless Ss and VPs should be allowed by the theory of grammar. It is pointed out there that more orthodox versions of the X-Bar Convention (in particular Jackendoff 1977) allow phonologically null heads in some constructions, which are there only for the semantic interpretation. On the other hand, if the projection from lexical category to syntactic category is automatic and not rule governed, we should not expect to find exocentric categories. We will take the latter view here for two reasons: (a) in a system of constituency rules such as that proposed by Falk (1983a), the rule introducing the head differs from all other rules in not introducing a maximal projection; and (b) the X-Bar Convention claims that syntactic categories like NP are not primitive but rather derived from lexical categories like N, which is senseless if all NPs are not headed by N.

What evidence is there for exocentric constructions? Bresnan (1982a) cites the constructions in (2).

\[
\begin{align*}
\text{(2) a.} & \quad \text{gerunds (NP headed by V)} \\
\text{b.} & \quad \text{sentences (S headed by V)} \\
\text{c.} & \quad \text{oblique argument PPs (PP headed by N(?))}
\end{align*}
\]

To these, we add (3).

\[
\begin{align*}
\text{(3) a.} & \quad \text{verbless VPs (VP without any internal head)} \\
\text{b.} & \quad \text{COMPless Ss (S with no C; headed by S (?))} \\
\text{c.} & \quad \text{other cases of } S (\text{S headed by } S \text{ with semantically empty COMP})
\end{align*}
\]

Bresnan's (2b) needs to be clarified in terms of our conception of the categorial status of Ss. In our framework, any S with a finite auxili-
ary is headed. However, two subcases of (2b) remain.

(2') b. (i) auxiliaryless sentences
(ii) sentences with the dummy auxiliary do

The cases in (2) and (2') break down into two subcategories: a. and b(i) involve truly headless expressions while b(ii) and c. involve constructions whose heads are configurationally present but do not function as heads semantically. In this latter type, there is what would normally be an argument (the OBJ NP in PP, the XCOMP VP in S) which serves as what Falk (to appear) calls a functional head. These are the cases that we have represented with the relational label HEAD. We leave it as an open question whether functional heads also bear a real grammatical relation to the configurational head, as we are unaware of any evidence that bears on this. In case (3c), which probably includes 5s headed by * that* and *to*, the HEAD "relation" can again be exploited.

The other subcategory of (2) and (2') represents a more serious problem for the theory of grammatical configurations. If there are truly exocentric structures then we cannot define syntactic categories in terms of lexical ones. If, as stated above, we adopt the approach of Chomsky (1981) and assume the entire X-Bar system of a given category to be supplied automatically without any language specific rule, there cannot be any truly exocentric categories. (This argument is made in detail by Stowell 1981.) Thus, the theory of grammatical configurations forces us to search for an endocentric account of these structures. Consider gerunds. The morpheme *ing* which turns verbs into gerunds is also a nominalizing suffix, forming what Chomsky (1970) calls "mixed nominals." It is probably not a synchronic coincidence that the suffixes are identical. The intuition we would like to capture is that *ing* (although not the *ing*
of participles) is a nominal suffix, but for some reason it is more nominalizing in some cases that in others, in that it sometimes creates nouns and other times it creates verbs which head NPs. On the assumption, not unreasonable, that category labels are subscripted with morphosyntactic features (such as case, finiteness, etc.), gerunds will represent the category V\_X, where \( X \) represents the formal morphosyntactic features of \( \text{ing} \). What does it mean to say that gerundive \( \text{ing} \) is a nominalizing suffix? One possibility is that \( X \) is, in fact, identical to the category label N. That is, gerunds are represented categorially as V\_N. We now adopt the following projection convention.

(4) For a lexical item of category \( X_Y \), where \( X \) and \( Y \) are both bundles of syntactic category features, the lexical item in a syntactic structure projects to \( X_Y', X_Y'', X_Y''' \), and then one further level to \( Y''' \).

We also adopt the following definitions.

(5) **Definitions**

a. The syntactic categories \( X', X'', X''' \), built over a lexical item of category \( X \) will be called *projections* of \( X \). \( X''' \) is the *maximal projection*.

b. The syntactic category \( Y''' \) built over \( X_Y'''' \) is called the *top projection* of \( X_Y \).

c. For a category \( X \) in the transitive closure of the relation HEAD with respect to a category \( Y \), we say that \( Y \) is a *functional projection* of \( X \). The maximal projection of a functional projection of \( X \) which is not itself the functional head of another category is the *maximal functional projection* of \( X \).

("A member of the transitive closure of the relation HEAD with respect to \( Y \)" means the HEAD (of the HEAD (...) of \( Y \).) The maximal projection of a gerund is the VP, and its top projection is NP. We can adopt a similar analysis for auxiliaryless Ss, taking the features of \( M \) to be the morphosyntactic features of finite inflection. If we take \( M \) to be a historical development of a category in Germanic representing finite verbs, as sug-
gested in the Appendix (§A.4.2), this is not implausible; in fact, it is very similar to making INFL[ection] the head of S. Finite verbs project to VP (maximal projection) and S (top projection).

We turn now to (3a,b). (3b), the complementizerless S, is actually rare; in general we consider COMPless clauses to be S (M'') not S (C''). One type of complementizerless S that is recognized in the present theory is the clause in which a preposed who phrase appears in English and similar languages. We will assume that the S in these cases is marked HEAD. However, we have no explanation for the lack of a lexical complementizer. It cannot be due to a universal constraint, because "doubly filled COMPs" were allowed in earlier stages of English. We will leave this for future research. We will treat matrix clauses, which also seem to be Ss without complementizers, as a special category: R[oot], not C''. The reasons for this are that (a) overt complementizers do not occur in matrix clauses, rendering an analysis of them as maximal projections of C questionable, and (b) there are, cross-linguistically, many examples of phenomena which are limited to matrix clauses (e.g. placement of the finite verb in German, left dislocation in many languages (including English), etc.) and having a special root node enables syntactic rules to refer to matrix clauses. What about verbless VPs? The existence of such structures in English is proposed by Falk (to appear) in relation to "main verb" be, which is claimed to be a member of the category Modal, with a non-VP XCOMP. The structure proposed there can be represented in CRG terms as (6).
Furthermore, Kaplan, Brenan, Peters, and Zaenen (1982) discuss a construction in Dutch in which VP XCOMPs appear without their verbs and all the verbs are stacked at the end of the sentence. Any surface theory which does not recognize discontinuous constituents has to recognize verbless VPs, at least in the Dutch case. We suggest that another special category like R is involved, one we call ARG[uments]. ARG has the same constituency as a VP but with no head V.

We can now provide an explicit syntactic representation, as opposed to the abbreviated ones in Chapters 2 and 3. For instance, the sentence *Miriam gave hamsters to Max* is represented as (7). (7a) is the complete representation of the sentence. The structure in (7b) represents a slightly abbreviated form which omits some nonbranching nodes. We will utilize these abbreviated representations in the remainder of this dissertation. Essentially, aside from maximal projections and top projections, we omit all nonbranching nodes which are dominated by nonbranching nodes. We will, however, include lexical category nodes unless there are no branches at all in the entire projection system of the lexical category in question, as with the NPs in (7).
The problem of exocentricity is a difficult one, and parts of our solution are better than others. We will maintain the approach of this section for concreteness, but acknowledge that it is highly sketchy. To
pursue it would take us too far afield.

Note that our approach to gerunds and auxiliaryless Ss can help us with a problem we noted in the discussion of the Dutch pseudo-reflexive zich (§3.4). Basing our analysis on Yang (1983), we noted that the S (C") seems to serve as the Dutch pseudo-clause. However, this analysis is possible only if a te-infinitive counts as a C"'. Unlike English to, te is a part of the verb and thus cannot be seen as a member of the Complementizer category. In the present theory, however, it is possible to treat a clause headed by a te-infinitive as a C", salvaging our analysis of zich. We consider the infinitive itself to be a member of the category V_C. It thus projects to V_C" and then C"'. On the assumption that finite verbs in Dutch belong to the category F, related to English M, as discussed in the Appendix, the syntactic representations of the sentences in (39) of Chapter 3, from Yang (1983); is (8). (Note that zich is [-local], so it is not coreferential with Karel in either sentence.)

(8) a. 

\[
\begin{array}{c}
\text{R} \\
\langle \text{HEAD} \rangle \\
\langle \text{SUBJ} \rangle \\
\langle \text{OBJ} \rangle \\
\langle \text{OBL} \rangle \\
\langle \text{XCOMP} \rangle \\
\end{array}
\]

\[
\begin{array}{c}
\langle \text{F} \rangle \\
\langle \text{N} \rangle \\
\langle \text{F} \rangle \\
\langle \text{ARG} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \\text{pseudo} \rangle \\
\langle \\text{over} \rangle \\
\langle \text{zich} \rangle \\
\end{array}
\]

\[
\begin{array}{c}
\langle \text{F} \rangle \\
\langle \text{N} \rangle \\
\langle \text{F} \rangle \\
\langle \text{ARG} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{pseudo} \rangle \\
\langle \text{over} \rangle \\
\langle \text{zich} \rangle \\
\end{array}
\]

\[
\begin{array}{c}
\langle \text{F} \rangle \\
\langle \text{N} \rangle \\
\langle \text{F} \rangle \\
\langle \text{ARG} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{pseudo} \rangle \\
\langle \text{over} \rangle \\
\langle \text{zich} \rangle \\
\end{array}
\]

\[
\begin{array}{c}
\langle \text{F} \rangle \\
\langle \text{N} \rangle \\
\langle \text{F} \rangle \\
\langle \text{ARG} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{pseudo} \rangle \\
\langle \text{over} \rangle \\
\langle \text{zich} \rangle \\
\end{array}
\]

\[
\begin{array}{c}
\langle \text{F} \rangle \\
\langle \text{N} \rangle \\
\langle \text{F} \rangle \\
\langle \text{ARG} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{V} \rangle \\
\langle \text{pseudo} \rangle \\
\langle \text{over} \rangle \\
\langle \text{zich} \rangle \\
\end{array}
\]
We can take this to be indirect evidence in favor of our analysis of gerunds and auxiliaryless sentences in English.

4.2.3. Syntactic Rules

The syntactic rules of CRG define well-formed configurational syntactic structures and correspond to the annotated phrase structure rules of Lexical-Functional Grammar. As discussed in §2.1, we follow Falk (1983a) in factoring out the two roles of traditional phrase structure rules — those of specifying dominance relations and precedence relations — into separate rule types, which we call c(onstituency)-rules and o(dering)-rules. For arguments that this is desirable, see Falk (1983a), as well as Gazdar and Pullum (1981) where the same conclusion is reached independently, although they treat c-rules (which they call I(mmediate)
D(ominance) statements) and o-rules (L(inear) P(recedence) statements) as metarules which define full phrase structure rules. (9) represents the form of a constituency rule, and (10a) the basic form of an ordering rule. We also allow rules which pick out specific positions for a constituent, the possibilities being specified by Universal Grammar: initial (10b), final (10c), and perhaps second (10d).

(9) \[
\begin{array}{c}
\alpha F1 \\
\beta F2 \\
\gamma F3 \\
\vdots \\
\vdots
\end{array}
\begin{array}{c}
\eta F4 \\
\xi F5 \\
\zeta F6 \\
\vdots \\
\vdots
\end{array}
\begin{array}{c}
\vdots \\
\vdots
\end{array}
\begin{array}{c}
<GR> \\
\vdots
\end{array}
\]

(10) a. A \gg B (read: "A precedes B")

b. A initial

c. A final

d. A second

(9) specifies that a category with the feature decomposition on the left may have as a daughter at projection level \( \eta \) a category with the feature decomposition on the right bearing the grammatical relation GR. We do not have to specify that the constituent on the right-hand side of the rule is a maximal projection, because it has to be in order to comply with the X-Bar Convention. As discussed in §2.1, the A and B in (10) can be grammatical relations, categories (in distinctive features), or both.

In §2.1 we assumed that o-rules could, as a marked option, refer to the category of the dominating node. It is unclear whether or not this ability is necessary, and we can create a much stronger theory if this option is not provided. For the most part, o-rules referring to the dominating node can be reanalyzed; for example, in §2.1 o-rules which mentioned grammatical relations proved to be empirically superior to categorial rules making reference to the dominating node. In Falk (1983a)
reference to dominating categories is restricted to preferred ordering rules, such as (11), which allows Heavy NP Shift but does not allow P-PP-NP order in the phrase headed by a preposition which is subcategorized for NP and PP.

(11) \[ NP \gg PP \] (preferred / / )

Restricting reference to dominating categories to preferred o-rules is tolerable because, among other things, it requires positive evidence (namely, the violation of an o-rule), whereas the restriction of an o-rule by dominating category would require negative evidence. Another case alluded to in Falk (1983a) relates to the positioning of V' adverbs (subcategorized adverbs and non-ly adverbs; see Jackendoff 1977), which must follow NPs. We could write a rule (12), but it would wrongly predict that all adverbs have to follow NPs. This is not true of sentential adverbs, which can either precede or follow the SUBJ NP.

(12) \[ NP \gg ADVP \]

Replacing (12) with (13) does not seem as elegant as (14).

(13) a. \[ OBJ \gg ADVP \]

b. \[ OBJ2 \gg ADVP \]

(14) \[ NP \gg ADVP / / V' \]

There are two possibilities for avoiding (14): V' adverbs (well, poorly, ...) may not be true adverbs (Grimshaw, personal communication) or OBJ and OBJ2 will prove to be collapsible in some feature-type notation. We would like to maintain the stronger theory, which is the one which does not allow reference to dominating nodes in o-rules. It is preferable to try to retain this by reanalyzing cases like (14).
4.3. The Lexical Component

As in Lexical-Functional Grammar, the lexicon plays an important role in Configurational Grammar. Much of the descriptive apparatus takes the form of lexical rules. Following Chomsky (1970), we assume a system of lexical redundancy rules. Bresnan (1978) shows that a larger set of lexical rules is needed than that envisioned by Chomsky, arguing that phenomena like passivization, dativization, and raising and control phenomena are lexically specified.

The nature of many lexical phenomena have been dealt with in this dissertation. Relation-changing rules are discussed briefly in §1.2; for more details, the reader is referred to Bresnan (1980, 1982b). The lexical aspects of raising and control are discussed in §3.7, and of syntactic binding in §3.8. To articulate the properties of these processes more fully, a better-structured theory of the lexicon would be needed. A promising approach is the one that has been developed by researchers in lexicalist phonology (Kiparsky 1982, Mohanan 1982a) in which the lexicon is partially ordered by means of domains of rule application called levels or strata. Falk (1983c) presents very preliminary evidence showing that this approach to the lexicon is promising from the point of view of lexicalist syntax.

All morphology, including inflection, is accomplished in the lexicon in CRG, as argued on independent grounds by Lieber (1980) and Kiparsky (1982). Verbal inflection and SUBJ agreement are not accounted for by an INFL constituent in the syntax with an AGR element which governs the SUBJ. The SUBJ is governed by a lexical item, the verb, just like all other arguments, and, as Bresnan (1982a) shows, there are languages which
display verb agreement with elements such as OBJs. The role of agreement inflection can be seen essentially as adding information to subcategorization. Thus, the suffix -s adds the specification [3 pers, sg] to the SUBJ position of the verb. The word *tries*, for instance, can be assigned the syntactic argument list (15).

(15)  { SUBJ , XCOMP}
     3 pers
     sg

In an explicit lexicalist theory, only idiosyncratic information about lexical items needs to be explicitly listed, with lexical (redundancy) rules filling in the remaining information. The input to the lexical component is a *minimal lexical entry*, called a "partial lexical entry" by Bresnan (1982b) and "underived lexical item" by Kiparsky (1982). This minimal lexical entry contains no predictable rule governed information (semantic, syntactic, morphological, or phonological). For instance, a simple transitive verb will not have the links between its SUBJ and OBJ and the semantic arguments to which they are linked specified, because it is predictable which semantic argument will link up to which syntactic one. Idiosyncratic information such as irregular syntax-semantics linkages will be listed in the minimal lexical entries and be prevented from being superseded by a version of the Elsewhere Condition which takes lexical entries as identity rules; see Kiparsky (1982) for details.

The question of the structure of the lexicon is a very complex one, and it needs to be studied in much greater depth. This section has simply outlined what seems to be a highly promising direction in which to proceed.
Footnotes for Chapter 4

1 The "..." in Figure 4.1 does not include Syntactic Relations, according to a similar chart in Johnson and Postal (1980), which uses slightly different terminology. It includes the relations borne by prepositions to their accompanying NPs (FLAG and others) and relations like L[abel] and L[inear]P[recidence].

2 The terms "pure" and "impure" are due to Johnson (1977), who attributes them to Davide Perlmutter and Paul Postal.

3 As stated, the Stratal Uniqueness Law only applies to the Term Relations (1, 2, and 3). However, Perlmutter and Postal state that it is quite likely that it should be extended to the obliques.

4 These "problems" are not universally recognized as such. For instance, Steele et al. (1982) claim that all languages have a category AUX, of which M is the English instantiation. A more volatile issue is the VP question. This researcher's view is that "nonconfigurational" languages like Japanese and Warlpiri do not have syntactic VPs: see Falk (1983a). Much of the recent evidence for VPs in superficially VP-less languages comes from "LF" phenomena like crossover effects. On the assumption that LF phenomena can be incorporated into the semantic level in a theory of semantics such as that of Jackendoff (1983), this is evidence not for a syntactic VP but rather for some semantic notion equivalent to the NP-VP structure. Jackendoff proposes, purely on semantic grounds, that action sentences have an ACTOR-ACTION structure, and in §3.5.3 we proposed that possessive sentences have a POSSESSOR-EVENT/STATE structure. Perhaps it is in semantic structures like these rather than phantom syntactic structures that VP phenomena should be localized.

5 By referring to a "flat" structure, I do not mean to rule out a structure like (i), which is more similar to that suggested by Emonds (1976).

(i)
```
   C''
  /
C''
  /
C'
  /
C    VP
  |
NP
  |
```

6 I'd like to thank Jane Grimshaw for encouraging me to support the stronger position.
Thus far, we have motivated the existence of a level of configurational representation in syntax. The question we have not addressed is whether or not this level is sufficient; that is to say, can it substitute entirely for LFG's levels of c-structure and f-structure? It should be clear by simple examination that there is nothing that can be defined over c-structure as envisioned in LFG that cannot be defined over CRG's syntactic representation as described in Chapters 1 and 4; by ignoring the relational annotations we are left with a structure identical in every way (except, perhaps, for the explicit X-Bar formalism of Chapter 4) to c-structures. The status of f-structure, however, is not so clear. In this chapter, we will examine the evidence that has been offered in the LFG literature for a level of f-structure, and conclude that an autonomous level of representation which deals only with grammatical relations is not necessary.

It is important to realize at the outset that some of the putative arguments for f-structure do not apply in Configurational Grammar. Thus, some of the arguments offered by Kaplan and Bresnan (1982, 231) are arguments against single-structure theories such as that of Gazdar (1982) in which there is no representation of grammatical relations. Kaplan and Bresnan point out that "functional control" (Δ-control in the analysis of §3.7) transmits purely syntactic (as opposed to semantic) elements: expletive it and there, syntactic case marking features (Andrews 1982), and semantically empty idiom chunks. Like LFG, CRG represents Δ-control as a relationally represented link at the syntactic level. In addition,
the universal principles behind anaphora and lexical rules are cited: as we showed in §§2.5 and 3.9.4, anaphora actually requires a level of configurational representation, and lexical rules do not, of course, apply to f-structures even in LFG. These arguments are essentially arguments for recognizing grammatical relations in syntactic theory and in the lexicon.

The arguments we will address in this chapter are arguments based on structural properties of f-structure (§5.1), the representation of discontinuous expressions (§5.2), and the interpretation of sluicing structures (§5.3).

5.1. Structural Considerations

The most obvious source of evidence for f-structure is the fact that this level has structural properties which differ from c-structure. Since our configurational representation is essentially an annotated c-structure, this is the first place to look to see if our representation is adequate. In §5.1.1, we will simply look at the structural differences and see how they are represented in CRG, without regard for the significance of these differences. In §5.1.2, we examine some of the uses to which the structural properties of f-structure have been put, and show that CRG is adequately equipped. In almost every case, the relational label HEAD and the terms functional head and (maximal) functional projection (§4.2.2) will play a significant role. Functional head, it will be recalled, refers to items with the relational label HEAD. We repeat the definition of (maximal) functional projection.
(1) **Definition**

For a category $X$ in the transitive closure of the relation `HEAD` with respect to a category $I$, we say that $I$ is a *functional projection* of $X$. The maximal projection of a functional projections of $X$ which is not itself the functional head of another category is the *maximal functional projection* of $X$.

In addition, the notion of argument list will play a role.

5.1.1. Nonisomorphism of Constituent Structure and Functional Structure

One argument for considering the representation of surface grammatical relations to be an autonomous level is that different structural properties emerge. As discussed in the LFG literature, primarily Kaplan and Bresnan (1982, 218ff), c-structures and f-structures are not isomorphic. There are c-structure terminals which are not associated with functional positions, and there are f-structure constituents that are not represented in c-structure. Thus, by imposing the hierarchical organization of one level on another, the appropriate structural relations will be skewed. In addition, there is no one-to-one mapping between maximal categorial nodes in c-structure and subsidiary f-structures in f-structure.

A legitimate question to ask prior to examining this nonisomorphism is whether it matters: are the particular structural properties of f-structure an artifact of the representational system and of the need to represent surface grammatical relations directly, or are they supported by linguistic facts? In fact, these differences do seem to be significant, and we will examine concrete examples in the next section. First, we will simply examine the structural differences in isolation from any discussion of consequences.

In order to facilitate comparison between f-structure and c-structure,
we present the LFG representations of a sample sentence in unorthodox fashion: the c-structure is conventional, but the f-structure is presented as a tree whose terminal nodes are meaningful lexical items (i.e. the values of the PRED feature) and in which "features" are omitted. In (3), the conventional f-structure is presented, with the omitted features circled.

(2)

(3)

[SUBJ [PRED 'Max']

[TENSE past]

[PRED 'say<(SUBJ)(COMP)>'

[COMP [INF TO]

[SUBJ [PRED 'PRO' [U +]

[PRED 'sell<(SUBJ)(OBJ)(OBLGoal)/p>]

[OBJ [SPEC 'the' [PRED 'hamster' [IJUM pl]]]]]]]]
Note also the representation ("relational network") associated with this sentence in Relational Grammar.

(4)

(The UN in (4) represents an unspecified argument, unexpressed phonetically in English but realized as *on* in French and *man* in German (Perlmutter and Postal, in press).)

The most noticeable difference between the c-structure in (2) and the various relational representations (tree f-structure in (2), conventional f-structure in (3), relational network in (4)) is the merging of the S and VP nodes, and of the VP and V nodes. Thus, *Max*, which is a daughter of S, and *to sell the hamsters*, which is a daughter of VP, are both arguments of the main verb *say*, and bear grammatical relations to it; this is apparent from (2)-(4) by virtue of the fact that there is no VP in the relational representations. Now consider the CRG representation of (2).
We can see from the argument list associated with it that *say* takes a
*SUBJ* argument and a *COMP* argument. The maximal projection of *say* is the
*V''* (*VP*) vertically over it, and although the *COMP* is within that
constituent, the *SUBJ* is not. However, the maximal functional projection
of *say* is the root node *R*, and the *SUBJ* is located within it. Thus, by
utilizing the notion of functional projection we can duplicate LFG's
merger of *S* and *VP* at *f*-structure. Similarly, the maximal functional pro-
jection of *sell* is *C''* (*VP*).²

There is one *f*-structure position in (2) which is not represented
at *c*-structure: the complement *SUBJ*, which is the empty pronominal *PRO*.
This is a potential problem for any representational system which only
recognizes phonologically relevant terminal nodes in the representation
of grammatical configurations, and thus an apparent argument for the
existence of another level of representation, such as LFG's f-structure.
However, CRG does incorporate into its syntactic representation an encoding
of grammatical relations other than the relational annotations on configu-
rational positions. This is the argument list, and it is there that PRO
is represented. In Chapter 3 we analyzed anaphors as a relation between
positions in argument lists; this proved particularly useful in handling
obligatory Equi in nominals (§3.7.3). Thus, by annotating nonconfigura-
tional anaphors to argument list positions, we avoid this argument for
a level of f-structure. In fact, a persual of Chapter 3 will show that
CRG recognizes several nonconfigurational anaphors in English.

Next we turn to c-structure constituents which are not represented
as subsidiary f-structures. There are three of these in (2): the dummy
auxiliary did, the infinitival complementizer to, and the Determiner
the. Interestingly, these are the types of morphemes whose analysis has
generally been ignored by relational grammarians (although certainly
not by all of them; cf. Johnson and Postal 1980). However, in LFG it is
not entirely accurate to say that these elements are totally absent from
f-structure in the same way that PRO is absent from c-structure. Although
they are not represented as meaningful relation-bearing elements, they do
contribute features to the f-structure. An interesting question is whether
there is evidence that any c-structure constituent has no effect at all
on f-structure. I suspect careful inquiry will reveal that the answer is
No. If this is so, it calls out for explanation, and none is forthcoming
from LFG, unless one adds a well-formedness condition on phrase structure
rules and/or lexical entries, requiring functional equations. In CRG,
on the other hand, there is a kind of primacy granted to phrase structure
representation, in that phrase structure configurations are the backdrop over which lexical items, grammatical categories, and grammatical relations are laid. Thus, anything represented configurationally (c-structure) is somehow going to be felt relationally (f-structure).

How does CRG deal with the specific examples in (2)? The dummy auxiliary did and the infinitival complementizer to are both the configurational heads of their phrases. They serve as inflectional markers on their functional heads. Determiners such as the are different. LFG makes a distinction between "features" and "functions": both are names of positions in f-structure, but functions have subsidiary f-structures as values, while features have simple atomic values. Determiners are associated with the feature SPEC, and therefore are not represented as bearing a grammatical relation (function) to the NP. In CRG, the most natural analysis is to assign them the Syntactic Relation SPEC. I know of no evidence that would favor one approach over the other, so Determiners provide no evidence for f-structure.

Thus, the nonisomorphism of c-structure and f-structure in LFG does not seem to be an insoluble problem for a framework recognizing only one level of syntactic representation, as long as the level is configurational and allows one to make reference to such notions as "functional head" and "functional projection." A purely configurational level, like a standard surface structure or an LFG c-structure, is not sufficient, and neither is a purely relational level, like a relational network or an LFG f-structure. We are not denying the significance of "f-structural" properties: it is one of the major insights of LFG and a very important contribution to linguistic theory. However, the conclusion that there must be two levels of representation is unfounded.
5.1.2. Definitions and Features at Functional Structure

5.1.2.1. Definitions

One of the primary reasons for caring about f-structure configurations is that there are important structural notions defined there, notions which represent important generalizations. We include in "definitions" other generalizations and processes that depend crucially on f-structure configurations.

The notion "clause nucleus" (the LFG equivalent of "clause") is a very important one in LFG. It is used for determining anaphora options, for filling argument positions, and for determining what arguments are possible controllers for ADJs and XADJs. It is defined at f-structure, and crucially depends on f-structure configurations. A clause nucleus has been defined as an f-structure with a SUBJ and a PRED (Bresnan 1979, 37; 1980, 108 (159 in Bresnan, ed., 1982); Neidle 1982, 405) or as an f-structure with a PRED that is a lexical form (has argument structure) (Bresnan 1982a, 361 (304 in Bresnan, ed.); Levin 1982, 647). Because the former definition excludes nominalizations (on the assumption that NPs lack SUBJs), we will adopt the latter one.

Clause nuclei define the domains in which pronouns are free and r-anaphors are bound (Bresnan 1979; class lectures, MIT, 1980-1981). The contrast in (6) shows that configurational structure is not sufficient.

(6) a. Miriam kept Max about her/*herself/*him/himself.
   
b. Miriam told Max about *her/herself/*him/himself.

As we have already seen (Chapter 3), facts like these can be handled in CRG without reference to f-structure. Glossing over the facts that led
to postulate the existence of an "Extended Local Binding Domain," the domains relevant to anaphora can be defined in terms of argument lists. Therefore, anaphora does not present us with an argument for retaining f-structure.

According to Bresnan (1982, 361 (304 in Bresnan, ed.))

The clause nucleus is the domain of lexical subcategorization in the sense that it makes locally available to each lexical form the grammatical arguments that are subcategorized by it.

In other words, argument positions are filled by positions in f-structure which are sisters to the PRED feature. Thus far, we have not really addressed the question of how to know which constituents can fill which argument positions in which argument list. Consider (7).

(7) Who did the hamster want to see e?

In LFG, the two representations of (7) are (8)

(8) a.
In LFG, the argument list of want (or its lexical form) calls for a SUBJ and a COMP, and see is subcategorized for a SUBJ and an OBJ. The arguments of see are filled locally, in the sense that neither is drawn from outside the maximal projection of want(VP or V'') (although, of course, its PRO SUBJ is not represented configurationally at all). In the matrix clause the COMP of want is in the VP that want heads, but the SUBJ is not. That position in the argument list has to be filled from outside the VP, by the NP which is located under the maximal projection of M, the category of did. In the CRG conception of syntactic binding constructions the situation is even worse, because, as discussed in §3.8, FOCUS is also a term in the argument list (a Lexical Relation). The FOCUS in (7) is filled by a constituent of the root node R (a special type of \( \tilde{S} \)), which is not only outside the phrase headed by want, but also the one headed by did. In f-structure (8b), however, FOCUS, SUBJ, and COMP are all sisters of the PRED feature of want. Thus, the clause nucleus seems to be the domain of subcategorization. All arguments are filled by sisters of the lexical form. How can this be accomplished in CRG? The configurational representation of (7) is (9).
As we have already stated, the notion of functional projection plays a very important role here. The FOCUS, SUBJ, and COMP of want are all immediate constituents of projections or functional projections of want. The arguments of see are likewise within its maximal functional projection (or, in this case, within its maximal projection). We state the following principle.

(10) **Argument Locality Principle**

For each argument position α in the argument list of a lexical item X, α must be filled by an immediate constituent of a (functional) projection of X.

Similarly, the assignment of controllers for adjuncts can be based on the notion "functional projection."

Another use that f-structure has been put to relates to the notion "f-command." Mohanan (1983b) states it as (11a) and Bresnan (1982a) as
(11b).

(11)  a. (=Mohanan's (104))

The antecedent of pro must c-command it.³

... [This condition] holds at the level of f-structure.

b. (=Bresnan's (47))

For any occurrences of functions α, β in an f-structure F, α f-commands β if and only if α does not contain β and every f-structure of F that contains α contains β.

Both Mohanan and Bresnan crucially use the level of f-structure for the f-command condition on PRO. However, as we showed in §3.9.3, f-command can be defined configurationally by making reference to the maximal functional projection; in fact, for reasons having nothing to do with f-structure vs. maximal functional projection, the CRG definition proved to be empirically superior.

5.1.2.2. Features

Another source of evidence for f-structure is that f-structure constituents serve as the domains of features. We illustrate this with two examples from Falk (to appear), both related to supportive do which has the unique property of requiring the following VP to merge in f-structure with the S node that it (do) heads.

The first feature that we will discuss is the TENSE feature, which usually appears in the clause nucleus associated with a main verb by virtue of its inflection. However, when a sentence is headed by do, do supplies the tense feature for the verb heading the VP, the functional head of the sentence. As an example, consider (12).⁴
The hamster did scare the cat

```
(12) a. The hamster did scare the cat

b. The hamster

 V''
 M''
 N''
 DET''
 the N did V
 hamster scare DET''
 N''
 cat
```

c. [[SUBJ

 [SPEC 'the']

 [PRED 'hamster']

 TENSE past

 PRED 'scare<(SUBJ)(OBJ)>'

 INF BARE

 OBJ [SPEC 'the']

 [PRED 'cat']

] The TENSE feature associated lexically with the N did "spreads" to the VP. This is important because of the behavior of used to in the author's dialect.

(13) a. Hamsters used to rule the world

 b. Did hamsters use(d) to rule the world?

 c. Hamsters didn't use(d) to rule the world.

 d. *Do hamsters use to rule the world?

 e. *Would hamsters use to rule the world?

 f. *Hamsters had used to rule the world.

It is proposed by Falk (to appear) that use (in the sense of used to) has the following functional equation in its lexical entry.

(14) (↑TENSE)=past

The "=" identifies this as a constraint equation: a filter on f-struc-
tures in which the verb appears. It states that the clause nucleus in which the verb appears must be marked for past tense. Consider the f-structures of (13a,b,d), where the features from the lexical entry of the relevant forms of use are marked with checks.

(15)  

```plaintext
(15)  

a. [SUBJ [PRED 'hamster']
   [PERS pl]
   TENSE past √
   PRED 'used-to<(XCOMP)>(SUBJ)' √
   XCOMP [ ]
   ]

b. [SUBJ [PRED 'hamster']
   [PERS pl]
   TENSE past
   INF BARE √
   PRED 'used-to<(XCOMP)>(SUBJ)' √
   XCOMP [ ]
   ]

c. [SUBJ [PRED 'hamster']
   [PERS pl]
   TENSE present
   INF BARE √
   PRED 'used-to<(XCOMP)>(SUBJ)' √
   XCOMP [ ]
   ]
```

In (15a), the inflected verb used supplies the past tense feature, while in (15b) it is supplied by did. In (15c), which corresponds to (13d), do supplies a present tense feature, which violates (14). The point is that
used to is good in exactly two environments: alone and when embedded after supportive did. What is significant about these two environments is that they represent the only two environments in which used to is marked with a TENSE feature having the value 'past'. This can be achieved only if the past tense of did can percolate to its functional head. In LFG this is automatic because they form a single f-structure constituent.

The other feature discussed by Falk (to appear) is AUX. Every verb which has Verb-like inflection (i.e. third person singular present tense agreement, unlike true Modals) and participates in "subject-auxiliary inversion" is assumed to have this feature. It serves to convert these Verbs to Modals when tensed. Each auxiliary has a unique value for Thus the value of AUX for be is 'be', for have it is 'have', and for do it is 'do'. With no further explanation it follows that do cannot co-occur with other auxiliaries. To see why, consider the f-structure of (16).

(16) *The hamster does be screaming.

(17) *The hamster does be screaming.

(17) is an ill-formed f-structure because the feature AUX has two conflicting values, which violates the Uniqueness condition. The only reason this
analysis is open to us is because the S and VP merge at f-structure.

The point of both these cases is that the TENSE and AUX have as their domains f-structure constituents. This constitutes a rather straightforward argument for f-structure as a level of representation. To examine phenomena such as these in CRG, it is necessary to consider exactly what the role of configurational heads of phrases with functional heads is. Essentially, their role is to assign some sort of "grammatical feature" to the functional head. For instance, "case marking" prepositions (Bresnan 1979) serve to mark the functional head with its thematic relation. The infinitival complementizer to serves essentially as infinitival inflection (in fact, in Government/Binding theory it is taken to be a realization of [-finite] INFL). What do serves for, primarily, is as a carrier of tense, which it assigns to its functional head. What does this mean? Formally, it is clear that tense is a feature of do itself: did is a past tense form. Until now, we have ignored the role that tense plays in the theory; we have not represented it as an inflectional subscript on the V node along with "M", which simply indicates finiteness. The identity of tense does not have any syntactic reflexes, which means that there is no reason to represent it as a feature that takes part in the categorial specification. We can call it a noncategorial feature. In a complete syntactic representation these features would be associated with the appropriate nodes, clearly distinguished from categorial features. We might, for instance, represent the V node of a present tense verb as (18).

(18) \[ V_M: \text{present tense} \]

While categorial features and inflectional subscripts percolate to X-Bar projections and heads, noncategorial features percolate to functional projections and heads. Thus, the notions of functional projections and
functional heads once again play the role of $f$-structure configurations.

We state the following convention.

\[(19) \text{ Noncategorial Percolation}\]

A noncategorial feature associated with a lexical item $X$ percolates down to $X$'s functional head and up to $X$'s maximal functional projection.

Technically, (19) means that these features are associated with every node from the maximal functional projection to the functional head. In order to simplify the representations slightly, the features will only be annotated to the maximal functional projection and every lexical category node. Note also that, as stated, (19) does not say that noncategorial features which percolate up to a maximal functional projection also percolate down to that projection's configurational head. This seems to be the simplest statement. Let us compare the LFG analyses given in this section with the CRG versions. First, the simple illustration of a sentence with \textit{do} looks like (20) in CRG.$^5$

\[(20)\]

```
past tense
R; - part
do aux

\textlt{HEAD}\n
M''

\textlt{SUBJ}\n
M''; past tense
\textlt{SPEC}\n
\textlt{OBJ}\n
\textit{hamster}  

\textit{scare}  

\{SUBJ, OBJ\}
```

Note that the features past tense and do aux spread to VP, as in the LFG analysis. Now consider \textit{used to}. In CRG terms, this lexical item requires
the noncategorial feature past tense. The three examples for which we produced f-structures in (15) are analyzed as (21).

(21) a. past tense

b. past tense
As in the f-structure account, the tense feature with the value past comes from the verb itself in (21a) but from did in (21b), where it percolates to use by Noncategorial Percolation. In (21c), do provides the feature pres tense, which violates the lexical restriction on used to. Finally, we present the configurational structure of the ungrammatical sentence with do followed by another auxiliary verb (16).

(22) pres tense
    R; - part
    | do aux+ be aux+
    
    <HEAD>
    M''
    | <SUBJ> M''
    |    | <HEAD> V non''
    |    |    | do aux+ hamsters V non'
    |    |    | pres tense <XCOMP>
    |    |    | V non; - part V non'
    |    |    | do aux+ be aux+ screaming {SUBJ}
    |    |    | be {SUBJ, XCOMP}
Note the conflicting values for the aux feature, marked by arrows. The value do comes from the lexical entry of does, and the value be from be; do percolates up to the maximal functional projection R and down to R's functional head (the V node of be). The value be originates with be and percolates up to its maximal functional projection, which is the root node once again.

Representations such as these are cumbersome, and there is no harm in using simplified representations without the noncategorial features. However, it should be remembered that f-structure features can be handled in CRG without recourse to a level of f-structure by means of Noncategorial Percolation.

5.2. Discontinuous Expressions

A different kind of potential argument for the existence of an autonomous level of functional structure comes from languages like Warlpiri (also spelled "Walbiri"). In these languages, sets of words which are represented as constituents in more familiar languages, such as nouns and their restrictive modifiers, may be separated in the surface string of words. For example, consider (23).

(23) Kurdu-jarra-rlu ka -pala maliki wajili-p1-nyi
    child-DUAL -ERG PRES -2dual dog(ABS) chase -NONPAST
    aspect
    wita-jarra-rlu.
    small-DUAL-ERG
    'The two small children are chasing the dog.'
    'The two children are chasing the dog and they are small.'

(Simpson 1983, Chapter 4 (52))

As is indicated, this sentence is ambiguous. The first gloss corresponds to another grammatical sentence where 'child' and 'small' are in the same NP, followed by the number and case marking. Nash (1980) proposes an
operation called "merger" which assembles these discontinuous expressions prior to semantic interpretation in order to achieve this first reading. In Simpson (1983), an account of this phenomenon is given in the framework of LF\&G, and the first meaning of (22) is assigned the following f-structure.

(24) \[
\begin{array}{c}
\text{SUBJ} \\
\text{PRED} \\
\text{NUM} \\
\text{PERS} \\
\text{ADJ} \\
\text{ASP} \\
\text{TENSE} \\
\text{OBJ} \\
\end{array}
\begin{array}{c}
\text{CASE} \quad \text{ergative} \\
'kurdu' \\
du \\
3 \\
\text{CASE} \quad \text{ergative} \\
du \\
'wita<\text{SUBJ}>' \\
\text{PRED} \quad \text{'PRO'} \\
\end{array}
\]

The advantage of a representation like this is that it provides a universal representation of restrictive modification, regardless of the surface constituency displayed by a particular language. However, there is an assumption behind the LF\&G analysis, namely that all languages must represent restrictive modification in the same way in the syntax! Suppose that a language like Warlpiri is unusual in that it allows restrictive modifiers to appear outside the phrase they modify. This could be due to
some unclear parameter of Universal Grammar, perhaps related to the structure of the lexicon.

We can assign (23) the representation (24).

By treating \textit{wita-jarra-rlu} as an XADJ we account for the transmission (or agreement) of case. It will be recalled from §3.7 that consequents (including Δ) agree in certain grammatical features with their antecedents; case is one of them in many languages. Simpson (1983) argues on the basis of the fact that there are clausal adjuncts whose null SUBJs can be demonstrated not to agree in case with the modified noun that this analysis cannot be correct, but, depending on the "complementizer" suffix, many of these adjuncts do not exhibit agreement anyway. It seems that one of the apparent problems with the Warlpiri sentence, discontinuous expressions, is not really a problem. As long as we concede the possibility of interpreting a structure like (24) as involving restrictive modification, and there is no reason internal to Warlpiri not to, this sentence, which is typical, shows no evidence for a level of f-structure.

We also must point out that the degree to which the argument for f-structure is convincing depends on one's belief that there are universal f-structures of the relevant types. If f-structure configurations are universal, nonrestrictive relative clauses can probably not be represented
as being outside of the phrases that they modify. In English, for example, there is good evidence (Jackendoff 1977) that nonrestrictive relative clauses are inside their NPs, PPs, etc. That is, f-structure (24) appropriately adjusted for English, would represent both glosses of (23). This, in fact, was one of the arguments that we gave (§2.3) for configurational input to semantics.

5.3. Sluicing

Perhaps the most intriguing argument for f-structure is the argument based on L. Levin's (1982) analysis of the "Sluicing" construction. Sluicing is the process involved in (26).

(26) a. Max read some book, but I don't know which one.
    b. Max ate, but I don't know what.
    c. The hamster stole a car, but I don't know whose.
    d. Hamsters evolved from higher animals, but I don't know in which prehistoric era.

In a framework which does not recognize deletions under identity, which includes all frameworks which are the theoretical heirs of the Extended Standard Theory (or "interpretive semantics"), Sluicing is an interpretive rule; that is, the sentences in (26) are base-generated and the interpretation process builds up a copy of the appropriate clause. The complications of Sluicing make it a very difficult rule to state, and Levin's account is very attractive.

What concerns us here is not the details of Levin's account, which we will assume are essentially correct. Our interest lies in the claim that the process of interpreting sluiced constructions cannot be stated in a framework that does not recognize a level like LF's f-structure. The claim is that Sluicing requires access to syntactic information, semantic
information, and lexical information. If Sluicing were to be handled as a syntactic rule, it should only have access to syntactic information, and if it were a purely semantic rule it would only have access to semantic information. However, the three types of information are all represented in f-structure, making it the ideal level at which to state the rule. To what extent can this argument be carried over to CRG?

It must first be pointed out that it is not really accurate to say that a semantic account of Sluicing would predict that the procedure has access only to semantic information. This is because the notion "semantic rule" can mean one of two very different things. If it means actual rules operating at the level of semantic representation (or conceptual structure (Jackendoff 1983)), such as the rules governing inference and judgments of synonymy, then Levin's observation is to the point. However, what is usually meant when a process is called semantic, however, is that it is part of the syntax-semantics mapping, not that it is internal to the semantic/cognitive system. Mapping (or projection) rules have access to purely grammatical information; for instance, Jackendoff (1972), who considers anaphora to be part of semantics and not syntax, allows the rules of anaphora to make reference to purely syntactic notions such as precedence and command. We will consider Sluicing to be a rule of this latter type.

Sluicing has access to information about case marking. (The following examples are from Levin 1982, 610).

(27) a. Someone kissed Janet, but I don't remember who.
    b. *Someone kissed Janet, but I don't remember whom.
    c. *Someone kissed Janet, but I don't remember whose.

(28) a. Janet kissed someone, but I don't remember who.
    b. ?Janet kissed someone, but I don't remember whom.
    c. *Janet kissed someone, but I don't remember whose.
(29)  a. It was hard to talk to one of the boys, but I don't remember who.
 b. ?It was hard to talk to one of the boys, but I don't remember whom.
 c. *It was hard to talk to one of the boys, but I don't remember whose.

(30)  a. *Someone's mother called, but I don't remember who.
 b. *Someone's mother called, but I don't remember whom.
 c. Someone's mother called, but I don't remember whose.

Although the data are somewhat confused by the status of who and whom, it is clear that there is case agreement. As Levin shows in further examples, the morphological case on the wh word cannot be traced to semantic properties. Although we have not explicitly discussed the representation of case in CRG, in discussing Warlpiri in the previous section, we represented it as a lexical feature associated with argument positions, much like agreement features. If this is so, then case is adequately represented in configuralational structure for the purpose of Sluicing: if, as Levin proposes, sluiced constructions are interpreted by copying the predicate and argument list and filling all arguments but one with the arguments in the antecedent clause, the case marking requirement associated with the predicate will be available for checking against the wh word.

One interesting aspect of Sluicing that seems to fit into LEG very nicely pertains to the access that the interpretation procedure has to lexical information; in particular, the mapping between syntactic arguments and semantic arguments.

(31)  (=Levin's (6))

a. Jenny was eating, but nobody knows what.
 b. Seymour sliced the salami, but it's not clear with what.

(32)  (=Levin's (60c))

*John moved, but I don't know what. (=I don't know what John moved)
In each of these cases, the argument lists of the predicates on the left are not exactly the same as the homophonous covert predicate on the right.

(33)  

\[ \begin{align*}
\text{a. eat} \\
\text{i. } & \langle \text{Agent, Theme} \rangle \\
& \{ \text{SUBJ} \} \\
\text{ii. } & \langle \text{Agent, Theme} \rangle \\
& \{ \text{SUBJ, OBJ} \} \\
\text{b. slice} \\
\text{i. } & \langle \text{Agent, Theme} \rangle \\
& \{ \text{SUBJ, OBJ} \} \\
\text{ii. } & \langle \text{Agent, Theme, "instrument"} \rangle \\
& \{ \text{SUBJ, OBJ, OBJ} \} \\
\text{c. move} \\
\text{i. } & \langle \text{Theme} \rangle \\
& \{ \text{SUBJ} \} \\
\text{ii. } & \langle \text{Agent, Theme} \rangle \\
& \{ \text{SUBJ, OBJ} \} \\
\end{align*} \]

(33c) differs from the other two pairs in that to go from one form to the other, a syntax-semantics link has to be broken: the SUBJ is linked to the Theme in (i) but the Agent in (ii). Levin calls (33a,b) compatible semantic forms, and (33c) and others like it conflicting forms. The claim is that f-structure encodes the difference between compatible and conflicting semantic forms by means of the lexical form. However this is information which has to be available to semantic interpretation in any theory.

Thus, Levin's analysis of Sluicing can be adapted by CRG without the addition of a level of f-structure. Once again, a piece of putative evidence for such a level proves not to exist.
5.4. Conclusion

We conclude that the theory of Configurational Grammar has no need for a level of representation which, like LFG's f-structures and RG's relational networks, represent only grammatical relations. A single level of configurational structure is sufficient. The notions of functional heads and projections, along with argument list representations, do the work of f-structure.

A general question that we have not addressed is this: can everything that is representable at f-structure be encoded in configurational structure? As we have seen in this chapter, there is a wide range of f-structure representations which can be simply translated into configurational representation. However, one might ask if there are any potential f-structure-based arguments which could not be accommodated. The one area that would give CRG problems is the statement of identity between two distinct f-structure positions. In LFG, this is done primarily in functional control (Δ-control) and constituent control (syntactic binding). However, the evidence supports the anaphoric analysis given in Chapter 3. The best apparent argument for the LFG view in the case of functional control is that a "raised" argument seems to act as if it were in both clauses for the purposes of f-command. However, we saw in §3.9.3 that this is true only of Raising-to-Subject. The discussion in that section strongly suggests that "multidomination" is the wrong approach for functional (Δ-) control. The only other proposal that I am aware of in which identity of distinct positions has been invoked in LFG is Bresnan's (class lectures, MIT, 1981) analysis of Extrapolation. Essentially, the lexical entries of predicates with extrapoosed sentential "subjects" are claimed to contain
the following equation.

(34) \((+\text{SUBJ}) = (+\text{COMP})\)

That is, the COMP, which represents the extraposed clause, and the pleonastic SUBJ \((it)\) are merged at the level of f-structure. While we do not have an alternative analysis of Extraposition to offer, it would be very distressing if (34) turned out to be correct. If linguistic theory allows SUBJ and COMP to be merged at f-structure, why not any other pair of arguments, such as OBJ and OBL, OBJ and COMP, or even SUBJ and OBJ? The consequences of allowing analyses like (34), which are not available in CRG, are far-reaching and undesirable.
Footnotes for Chapter 5

1 There are two assumptions underlying this discussion, neither of which is pretheoretically obvious. The first, an assumption shared by LFG and CRG, is that SUBJs and OBJs have the same representational status, i.e. that S and VP do merge in the relational representation. This view has been challenged by Chomsky (1981), and, in more detail, by Marantz (1981). For discussion, see Bresnan (1982a). The other assumption is that this merger is surprising. It is if we are correct in considering S and VP to be projections of different categories, but if one follows Jackendoff (1977) in considering S and VP to both be projections of V, there is nothing surprising here and we are not dealing with a real difference between c- and f-structure.

2 This merger between VP and VP is unsurprising under Bresnan's conception of grammatical categories, because she considers VP to be the maximal projection of V.

3 Mohanan's "pro" is the same anaphor as our "PRO." He reserves "PRO" for what we call "A."

4 We are ignoring the fact that (12a) is grammatical only on an emphatic reading, because we take this not to be a syntactic fact, but rather a pragmatic one similar to Chomsky's (1981) "Avoid Pronoun Principle." For discussion, see Falk (to appear).

5 We are using the subscript 'non' for nonfinite verbs, i.e. as the opposite of 'N' as an inflectional subscript. The noncategorial feature for infinitives is assumed, without argument, to be -part[iciple].

6 One advantage of treating case as lexically required rather than structurally is that it makes ergative case marking easier to handle. If case were a feature (categorial or noncategorial) on NPs, as in Bresnan's (1982a) "nonconfigurational encoding," the c-rule introducing SUBJs would have to be context-sensitive and specify ergative case marking if there were an OBJ in the same clause and absolutive otherwise. The lexical entry for a verb will certainly have information it whether or not a verb is transitive. In fact, ignoring ditransitives, we could state ergative case marking as (1).

   (i) Mark the lowest Pure GR on the relational hierarchy to be absolutive and mark the other one ergative.
   In a transitive lexical entry, the OBJ will be absolutive and the SUBJ ergative, while in an intransitive one, the SUBJ, as the only Pure GR, will be he lowest and therefore absolutive.
APPENDIX: TOWARDS AN ANALYSIS OF GRAMMATICAL CATEGORIES

Generative grammar's greatest success has been in formally characterizing certain notions of traditional grammatical description in interesting ways. This is true of the many varieties of generative grammar currently in existence. For example, the traditional notion "passive" has been formalized in various frameworks as Case-absorption (Chomsky 1981), a change of grammatical relations borne by the NPs in a sentence (Perlmuutter and Postal 1977), and a change in the (relational) subcategorization (Bresnan 1982b) in various current approaches. Similarly, the theory of autosegmental phonology has allowed a revealing formal description of phenomena such as vowel harmony and consonantal roots in Semitic languages.

One problem are in this regard has been what is a central issue in most theories of syntax: the identification of grammatical categories. The earliest work (Chomsky 1955, 1957, 1965) assumed that categories could be identified but did not attempt to define them. The birth of Generative Semantics (circa Lakoff 1970, written in 1965) can be seen, in part, as an effort at such a definition in terms of logical notions such as "predicate."¹ The X-Bar Convention (henceforth XBC) of Chomsky (1970) suggested an outline for the definition of categories in terms of a less semantically oriented approach to syntax. More recently, Steele, et al., (1981) have suggested a more functional approach to the identification of categories, concentrating on the putative category AUX as it is (in their view) instantiated in various languages. The most successful of these approaches has been the XBC, but few studies which claim to be in that framework actually propose distinctive features, and many of these...
are not strictly consistent with the XBC. We will discuss several past feature proposals in §A.2.

Although a successful rigorous definition of categories has not yet been achieved, syntacticians experience no pangs of conscience in making sweeping cross-linguistic generalizations about categories. For instance, an oft-heard argument against separating English auxiliaries into a separate category (AUX, M, or whatever) is that the proposed category does not exist in other languages. While one might disagree with the specific conclusions, the study of Steele et al. (1981) is very valuable because it shows that it is a mistake to make blanket statements of this nature without a clear definition of category membership.

This Appendix represents an attempt at realizing the promise of the XBC as regards the definitions of categories. Half of the job is done by the formalism: phrasal categories are defined as projections of lexical categories. Thus, lexical categories are more primitive, with phrases defined in terms of them. Note that this suggests very strongly that categories be defined in terms of lexical properties. To our knowledge, no feature system proposal to date has ever attempted to meet this criterion. We will attempt it, but one of our features will reflect a clearly nonlexical syntactic property, and another also seems to refer to syntactic constituency.

If we are successful, we will have vindicated the XBC and shown that an explanatory model of grammar can incorporate the large number of categories in evidence in human language. To the extent that assuming all these categories results in elegant explanations of linguistic phenomena, we will be showing such a theory to be superior to a framework such as Generative Semantics, or a framework which denies the importance of
categories in syntax, such as Relational Grammar or the framework of Brame (1979).

\section*{A.1. Distinctive Features}

The notion of binary distinctive features did not, of course, originate with the XBC. It goes back to the phonological work of the Prague School earlier in this century. The definitive pre-generative work on phonological distinctive features is Jakobson, Fant, and Halle (1952), henceforth JFH. Within the generative tradition, the point of reference for almost all things phonological is Chomsky and Halle (1968), henceforth SPE. By studying the nature of features in these studies, we can come to an understanding of how distinctive feature systems work.

What is the role of phonological distinctive features? They serve primarily to define the phonological properties of the phonemes of a language. For instance, the consonants in Spanish which undergo intervocalic spirantization can be defined by the feature complex [+consonantal, -vocalic, +voice]. The statement of the English rule of Vowel Shift (SPE) involves changing the values of the features [thigh] and [tlow]. In JFH, the determination of whether voicing or tenseness is "phonemic" in a given language is dependent on its phonological properties. Nevertheless, the features are stated and defined in terms of physically measurable characteristics, articulatory and acoustic. Voicing may be a feature relevant for the operation of spirantization in Spanish, but it is defined in terms of what the vocal cords are doing during the articulation of the segment. SPE points out that at the purely phonetic level, the distinctive features are not really binary, as there can be degrees of voicing, nasalization, etc. Nevertheless, the features do represent real physical properties.
It is this duality of the phonological features that makes things difficult for syntacticians. The physical properties of phonemes are a constant no matter what view one takes of phonology. For instance, the view of phonology in JFH is different from that in SPE, yet the feature systems are very similar. Lexical (or syntactic) categories do not have physical correlates (unless one figures out how they are represented in the neural circuitry of the brain). There is no theory-neutral way to determine what the features should be. Nevertheless, for the features to be empirically meaningful, they must represent concrete syntactic properties.

In both JFH and SPE, the initial set of features ([tconsonantal] and [tvocalic] in the former; these two and [tsoncrant] in the latter) represent broad groupings of phonemes. Thus, in SPE, general categories such as vowels, consonants, nasal consonants, liquids, etc., are distinguished by these features. There are then more specific features which break up these general groupings by such criteria as manner and place of articulation. This basic pattern of general "major division" features and more specific features should be followed in syntax as well, but it generally has not. Another point to keep in mind is that in phonology the goal has generally not been economy of features. For instance, in SPE nasal consonants can be distinguished without the feature [±nasal]. Nevertheless, it is needed for phonological reasons. In phonology, as discussed above, we are guided by phonetic properties. It is harder in syntax, but we must not fall into the feature-counting trap. Just as there are holes in the phonemic inventory (JFH discuss some of these from a cross-linguistic perspective), we can expect holes in the category inventory as well.
A.2. Past Approaches

We will now briefly examine and reject three proposals for category distinctive features which have been proposed as part of the XBC. The discussion in §A.1 should be kept in mind throughout.

A.2.1. Chomsky and Stowell

The first proposal of features was Chomsky's. He proposed a feature system that split up the "major categories" N, V, and A. P has since been added to the system. The feature system is based on the traditional notions of "nominal" and "predicate," represented as [±N] and [±V], respectively. Nouns are [+N, -V], verbs are [-N, +V], adjectives are [+N, +V], and prepositions are [-N, -V]. However, no definition of these features in terms of formal properties was given in early work.

This feature system has been adopted in the Government/Binding framework (Chomsky 1981). In a study of phrase structure based on this framework, Stowell (1981) has extended the feature system. We will discuss Stowell's version, which is presented in Figure A.1.

The features and meanings of positive values

- [N]: cannot assign abstract Case (also if unspecified)
- [V]: ?
- [Tense]: phrase has tense reference and INFL (only for [+N, -V])

Feature decomposition of categories

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>V</th>
<th>Tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>gerund</td>
<td>+</td>
<td>-</td>
<td>Ø</td>
</tr>
<tr>
<td>passive</td>
<td>Ø</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Fig. A.1. The Chomsky-Stowell feature system (Stowell 1981)
There are several objections to this system. In the first place, although the feature [+N] has a definition in GB, and this definition reflects a lexical property, [+V] does not. The feature [+Tense] does not reflect a lexical property, but rather a syntactic and semantic property. Secondly, the use of unspecified features (the Øs in Figure A.1) to define separate categories recalls Stanley's (1967) objection to the use of partially specified feature matrices. While Kiparsky's (1982) approach to markedness uses partial specification in the minimal lexical entry, it is permitted specifically because the unspecified features get specified by rules. Thirdly, the feature system only includes "major categories." Adverbs, determiners, etc., are not part of the feature system. In a sense, one could say that the major division feature is [+feature], with no further specification of the [-feature] categories. There is no precedent for this in phonology. No phoneme, no matter how rare, is denied a full feature specification. Even boundary symbols (#, +, =), which are considered phonological entities in SPE, are assigned feature specifications there.

The biggest problem with Stowell's system is the treatment of for-infinitives. 5 is assigned the feature specification [+N, -V, +Tense] to show that it appears in NP-type positions but is distinguished by having tense specification. Stowell argues that 5 is a projection of COMP, primarily for reasons internal to GB. However, the COMP for is an "exceptional Case assigner" (see Chomsky 1981) and must bear the feature [-N]. Thus, it is either the case that 5 and its head have a different feature specification, or that they have contradictory features (both [+N] and [-N]). Neither of these is consistent with the XRC.

In conclusion, the Chomsky-Stowell feature system does not define all
categories and is not based on specific properties. Stowell's version of it is also internally inconsistent with respect to for-infinitives.

A.2.2. Bresnan

Bresnan (1982a) presents a feature system consistent with the framework of Lexical-Functional Grammar. The feature system has two features: \([\text{predicative}]\) and \([\text{transitive}]\), based on lexical properties and f-structure representation. The values and definitions of the features are presented in Figure A.2.

As we can see in Figure A.2, there are several problems with this feature system. In the first place, the predicativity feature is not really binary. It can have three values which are (apparently) distinct: '+', '-', and '±'. This is necessary by the definition given for the feature (which is, incidentally, a structural one, not lexical). If '±' were not distinct from '+' or '-', predicative NPs and PPs would be indistinguishable from APs and VPs, respectively, and nonpredicative NPs and PPs should be indistinguishable from S. In the second place, like

The features and meanings of positive values

[predicative]: the SUBJ can be functionally controlled
[transitive]: can have an OBJ as a syntactic argument

Feature decomposition of categories

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>±</td>
<td>-</td>
</tr>
<tr>
<td>V</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>A</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>P</td>
<td>±</td>
<td>+</td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Fig. A.2. The feature system of Bresnan (1982a)
the Chomsky-Stowell system, "minor" categories are excluded from the system, which only defines N, V, A, P, and the lexically headless $\bar{S}$, which Bresnan considers to be the maximal projection of the intermediate category $S$, which is exocentric.

An advantage of Bresnan's system is that the features are clearly defined in terms of identifiable properties which can be tested in other languages so as to identify categories. Ignoring $\bar{S}$, the features make the same division as in Chomsky's system, although the definition for $[\text{tpredicative}]$ does not translate into GB.

A.2.3. Jackendoff

Jackendoff (1977) presents a thorough and very valuable analysis of English phrase structure within the general Extended Standard Theory framework that gave rise to approaches like GB and LFG. He proposes a new feature system, along with several other revisions of the XBC, such as adding a third bar-level. His feature system, unlike the ones we have discussed thus far, includes all categories except complementizers, which he considers to be nonlexical inflectional elements introduced by a special "deverbalizing" phrase structure rule, and conjunctions, which are also introduced by a special rule.

The fact that all lexical categories are covered by Jackendoff's system means that it is possible to group the features as is done in phonology. There is one major division feature: $[\pm\text{Comp}]$. This distinguishes the "major" categories N, V, A, and P, from the various minor categories. The second set of features is like Chomsky's $[\pm I]$ and $[\pm V]$, and Bresnan's $[\pm p]$ and $[\pm t]$: it slices up the major categories. Jackendoff's features refer to the ability of the category to take a subject and (surface) object.
The features and meanings of positive values

[Comp]: can have complements
[Subj]: can have subject (or associated with major category that can)
[Obj]: can have surface object (or associated with maj. cat. that can)
[Det]: X'' specifier(?)

Feature decomposition of categories

<table>
<thead>
<tr>
<th></th>
<th>Comp</th>
<th>Subj</th>
<th>Obj</th>
<th>Det</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>V(S,S)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>PRT</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>ADV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Q</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DET</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DEG</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Fig. A.3. The feature system of Jackendoff (1977)

On the assumption that S is a projection of V and that possessive NPs are subjects, N and V are [+Subj] while A and P are [-Subj]. The [+Obj] feature is the same as Bresnan's [+t] and the inverse of Chomsky's [+N]. Although in Jackendoff's framework [+Obj] cannot be defined lexically (Ns and As can be subcategorized for objects and then require of insertion in the syntax), it is identical to features which are defined lexically in current frameworks. The [+Subj] feature, which makes the split differently from Chomsky's and Bresnan's systems, is justified on the basis of the fact that syntactic processes seem to treat N and V similarly. Whether or not [+Subj] can be defined lexically depends on whether or not subjects are subcategorized for.

Internal to Jackendoff's theoretical framework, the major difficulty with the system is the treatment of the [-Comp] categories. Jackendoff's strategy is to pair each one up with a [+Comp] category and assign it the
same values for \([±Subj]\) and \([±Obj]\). Thus, modals are paired up with verbs, particles with prepositions, determiners and quantifiers with nouns, and degree modifiers and adverbs with adjectives. In the latter two cases, it is observed that determiners and degree modifiers occur together, and a fourth feature, \([±Det]\), is added without explanation. In Figure A.3, we have defined it in terms of the syntactic reasoning that led Jackendoff to propose it. The problem with this approach is that it does not explain anything about the \([-\text{Comp}]\) categories, none of which (in Jackendoff's framework, at least) take subjects or objects or any other kind of argument. What sense does it make to call determiners \([+\text{Subj}]\) or particles \([+\text{Obj}]\)? The motivating force behind this approach is feature economy; the only alternative would be to invent additional features. However, as we discussed in §A.1, feature economy is not an appropriate criterion. The features of each category must specify the properties of that category, not of some category to which it is hypothesized to be related. In the present case, universal feature conditions should specify that \([+\text{Obj}]\) is as incompatible with \([+\text{Comp}]\) as \([+\text{high}]\) is with \([+\text{low}]\) in phonology.

External to Jackendoff's framework, doubts have been raised within both GB (Stowell 1981) and LFG (Rappaport 1983) concerning the generalization that Vs and Ns pattern together in allowing subjects. In GB, S is not a projection of V, and Stowell discusses some of the differences between possessives (which he does consider subjects) and subjects of sentences. Stowell also argues that other categories can have subjects. In LFG, doubts have been raised as to the subjecthood of possessives. In addition, S is not a projection of V in LFG or CRG either. Jackendoff himself notes that even under his assumptions there are languages which do not allow
subjects in NPs, yet one would want the features to be the same.

Although we have been somewhat negative about Jackendoff's features, his analysis may be one of the most important studies of phrase structure and categories. It is easy to criticize because of its (unfortunately rare) explicitness. By including the "minor" categories, he has made the job of proposing an alternative much easier. The view of categories to be taken here has been heavily influenced by Jackendoff's work, and this debt should be acknowledged.

A.3. A New Attempt

A.3.1. The General Framework

The syntactic framework to be presupposed here is that motivated in the body of this dissertation: the theory of Configurational Grammar. The motivation for this framework will not be repeated here.

The list of categories we will be assuming is set forth in (1).

(1)  N  (N''=NP)  
     V  (V''=VP)  
     A  (A''=AP)  
     P  (P''=PP)  
     M  (M''=S)  
     C  (C''=S, VP)  
     PRT  
     ADV  (ADV''=ADVP)  
     Q  (Q''=QP)  
     DET  
     DEG  (DEG''=DEGP)
The departures in the verbal and sentential system from the views of Jackendoff (1977) are discussed in §4.2.1.

One issue not discussed in §4.2.1 but important here is the internal relational structure of for-infinitives. A very attractive possibility is to treat for as a semantic predicate that takes a propositional argument. The NP would then be the nonthematic OBJ of for, and A-control the SUBJ of the VP. Alternatively, the embedded infinitive could be the functional head of the entire construction, with the NP assigned the relation SUBJ. The alternatives are diagrammed in (2).6

\[
\begin{align*}
(2) & \\
\text{a.} & C'''
\end{align*}
\]

\[
\begin{align*}
\text{b.} & C'''
\end{align*}
\]

What makes (2a) attractive is that if the rule introducing the NP is generalized to VP and PP objects, nothing has to be stipulated. The morphological case marking will be correct, and control follows automatically, with the nonthematic OBJ being lexically marked to take a consequent. One problem with it is that if for clauses in English are obviative (in the sense of Bresnan 1982a), treating the NP following for as an OBJ will complicate the statement of SUBJ-SUBJ Obviation in UG (Bresnan, personal communication). Another problem is that our feature system will be unable to collapse V, P, and C, without including M. It is unclear how to prohibit Modals from having OBJs. We leave the issue unresolved, but adopt (2a) as a working hypothesis.
A.3.2. The Major Division Features

As we pointed out in §A.1, the distinctive feature systems of JFH and SPE are based on a set of major division features: [±consonantal], [±vocalic], and, in the latter, [±sonorant]. We will propose two major division features: [±predicate] and [±grammaticalized]. They split up the categories into the four general groups in (3).

(3) +pred | -pred
     +gram | lexical grammatical formatives, specifiers
     -gram | "major" categories, quantifiers

These descriptions will become clearer in the coming two subsections.

A.3.2.1. [±predicate]

The first division among the various categories is between those that can and those that cannot define predicates. For this, we use the feature [±predicate]. We will go down the list in (2) and determine which ones are [+predicate] and which are [-predicate].

Nouns can obviously define predicates, although they need not, so they have the value '+'.

Verbs must, so they are also '+'.

Adjectives do as well; in many frameworks, including CRG, they are viewed as having SUBJ's, and they can take argument PPs, as in afraid of hamsters. Prepositions can also take arguments, at least those cases of PP which define location, etc. In some frameworks, including CRG, "case marking" prepositions do not take arguments, but being marked '+' does not mean that every lexical item of the category takes arguments. In the framework adopted here, modals also take arguments, as do complementizers. Particles are very troublesome to analyze, especially because they are hard to
distinguish from intransitive prepositions. One distinction is that particles may intervene between the verb and its OBJ while intransitive prepositions may not (see Falk 1983a). It seems that there are some cases of arguments for particles.

(4) I threw out the hamster.

We will therefore assign particles the value '+' for this feature. Adverbs do not take complements, but on the assumption that adjectives take SUBJS, although not lexical ones, the related adverbs must as well. They too are [+predicate]. Determiners are [-predicate], as are degree modifiers, which are the non-nominal equivalent of determiners (Jackendoff 1977). Finally, we will consider quantifiers not to be predicates.

A.3.2.2. [±grammaticalized]

The other major distinction is between those categories whose major role is to add some "meaning" to the sentence, and those which assign some grammatical (syntactic or semantic) feature to some element of the sentence. For instance, verbs, nouns, adjectives, adverbs, and quantifiers are always meaning-bearing elements. To the best of my knowledge, there are no members of any of these categories which can be identified as grammatical feature assigners. What about the remaining categories? Modals, although they are primarily predicates, do also sometimes assign grammatical information, such as when they are used in a "subjunctive" sense. Also, supportive do, which is a Modal at last when it emerges from the lexicon (Falk, to appear) serves only to assign tense to the main verb (functional head). Therefore, we consider modals to have the value '+'.

Similarly, prepositions and complementizers are carriers of feature information. Prepositions often serve to specify the thematic relation that
their "objects" bear to the governing predicate, and complementizers specify definiteness and the like (Bresnan 1972). Particles, and sometimes prepositions, are grammaticalized in a different way: they are often idiosyncratically selected, as in the infamous verb-particle construction. Determiners may assign such features as specificity, although, as shown by Jackendoff (1977), some semantic quantifiers are syntactic determiners. Degree modifiers, as the non-nominal equivalent of determiners, are also [+grammaticalized].

We can now be more specific about the divisions in (4).

(5)  [+pred, +gram]: lexical grammatical formatives: M, P, C, PRT
     [+pred, -gram]: "major" categories: V, N, A, ADV
     [-pred, +gram]: specifiers: DET, DEG
     [-pred, -gram]: quantifiers: Q

A.3.3. The Argument List Features

The next set of features to be discussed concerns the kinds of arguments a predicate can take, where type is defined syntactically. In English, these features are limited to the [+predicate] categories, but this is not true universally. For instance, there are languages (Warlpiri (Nash 1980, Simpson 1983) and Luiseño (Akmajian, Steele, and Wasow 1979) are two of them) in which there is an auxiliary element which is not itself a predicate but specifies certain information (such as number and person) about some of the arguments of the verb. Such auxiliaries are best analyzed as being [-predicate, +grammaticalized] but with '+' values for both of the features to be considered here.

A.3.3.1. [±P(ure) G(rammatical) R(elations)]

All approaches to syntactic distinctive features in the X-Bar tradition have included a feature to distinguish verbs and prepositions from nouns
and adjectives. The primary distinction is that verbs and prepositions take plain NP complements (objects) while nouns and adjectives do not. Within a framework like CRG, this is a lexical property and one which is very important for the operation of lexical rules like nominalization (Rappaport 1983, Falk 1983c). In our framework, it is also likely that complementizers take objects (§A.3.1).

What about modals? As is well known, modals do not take objects. Nevertheless, it is a priori preferable for verbs and modals to be similar both for historical reasons and because (as shown by Falk, to appear) there are certain aspects of the synchronic analysis of finite inflection which rely on these two categories being very similar. There is a property of modals which seems to correlate with the ability of verbs, prepositions, and complementizers to take objects: only modals can have lexical subjects. If we view modals as the equivalent in a lexicalist grammar of Chomsky's (1981) INFL, the unifying property is the ability to "assign abstract Case."

In fact, Marantz (1981) has suggested that the subjects of sentences be viewed as the object of INFL. Translating this intuition into the current framework is difficult because many other categories can have controlled subjects. We will define the feature [±PGR] as referring to the ability to have lexical arguments bearing Pure GRs, but the notion "lexical argument" is syntactic, not lexical. If this difficulty cannot be overcome, it may mean that the criterion of lexical definitions for features may be too strong.

To review, the feature [±PGR] has the value '+' for V, M, P, and C, and '-' for N, ADJ, ADV, and PRT.
A.3.3.2. [±O(bligatory) A(rguments)]

This feature separates those categories whose members must have arguments from those where they need not. Verbs are clearly [+OA] and nouns are clearly '−'. Prepositions often appear as case-markers (the ones that head OBLs) rather than argument-taking predicates, so they are [-OA]. Similarly, of the four basic complementizers in English (that, for, to, and as), at least half (that and to) do not take arguments. (The percentage is even greater if we add if and whether.) Therefore, complementizers are also '−'. Particles, even if some of them take arguments, are basically argumentless; another '−'. Modals are somewhat more problematic. The problem is supportive do, which is a verb at the level of the minimal lexical entry, but enters syntax as a modal. It takes no arguments. However, it is certainly an exception, unlike the case with P, C, PRT, and even N, where there are many cases of non-argument-takers. We will treat modals as being [+OA]. Adjectives and adverbs are also '+': they are modifiers, and what they modify usually bears the relation SUBJ to them.

A.3.3.3. Division of the Major Categories

If we plot the [+predicate] categories on a chart specifying their PGR and OA values, we get (6).

<table>
<thead>
<tr>
<th></th>
<th>+PGR</th>
<th>-PGR</th>
</tr>
</thead>
<tbody>
<tr>
<td>+OA</td>
<td>V, M</td>
<td>A, ADV</td>
</tr>
<tr>
<td>-OA</td>
<td>P, C</td>
<td>N, PRT</td>
</tr>
</tbody>
</table>

The italicized categories in (6) are the ones which have received the most attention in the feature literature. The division of these categories as given in (6) agrees with Chomsky and Bresnan. Jackendoff, in arguing
for his feature system, presents some evidence against this, and in favor of a system where nouns and verbs pattern together. However, his major argument is the generalization of the subject relation, a theory-internal approach which is not valid if nouns do not take subjects. Stowell (1981) tries to defend Chomsky's organization, but not very convincingly. For some unknown reason, the $\pm PGR$ feature, which everyone agrees to in some guise, is easier to argue for. Making the split where we have, however, seems to suggest that any language which neutralizes one of these two features would still maintain a distinction between nouns and verbs.

A.3.4. Additional Features

The features proposed thus far are not sufficient. Three pairs of categories remain undistinguished: $P$ and $C$, DET and DEG, and $A$ and $ADV$. We will propose two additional features here, one lexically defined and the other syntactically.

A.3.4.1. $\pm$basic

The first of the two features distinguishes those lexical categories which can be involved in the word-formation process from those which cannot. In accordance with the lexicalist theory of morphology (Lieber 1980, Kiparsky 1982), word formation is taken to include inflection. We designate this feature as $\pm$basic. Nouns, verbs, and adjectives clearly have the value '+’ as they are the input to both derivation and inflection. On the assumption of a rule like "Verb-Preposition Incorporation" (Bresnan 1982b), prepositions are also '+’. Complementizers, particles, etc., obviously do not participate in word formation, so they are '-’. The only unclear category is adverbs, because their status depends on how one views
superlatives and comparatives in which *ly* has been truncated. If they are actually derived from adverbs by a rule of truncation (Aronoff 1976) then ADV must be [+basic], but if one adopts Kiparsky's (1982) analysis in which comparative and superlative adverbs are derived directly from adjectives, then ADV may be [-basic]. We leave this unresolved. Finally note that modals are [-basic]. Modals never enter into derivational processes. For instance, there are no demodal nouns corresponding to the large class of deverbal nouns (with the possible exception of *must*).

Modals also have no regular inflectional paradigms; none of them exhibit SUBJ agreement in the present, and those that have past tenses have irregular ones that seem to be drifting away in meaning from their present tense forms. Thus, N, V, A, P, and maybe ADV are [+basic] while everything else is '-'.

A.3.4.2. [±nominal]

Our feature system now distinguishes almost all categories. However, DET and DEG are still identical, as well as A and ADV if ADV is [+basic]. As the analysis of Jackendoff (1977) makes clear, these two pairs of categories are related: in each pair, the first element is a noun modifier while the second one modifies all other categories. This is, of course, not a lexical property. Nevertheless, we use it for our last feature [±nominal]. All other categories are unmarked for this feature.

A.3.5. Summary

The feature system is summarized in Figure A.4
The features and meanings of positive values

[predicate]: can be a predicate (i.e. take logical arguments)
[grammaticalized]: assigns grammatical (syntactic/semantic) information
[PGR]: may take lexical arguments bearing Pure GRs
[OA]: must have syntactic arguments
[basic]: can be the input to word formation (derivation or inflection)
[nominal]: may only modify nouns (and if '-', may not)

Feature decomposition of categories

<table>
<thead>
<tr>
<th></th>
<th>pred</th>
<th>gram</th>
<th>PGR</th>
<th>OA</th>
<th>basic</th>
<th>nominal</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>P</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>H(=S)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>C(=S)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PRT</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ADV</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>Q</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DET</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DEG</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Fig. A.4. The feature system proposed here.

A.4. Syntax

The test of a syntactic category feature system lies in how well it expresses syntactic facts. In this section, we will first examine a small set of facts in English and write c-rules and o-rules, comparing our feature system with the others discussed. We will then see how our feature system accounts for the development of the category Modal in modern English.

A.4.1. Modern English Syntax

A.4.1.1. The Distribution of S and  S Complements

Sentential complements in English can take several forms. Superficially, there are four types: plain S, that-S, for-NP-to-VP, and to-VP. Relationally, these split into two groups: XCOMP, which encompasses some
instances of to-VP, and COMP, which includes all instances of the other three types and the remaining cases of to-VP. As Bresnan (1972) observes, for-NP-to-VP clauses have slightly different meaning than clauses involving full Ss; Erteschik Shir (1973) also points out that there are pragmatic conditions on that-lees Ss.

Here, we are concerned with the categorial differences between these types. We will therefore abstract away from semantic, relational, and pragmatic distinctions. In the present framework, plain S is the maximal projection of \( M \) while the other three types are maximal projections of C.

One would expect to find an asymmetry between these two. We do.

(7) a. N: plain S:  
   that-S:  
   for-NP-to-VP:  
   to-VP:  
   *Max's hope his hamster will return.  
   Max's hope that his hamster will return.  
   Max's hope for his hamster to win the Nobel Prize.  
   The hamster's hope to find Max.

b. A: plain S:  
   that-S:  
   for-NP-to-VP: \(^{10}\)  
   to-VP:  
   *Max is happy his hamster left home.  
   Max is happy that his hamster left home.  
   It is pleasant for Max for his hamster to earn a living.  
   Max is happy to be supported by his hamster.

c. V: plain S:  
   that-S:  
   for-NP-to-VP:  
   to-VP:  
   Max hopes his hamster will return.  
   Max hopes that his hamster will return.  
   Max hope for his hamster to win the Nobel Prize.  
   The hamster hopes to find Max.

d. P: plain S:  
   that-S:  
   for-NP-to-VP:  
   to-VP:  
   before Max wrote his term paper  
   *before that Max wrote his term paper  
   *before for Max to write his term paper  
   *before to write a term paper

As (7) shows, nouns and adjectives only take C''', prepositions only take M''', and verbs take both.\(^{11}\) There has been some effort, particularly on the part of researchers working within the GB framework, to show that some apparently arbitrary facts, such as the difference between what we are calling C' and M'', derive from general principles of UG; Stowell is an
outstanding example of this approach. We will take a somewhat more conservative view here and assume that the distribution in (7) is due to English rules of syntax. That is to say, we will explain how it is that we get the distribution in (7), but not why. It has not been shown in the literature that the UG approach is empirically sound. For instance, the very existence of complementizer-less clauses is unattested in many of the familiar non-English languages that have been studied. Furthermore, it is possible that complementizers do not form a single class cross-linguistically. Just as 'AUX' can refer to either a predicate with some grammatical role in addition to supplying its meaning to the sentence (as in English) or a specifier on the sentence (Warlpiri, Luiseno), so we might expect both of these possibilities to be realized with complementizers. If this is true, then in some languages the identification of $\bar{S}$ as a further projection from $S$, which we rejected for English, will be correct. The issue is the ine that motivated the study in this Appendix: the cross-linguistic identification of categories. Until we have some concrete evidence that (7) really does involve deep universal principles other than the metatheoretical concerns of GB theorists, we are better off taking the more conservative approach.

We take (8) to be the most concise statement of the generalization that we can expect a feature system to provide.

(8)  

a. N/A/V': $<\text{COMP}>C$

b. V/P': $<\text{COMP}>M$

c. $M/C''$ final (preferred)$^{12}$

How does (8) work out in features? In particular, how does our system compare with the other three? One problem we have to deal with here is the difference in views of what the distinction between $S$ and $\bar{S}$ is.
Chomsky (1981) and Stowell (1981) consider S to be a projection of INFL(ec- 
tion) and S a projection of COMP, but they consider every complement clause to have a COMP, even if it is not overt, and do not supply a category decomposition for S (i.e. INFL). Bresnan (1982a) considers S to be a further projection from S, so they belong to the same category and, in fact, every complement clause has an S node at its root. Jackendoff (1977) considers S and S to be alternate realizations of the maximal projection of V. Thus, there is no way to record the right-hand side of (8a,b) in either of these theories, and (8c) is trivial because it involves a single category. The left-hand side of (8b) is simple in all four feature systems, but that of (8a) involves a disjunction in every case except ours.

(9) Chomsky/Stowell

a. \{[+N]’’ \[+V]’’ \}: <COMP> ...

b. \[-V]’’ : <COMP> ...

(10) Bresnan

a. \{'+[p]\}' : <COMP> ...

b. \{[-t]\} : <COMP> ...

(11) Jackendoff

a. \{[+Subj]\}’’ : <COMP> ...

b. \{[+Obj]\}’’ : <COMP> ...

(12) Falk

a. \{[+pred]\}’’ 
   -gram 
   +nom \} : <COMP> C

b. \{[+PGR basic]\}’’ : <COMP> M

c. \{[+PGR basic]\}'' final (preferred)

Our ability to state (8a) as (12a) is due to our decision not to minimize the number of features.
A.4.1.2. Subjects

SUBJs in English can be NPs, PPs, and Ss. NPs and Ss are familiar. (13a) is an example of a PP SUBJ from Jackendoff (1977) and (13b) is one from Stowell (1981).

(13) a. From Groundhog Day to Purim is not too long.

b. Under the stars is a nice place to sleep.

Note that while S (C'' in our framework) can be SUBJ, S (M'') cannot. As with the distinction between C'' and M'' complements, we take the position that this fact must be a part of the grammar of English. As a glance at Figure A.4 will indicate, N, P, and C are three of four categories with the partial decomposition [+pred, -DA], the fourth being PRT. Particles are a somewhat bizarre category, and, considering the limited possibilities for PP SUBJs, there may some independent reason for excluding PRTs from SUBJ position. Another possibility is that our feature specification for particles is not quite right. If neither of these explanations works, we could resort to a more complex feature bundle for the SUBJ c-rule, [+pred, -OA, agram, aPGR], but this should be the choice of last resort.

A.4.1.3. Degree Modifiers and Determiners

Jackendoff (1977) guesses that degree modifiers like too constitute a category similar in its syntactic properties to DET, what he and we have called DEG. The c-rule that Jackendoff's data suggest is (14), where the subscript indicates that DET is chosen for N and DEG elsewhere.

(14) A/ADV/Nqx/Q'': <SPEC> DETqx/DEG

Jackendoff's feature version of (14) is (15).
(15) \[
\begin{align*}
\{\text{+Subj} \} & \quad , \\
\{\text{+Comp} \} & \quad <\text{SPEC} > \quad <\text{+Subj} > \\
\{\text{-Obj} \} & \quad <\text{-Obj} > \\
\{\text{-Det} \} & \quad <\text{-Det} > \\
\end{align*}
\]

(15) is fairly straightforward. One problem with it, however, is the fact that it takes two features on the left and one on the right all enclosed in angle brackets to indicate that nouns take determiners instead of degree modifiers. We would like to suggest a notational innovation limited to the feature [±nominal]: if the right-hand side of a c-rule includes the feature [±nominal] and 'a' does not appear elsewhere in the rule, then a will be interpreted as '-' for all dominating categories except N. We can state the rule in our features as (16).

(16) \[
\begin{align*}
\{\text{-gram} \} & \quad , \\
\{\text{-PGR} \} & \quad <\text{SPEC} > \quad <\text{-pred} > \\
\{\text{-pred} \} & \quad <\text{-pred} > \\
\{\text{-gram} \} & \quad <\text{gram} > \\
\{\text{-gram} \} & \quad <\text{-gram} > \\
\{\text{-gram} \} & \quad <\text{-gram} > \\
\end{align*}
\]

4.1.4. Adjectives and Adverbs

Jackendoff (1977) suggests that adjective–phrases modifying nouns and adverb phrases modifying verbs and other categories should be collapsed by the XBC. He shows that ADVPs can occur as modifiers on the "double bar" level for V, A, ADV, P, and even PRT. These ADVPs serve as adjuncts which modify the head of the phrase. We can state the following c-rule.

(17) \[
\begin{align*}
V/N_a/A/ADV/P/PRT" : <\text{ADJ} > A_a/ADV \\
\end{align*}
\]

Jackendoff does not actually state a distinctive feature phrase structure rule. The best available statement seems to be (18).

(18) \[
\begin{align*}
\{\text{-Subj} \} & \quad , \\
\{\text{-Obj} \} & \quad <\text{-Obj} > \\
\{\text{-Det} \} & \quad <\text{-Det} > \\
\end{align*}
\]

\[
\begin{align*}
\{\text{+Comp} \} & \quad , \\
\{\text{+Subj} \} & \quad <\text{SPEC} > \quad <\text{+Subj} > \\
\{\text{-Comp} \} & \quad <\text{-Comp} > \\
\{\text{-Subj} \} & \quad <\text{-Subj} > \\
\{\text{-Det} \} & \quad <\text{-Det} > \\
\end{align*}
\]
The top half of the disjunction in (18) picks out N, V, A, and P, with special mention made of N; the bottom half collapses ADV and PRT.

Before determining how to handle (17) in our framework, we must first note a problem mentioned by Jackendoff. He observes that PRTs can only take an ADVP when not preOBJ. As particles are morphologically identical to prepositions, it is possible that the cases where it appears that a PRT is being modified by an ADVP that it is really a preposition. This is one of those instances where we can let the theory decide. In the case of Jackendoff's feature system, it is better to include PRT: in our case it is better to treat is as P. On the assumption that ADVs are [+bas.c], we get (19).

(19)  [+basic]: \[<\text{ADJ}> \begin{bmatrix} +\text{OA} \\ -\text{PGR} \\ \text{anom} \end{bmatrix}] \\

We can take (19) as evidence bearing on the feature decomposition of ADV.

There is one problem we should note here. According to the analysis of Falk (1983a), based on Jackendoff (1972, 1977), V adverbs are not positioned by an o-rule. Instead, they have free ordering among their sister constituents. In AP, ADVP, NP, and PP, however, APs and ADVPs are obligatorily prehead. There is no feature bundle in any feature system we are aware of that will collapse A, ADV, N, and P. Perhaps the o-rule should not include all four categories, but rather have more than one o-rule.15

A.4.2. The Evolution of the Category "Modals"

The only generative study prior to this one to examine the question of cross-linguistic equivalences of category, Steele et al. (1981), concentrates on the universality of the category AUX. The reason for this is obvious: there is a large volume of literature attempting to show that
there is no such category, in English or in any other language. What we
will attempt to do here is to show how the category M evolved in the
history of English, in terms of the feature system proposed in this
Appendix.

Old English (OE) did not have a separate category consisting of
auxiliaries that could not appear in nonfinite form. However, the ances-
tors of today's modals (which Lightfoot 1979 calls "premodals") did
possess one property reminiscent of their descendants: they did not
have the regular third person singular suffix in the present tense. They
were part of a class that Lightfoot (1979) calls "preterite-present,"
because their present tense forms were originally past tense.

It is well-known that it is a property of Germanic languages as a
class that finite and nonfinite verbs display different syntactic proper-
ties. Finite verbs are subject to the "Verb Second Constraint" (Maling
and Zaenen 1981) while nonfinite verbs were either final or V' initial,
depending on the language. In some languages, such as OE and modern German,
the Verb Second Constraint only applies to main clauses (or only optionally
in subordinate clauses, as may have been the case in OE), obscuring the
general distinction between finite and nonfinite verbs. However, the
general Germanic property is that finite and nonfinite verbs are distinct.

Suppose that finite verbs constitute a distinct lexical category in
(most) Germanic languages, which we will call F, saving V for nonfinite
verbs. V would have the same feature decomposition as in English: [+pred,
gram, +PGR, +OA, +basic]. What are the features of F? There is certainly
no motivation for changing the values of [+pred], [+gram], [+PGR], or
[±OA]. Recall, however, that it is a general property of inflected forms
that they do not enter into word formation processes (Kiparsky 1982).
Therefore, we can assign this hypothetical category \( F \) the feature decomposition \([+\text{pred}, -\text{gram}, +\text{PGR}, +\text{OA}, -\text{basic}]\). This category is a lexically derived one: at the level of the minimal lexical entry, all verbs belong to the category \( V \). Although, as mentioned above, the evidence for the category \( F \) was somewhat obscured in OE, it is not unreasonable to suppose that it was clear enough to the native speakers of the language. In modern English (NE), as analyzed by Falk (to appear), the category \( M \) (which we have analyzed as \([+\text{pred}, +\text{gram}, +\text{PGR}, +\text{OA}, -\text{basic}]\)) includes modals and the finite forms of other auxiliaries such as \( \text{have} \), \( \text{be} \), and \( \text{do} \), which are, at the level of the minimal lexical entry, verbs with a lexical AUX feature. This feature also prevents auxiliaries from cooccurring with supportive \( \text{do} \). As the reader may suspect, we view \( M \) as a descendent of the Germanic category \( F \). Its syntactic positioning is roughly similar to that of \( F \) in other Germanic languages, and the category-switching behavior of nonmodal auxiliaries is a relic of the change of category that all verbs underwent in earlier stages of the language.

Our description of the changes in the English verbal system is a composite of the descriptions of Visser (1969), Traugott (1972), Hausmann (1974), and Lightfoot (1979). The basic changes correspond fairly closely to the traditional division between OE, Middle English (ME), Early Modern English (ENE), and NE. The division between OE and ME is roughly in the mid-12th century. ENE begins c. 1500 and NE c. 1700.

From our perspective, ME was much like OE. The only real difference is the development of the dummy auxiliary \( \text{do} \), which apparently developed first in poetry and then spread to prose. It is unclear where this use of \( \text{do} \) originated (see Visser 1969 and Hausmann 1974), but it seems not to have induced any other immediate changes. According to Visser (1969), \( \text{do} \)
did not cooccur with any auxiliary verbs, such as have (in all its uses), be, used to, and others. This suggests that the AUX feature was already present in the grammar. It is not implausible that languages have some sort of formal representation of auxiliary verb regardless of category, and the fact that there has historically been some leakage from the class of auxiliaries in English implies that it is a purely formal feature.

Around the beginning of the ENE period, the first major restructuring of the verbal system took place. During ME, the nonpremodal preterite-presents began dropping out of the language or developing regular paradigms. In the last decade of the period, those premodals that were able to take OBJ arguments lost that ability (with the exception of can, which hung on until 1641). What happened then at the beginning of the ENE period was that the premodals, which were now clearly distinguished from other verbs, came to be identified as F at the level of the minimal lexical entry. In fact, it is possible that at least some speakers may have made the transition from [-gram] to [+gram] (i.e. from F to N) even at this early stage. There is no way to tell from the data that this happened, but with the existence of a purely grammaticalized verb do in ME, there may have been some tension to create a [+gram] verbal category. In addition, it is highly marked for a [-gram] category to be [-basic]. At any rate, this reanalysis led to a succession of changes: premodals lost their ability to appear in nonfinite form, the use of do increased, dare developed a second present tense with -s in the third person singular which, a generation later, came to be used only with to-infinitives while the original dare continued to be used only with bare infinitives, and need developed a preterite-present-like present tense alongside the regular original. These changes, particularly need and dare, follow naturally from the crea-
tion of the new category.

A further change took place at the end of the ENE period. It is signalled by the inability of ordinary verbs to invert with the subject NP in questions. The grammatical change that caused this empirical alteration was the loss of the automatic change of category for finite verbs. Most verbs were now always members of the category V, and modals were always members of the category M. Only a small class of nonmodal auxiliaries (the Helping Verbs of Falk, to appear) have preserved the old Germanic pattern.

Our account does not account for every fact about the evolution of the category "modal," but it does a lot of work. It shows that we have a basis for explaining category changes by means of our feature system.

A.5. Conclusion

The XBC represents the best attempt at defining grammatical categories and acceptable grammatical configurations to date. The distinctive feature system proposed here, which takes seriously the analogy with phonology, is an attempt to contribute to the XBC literature and, perhaps, improve our understanding of what it means to be a noun, etc. Note that under a serious view of categorial features, it makes no more sense to talk of the proliferation of categories than it does to talk of the proliferation of phonemes.

Further work on other languages is needed to confirm or disconfirm the features we have proposed here. It is important to examine each language on its own merits. For instance, as we discussed in §A.4.2, there is reason to suspect that finite verbs in many Germanic languages belong to a different category than nonfinite verbs; this would not be
a possible analysis if we assumed that category labels are universal atomic primitives. Similarly for AUX in English on the one hand and Warlpiri and Luiseño on the other. To cite but one additional example, prepositions in Biblical Hebrew resemble nouns, both historically and in terms of lexico-syntactic properties. In particular, the preposition-"object" relation is an instance of a Semitic construction termed *ṣmikut* by Hebrew grammarians, which is characteristic of nouns and some of their arguments, primarily possessors. Thus, on the assumption that nouns do not take OBJs in Hebrew, prepositions do not either and are [−PGR].

(For some discussion of the Biblical Hebrew prepositions from a descriptive point of view, see Gesenius, et al. 1910.)
Footnotes for Appendix

1 This is clearly not the only way of interpreting the Generative Semantics literature. For instance, McCawley (1981) suggests that the difference between the view of categories in Generative Semantics and the Extended Standard Theory is that the former deals with "deep" (logical) categories and the latter with surface categories.

2 Two examples will suffice as an illustration. First of all, the ever-popular NP-S analysis of relatives is inconsistent with a strict (i.e. restrictive; i.e. a priori preferable) reading of the Convention, as discussed by Jackendoff (1977). Secondly, the XBC does not allow nonmaximal categories to appear as complements, but such analyses have been proposed, such as Gazdar, Pullum, and Sag's (1982) analysis of auxiliary verbs, which they view as verbs which take V' complements in a framework which assumes two bar levels.

3 It may seem that this is slightly overstated. After all, it is just as easy to interpret the XBC as defining lexical categories on the basis of phrasal categories (N is the head of NP, etc.). However, this alternate view is misdirected. Lexical categories have properties whether they are in phrases or not; namely, lexical properties. Phrasal categories in general do not exist independently of the categories which head them and of which they are projections. This last statement has the qualification "in general" because there are a few cases of apparent exceptions. For a treatment of some apparent exceptions, see §2.7.2.

4 Stowell's system is somewhat inconsistent. The meaningful value of the feature [±N] is actually '-' in terms of current GB theory rather than '+'; for [±Tense], however, it is the expected '+'. What this means is that when [±Tense] is unspecified (gerunds) the phrase behaves as if it had the '-' value (i.e. does not display tense reference and has no INEL) while when [±N] is unspecified (passives), the phrase behaves as if it were [±N] (i.e. does not assign Case). This fact about [±N] is, of course, due to historical accident. Originally, [±N] could be defined as governing of-insertion (which passives also do not do).

5 Jackendoff uses the term ART(icle) instead of DET(erator).

The structures in (2) arbitrarily show the NP and V̅P as sisters. This does not mean that we reject arguments that [for NP] forms a constituent, perhaps C', which V̅P is not a part of. We simply intend to be neutral on this point.

6 In determining the features of DEG, I am making extensive use of Jackendoff's (1977) observation of a correlation between DEG and DET.

8 The term "grammatical feature" here may reflect a bad choice of
terminology. For instance, within the GB framework, verbs assign abstract Case to their arguments, and Case can be considered a "grammatical feature," but we would not want to analyze V as [+gram]. What this feature is intended to do is formalize the intuitive notion "function word."

In §A.4.1.4, we will find some syntactic evidence for analyzing ADVs as [+basic].

Complements of the form for-NP-to-VP are actually rather rare with adjectives. This may be because most of the adjectives that take complements are factives, and the semantics of for infinitivals is nonfactive.

We are omitting the category C from the list of categories that take complements at the X' level. It has been suggested (e.g. by Emonds 1976) that for-NP forms a constituent that excludes to-VP. In our framework this would indicate that this complement is a daughter of C'.

It would not do to simply state (i) as the o-rule for clausal complements.

(i) COMP final
The reason is that subjectless infinitives serving as XCOMP's also must be final when they are to infinitives (as opposed to bare infinitives, which are plain VPs). The theory does not provide a notation for collapsing grammatical relations, and before weakening the theory to allow that it needs to be shown to be necessary. In this case, it is obviously not. (8c) is a more accurate version of (19a) of Chapter 2.

We are faced here with the need to invent a notation because of the fact that Bresnan's feature [±p] has three possible values. For these collapsed rules, the notation [ap] collapses [ap] and [tp], while the notation ['a'p] means only [ap].

We are assuming that the other necessary feature, [+Comp], is supplied by UG for any rule defining a daughter of X'.

One piece of evidence that may bear on this is the fact that adjectives can follow their head nouns under some poorly understood conditions. The clearest case is an NP like (i).

(i) the man proud of his son
The status of this in the theory is unclear, but one possibility is that AP and N are not ordered by an o-rule, but rather that there is an interpretive principle that refers to their relative order. In this case, the ordering rule would be (iia), which can be collapsed as (iib).

(ii) a. ADV >> A'/ADV'/p'
b. ADV >> aPGR'

For the CRG account, see §4.1.2.2.

Lightfoot (1979) lists these and several other changes. The relevance of the other changes he lists is not clear, so I have not
bothered to list them. We agree that there was a series of accidental changes that succeeded in isolating the premodals as a distinct category. It should be noted that Lightfoot, working without a system of distinctive features, has no explanation for the properties of the resulting category M.

18One problem with this conjecture is that at the level of the minimal lexical entry, do is a Verb, not a Modal. See Falk, to appear.
ABBREVIATIONS


Akmajian, Adrian, Susan Steele, and Thomas Wasow (1979) "The Category AUX in Universal Grammar." Linguistic Inquiry 10: 1–64


Baker, Mark (1983a) "Objects, Themes, and Lexical Rules in Italian." in Levin, Rappaport, and Zaenen, eds.

——— (1983b) "'Assume GF' and the Order of Morphemes." in Haik and Massam, eds.


——— (1979) "Realistic Grammar." in Brame, ed.


——— (1978) "A Realistic Transformational Grammar." in Halle, Bresnan, and Miller, eds.

——— (1979) "A Theory of Grammatical Representation." Duplicated lecture notes, MIT.


(1973) "Conditions on Transformations." in Anderson and Kiparsky, eds.

(1977) "On Wh-Movement." in Culicover, Wasow, and Akmajian, eds.


Falk, Yehuda N. (1979) "The Status of S in the X-Bar Convention."
Unpublished paper, Brandeis University.

Linguistic Analysis 11: 331-360.

——— (1983b) "Subjects and Long-Distance Dependencies." Linguistic
Analysis 12: 245-270.

——— (1983c) "Conjectures on Lexical Domains." in Haïk and Massam, eds.

——— (to appear) "The English Auxiliary System: A Lexical-Functional
Analysis." Language 60(3).

Fassi Fehri, Abdelkader (1980) "Some Complement Phenomena in Arabic,
Lexical Grammar, the Complementizer Phrase Hypothesis, and the
Non-accessibility Condition." Part I. Unpublished paper,
University of Rabat.

Gazdar, Gerald (1981) "Unbounded Dependencies and Coordinate Structure."


Order, and the Notion 'Head'." in Moortgat, Hulst, and Hoekstra, eds.

———, Geoffrey K. Pullum, and Ivan Sag (1982) "Auxiliaries and Related

Gesenius, Wilhelm, E. Kautzsch, and A. E. Cowley (1910) Gesenius' "Hebrew

Green, Georgia M. (1976) "Main Clause Phenomena in Subordinate Clauses."
Language 52: 382-397.

Grimshaw, Jane (1979) "Complement Selection and the Lexicon." Linguistic
Inquiry 10: 279-326.


——— (1982b) "On the Lexical Representation of Romance Reflexive
Clitics." in Bresnan, ed.

Gruber, Jeffrey S. (1965) Studies in Lexical Relations. Doctoral
dissertation, MIT. Published in Gruber (1976).

——— (1976) Lexical Structures : i Syntax and Semantics. Amsterdam:
North Holland.


Kiparsky, Paul (1973) "Elsewhere in Phonology." in Anderson and Kiparsky, eds.


——— (1978b) "Why Subject Sentences Don't Exist." in Keyser, ed.


-------- (1982a) "Whither 'Move NP'?" in Marantz and Stowell, eds.

--------(1982b) "Grammatical Relations and Explanation in Linguistics." in Zaenen, ed.

(1982b) "Grammatical Relations and Anaphora in Malayalam." in Marantz and Stowell, eds.

(1982c) "Grammatical Relations and Clause Structure in Malayalam." in Bresnan, ed.

(1983a) "Lexical and Configurational Structures." Unpublished paper, MIT.


Neidle, Carol (1982) "Case Agreement in Russian." in Bresnan, ed.


———, and Paul M. Postal (in press) "Impersonal Passives and Some Relational Laws." in Perlmutter and Rosen, eds.


Rappaport, Halka (1933) "On the Nature of Derived Nominals." in Levin, Rappaport, and Zaenen, eds.


Wanner, Eric, and Michael Maratsos (1978) "An ATN Approach to Comprehension." in Halle, Bresnan, and Miller, eds.


Woolford, Ellen (1982) "Island Constraints, Subjacency, and ECP." in Marantz and Stowell, eds.


Zubizaretta, Maria-Luisa, and Jean-Roger Vergnaud (1982) "On Virtual Categories." in Marantz and Stowell, eds.
BIOGRAPHICAL NOTE

The author was born Eugene Norman Falk in New York City on 14 November 1958. He attended public schools in Laurelton (Queens), New York, and Lynbrook, Long Island, New York, and graduated from Lynbrook High School in June 1976. He subsequently attended Brandeis University, from which he graduated magna cum laude with highest honors in linguistics in May 1980. At MIT, he participated in two research groups: the LFG Research Project and the Lexicon Seminar. Following his graduation from MIT, the author will begin teaching linguistics in the Department of English at the Hebrew University of Jerusalem in Jerusalem, Israel.

The following is a list of the author's publications:


