Joan Bresnan

Control and Complementation

This article develops a theory of control and complementation based on the lexical–functional theory of grammar, LFG, presented in Bresnan (1982a). A formal theory of grammar, such as the theory of LFG, is not itself a substantive linguistic theory. Rather, it is a language for precisely expressing descriptive rules and universal postulates of grammar. The choice of such a formal theory of grammar is extremely important. If the formal theory contains the appropriate concepts and representations, then linguistic principles and grammatical descriptions expressed within it will immediately generalize along the right dimensions, simplifying both descriptive rules and theoretical postulates. An inappropriate formal theory will require a host of auxiliary concepts and definitions, and may obscure the underlying regularities that optimal grammars must express. Having no formal theory at all will lead to vague and inconsistent formulations at both the theoretical and the descriptive levels. Despite its importance, however, a formal theory of grammar is only one step in the construction of a substantive linguistic theory of universal grammar. The present work adds to the formal theory of LFG a set of substantive postulates for a universal theory of control and complementation. Not all of the relevant phenomena could be treated here, but the central concepts of function, category, syntactic encoding of function in categorial structure, government, and functional and anaphoric control are discussed, and comparisons are made with alternative theories including the government and binding theory of Chomsky (1981).

The basic assumptions of the theory of control are articulated in the initial sections: first, grammatical functions are universal primitives of syntax, not derived from phrase structure representations or from semantic notions; second, grammatical functions are lexically encoded in predicate argument structures in varying ways; third, constituent structure categories are universally decomposed into features and types, the features being definable in terms of the primitives SUBJ, OBJ; and fourth, grammatical functions are syntactically encoded directly in surface representations of phrase structure, according to structural configurations or morphological features. In lexical–functional theory, the grammatical relations of a sentence—that is, the set of associations between its

This article is part of a work published by the MIT Press (J. Bresnan, ed., *The Mental Representation of Grammatical Relations*) and appears here with their kind permission.

I wish to thank K. P. Mohanan, Per-Kristian Halvorsen, Ronald Kaplan, and Marilyn Ford for detailed comments on earlier drafts of this study. I am also grateful to Jane Grimshaw, Lorraine Levin, Malka Rappaport, and Jane Simpson for suggesting several improvements. This study is based upon work supported by the National Science Foundation under Grant No. BNS 80–14730 to MIT and also by the Cognitive and Instructional Sciences Group of the Xerox Palo Alto Research Center.

Linguistic Inquiry, Volume 13, Number 3, Summer 1982
0024-3892/82/030343–92 $02.50/0

surface syntactic constituents and its semantic predicate argument structure—are represented by a functional structure (*f-structure*), which is separate from the representation of constituency relations (*c-structure*). F-structures represent grammatical relations in a universal format, while constituency relations vary radically across languages and even across constructions within a single language. In this article it will be argued that the syntactic conditions on government and control are properly stated in terms of f-structure representations, not c-structure representations. Evidence will be presented to choose between alternative theories of government and control, and a number of major generalizations will be derived from the present theory.

This study presupposes previous work in LFG, including Kaplan and Bresnan (1982), Grimshaw (1982), and Bresnan (1982b); these, together with related work, are collected in Bresnan (1982a).

1. Grammatical Functions

In the lexical–functional theory, grammatical functions are universal, syntactically primitive elements of grammar. Although grammatical functions are basic concepts in traditional grammars as well as in recent studies of a wide variety of languages (Perlmutter (to appear)), some researchers have objected to taking grammatical functions as primitives at the level of syntactic representation.

The theory of universal grammar, like other theories, can be axiomatized as a set of basic concepts and a set of postulates from which the explanatory results of the theory are deducible as theorems; all grammatical concepts of the theory can be defined in terms of the basic set of concepts, which are taken to be primitive, or undefined. In general, there are many alternative sets of primitives in terms of which theories can be axiomatized. However, Chomsky (1981) asserts that we are not free to choose our primitives in constructing an explanatory theory of universal grammar; in order to explain the fact of language acquisition, we must require that the primitives of our formal grammatical theory correspond directly to categories of prelinguistic experience. This he calls the criterion of *epistemological priority*. He then suggests that the primitive concepts in terms of which constituent structure has been defined may be "epistemologically prior" to the primitives in terms of which grammatical relations have been defined. This is the basis of Chomsky's (1981, 10) conclusion that "we should . . . be wary of hypotheses that appear to assign to grammatical relations too much of an independent status in the functioning of rule systems."

However, Chomsky's suggestion that the primitives of categorial or constituent structure are epistemologically prior to the primitives of relational or functional structures is unsupported; no evidence or argumentation is offered for this view. Indeed, this view is implausible, for if it is assumed that the primitives of constituent structure are epistemologically prior, and that grammatical relations are derived from them as defined concepts, then it is difficult to explain the invariance of grammatical relations across the radically varying constituent structures of different languages. If constituency prim-

itives were epistemologically prior, the appearance of universal grammatical relations in languages which lack sufficient constituent structure to support configurational definitions (see Perlmutter (to appear), Hale (1979), Mohanan (1982a)) would be completely mysterious.

A more fundamental problem is the criterion of epistemological priority itself. Pinker (1982) and Pinker and Lebeaux (1981) have shown that there is a natural inductive basis for the elementary concepts of category and function that are assumed in the lexical–functional theory of syntax; for example, the notion of SUBJ is grounded in the concept of animate agent or actor. However, this inductive basis is supplemented by a distributional learning algorithm that can infer the existence of nonagent subjects. As a result, according to the Pinker–Lebeaux model, the mature grammar eventually acquired by the language learner is *not* represented in terms of formal primitives that correspond directly to prelinguistic notions such as animate agent or actor. Pinker and Lebeaux argue that this model can not only explain the fact of language acquisition, which is the goal of learnability models, but also account for the course of language acquisition; that is, the model provides a consistent and coherent account of the evidence from studies of child language. In short, since well-motivated language acquisition procedures can induce classes of formal grammars whose primitive concepts bear no direct correspondence to prelinguistic categories, the criterion of epistemological priority is simply irrelevant to the goal of constructing an explanatory linguistic theory.

The doctrine that theories of higher cognitive processes should be formulated in terms of primitives that correspond directly to categories of primary experience derives from empiricist theories of knowledge that have long since been rejected in the psychology of language and perception, and which in fact have never been accepted in generative linguistic theory. The evidence suggests, on the contrary, that the human mind is innately endowed with powerful conceptual structures and operations that are only abstractly and indirectly related to "primary experience" (Miller and Johnson-Laird (1976)); elsewhere, Chomsky has also championed this view (Chomsky (1969)). The fact is that many elementary linguistic concepts play an important theoretical role in rule systems and yet cannot be reduced to primary experience as required by the criterion of epistemological priority. For example, *stress* is a basic concept of Chomsky and Halle's (1968) theory of English phonology, although phonetic research had already shown that stress has no reliable acoustical correlate in amplitude, duration, or fundamental frequency of sound (Lehiste (1970)). What *is* required of an explanatory theory of grammar is not that it be axiomatizable in terms of such "epistemologically prior" primitives, but rather that its knowledge representations can be naturally embedded in substantive theories of the processes of language acquisition and use. The theory of syntax adopted here has already begun to meet this requirement in the work of Pinker (1982), Pinker and Lebeaux (1981), Ford (1982), and Ford, Bresnan, and Kaplan (1982).

A different form of objection to considering grammatical functions as primitives could be based on *constructive elimination*, by constructing from constituency primitives alone a syntactic theory which is capable of explaining the universal properties of gram-

matical relations. Research in relational grammar (Perlmutter (to appear)) as well as studies in Bresnan (1982a) and elsewhere (Mohanan (1981a), Marantz (1981), Rappaport (1980), Roberts (1981), Simpson and Bresnan (in preparation)), shows that the invariance of grammatical relations across languages and across constructions with radically varying constituency relations is not explained by purely structure-based theories. In fact, to express certain universals, recent versions of transformational grammar now make explicit or tacit appeal to grammatical relations that cannot be eliminated through definitions in terms of independently motivated constituent structure configurations (see Chomsky (1980a, 1981); Mohanan (1981a)).

These objections to considering grammatical functions as primitives at the level of syntactic representation can therefore be rejected. The question remains, however, whether grammatical functions can be reduced to nonsyntactic notions at other levels of linguistic representation. The important work of Dowty (to appear) bears on this question. Dowty proposes a definition of the grammatical functions in terms of their universal roles in the semantic composition of sentences (assuming the framework of Montague (1973)). For example, the *object* is the argument that combines with a verb meaning to produce an intransitive verb phrase meaning, and the *subject* is the argument that combines with an intransitive verb phrase meaning to produce a sentence meaning.

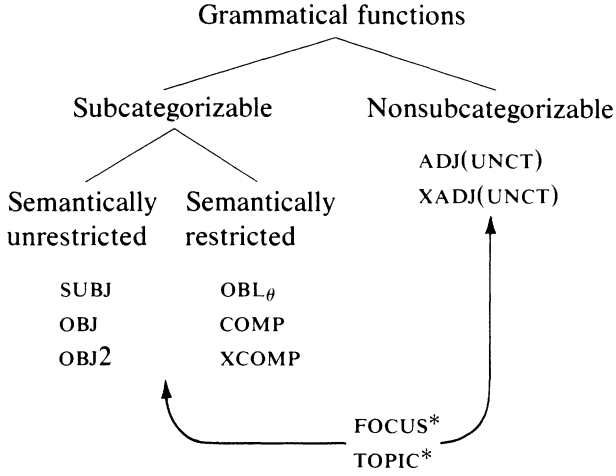
A successful definition in these terms would be a welcome advance in grammatical theory. However, while it is possible to identify semantic notions that can be put into correspondence with syntactic concepts, what must be shown in order to claim that the syntactic concepts can be eliminated is that this correspondence is explanatory rather than arbitrary. In other words, the syntactic properties of grammatical functions—such as their roles in reflexivization, causativization, passivization, and control—should not be arbitrarily associated with the proposed semantic definitions, but should be shown to follow from the appropriate semantic characterizations. Unless this is done, syntactically primitive grammatical functions will not have been eliminated by these definitions, but only renamed. For example, in terms of the semantic composition of sentences, the subject plays the same semantic role with both transitive and intransitive verbs. Yet causativization in Malayalam (as in many languages) affects the subjects of transitive and intransitive verbs in syntactically distinct ways: the subject of an intransitive base verb behaves syntactically like the object of the causativized form of the verb, but the subject of a transitive base verb behaves syntactically like an oblique instrumental of the causativized verb (Mohanan (1981a)). Since the semantic properties of the causativized subjects are the same regardless of whether they function syntactically as objects or as oblique instrumentals, there is no principled semantic distinction between these two grammatical functions. The proposed semantic definitions of grammatical functions do not explain the syntactic properties of objects and obliques (although one could of course stipulate arbitrary differences in the semantic composition of causativized transitives and intransitives in order to match the syntactic differences). We therefore conclude that the definitions do not eliminate grammatical functions, but merely rename them in semantic terms. (See Mohanan (1981a) for a similar argument

with respect to recent attempts to define grammatical functions in terms of Chomsky's government and binding theory.)

Because grammatical functions in the present theory play an important role both in lexical representations and in semantically interpreted syntactic representations, it is of course possible that lexical or sentential semantics (or both) may ultimately provide true nonsyntactic definitions of these concepts, just as many of the elementary features of phonology are definable in terms of phonetic theory. But the elimination of functions as primitives is no more necessary to an explanatory theory of grammar than is the elimination of phonological features.

The grammatical functions assumed in this study are classified in figure 1. The functions XCOMP and XADJ(unct) are distinguished as *open grammatical functions*; all others are *closed*. (Slightly different notations are used elsewhere in Bresnan (1982a); for example, Mohanan (chapter 8) uses the symbol COMP for XCOMP, and Kaplan and Bresnan (chapter 4) and Bresnan (chapter 1) use the symbols VCOMP and SCOMP for XCOMP and COMP, respectively.) The distinction between open and closed functions plays a role in the theory of control (section 8). The *subcategorizable* functions in figure 1 correspond to the *governable* functions of Kaplan and Bresnan (chapter 4 of Bresnan (1982a)); these are the only functions to which lexical items can make reference.

Not all aspects of this classification are universal; in particular, the subcategorizability of TOPIC and FOCUS appears to be a typological property that distinguishes "subject-oriented" from "topic-oriented" languages. In addition to these functions, there is assumed to be a universal set of functional features, including CASE, TENSE, NUMBER,



*The subcategorizability of these functions is a parameter that distinguishes "subject-oriented" from "topic-oriented" languages.

Figure 1
Classification of grammatical functions assumed in this theory

GENDER, PERSON, FINITE, PRED, and SPEC. In the formal representation of functional structures (Kaplan and Bresnan (1982)), function and feature names are formally distinguished by the types of their values.

The role of grammatical functions in the lexical–functional theory of syntax is to provide the mapping between surface categorial structure and semantic predicate argument structure. This is done by assigning the grammatical functions semantic roles in the lexicon and syntactic realizations in the categorial component of the syntax. A fundamental constraint on all references to functions and functional features, whether by lexical or by syntactic rules and representations, is the *principle of functional locality*: designators in lexical and grammatical schemata can specify no more than two function applications (Kaplan and Bresnan (1982)). This principle means that function designators can refer only to functions and features or to immediate subfunctions and features of the f-structure to which they apply.

2. Lexical Encoding of Grammatical Functions

In the lexical–functional theory of syntax, lexical items subcategorize for functions, not constituent structure categories, and lexical items exert their selectional restrictions on a subset of their subcategorized functions. Lexical subcategorization for function rather than structure explains the fact that when a function is freely realizable in a set of different phrase structure positions, the subcategorization restrictions of any lexical item that subcategorizes that function are satisfied by the entire set of positions (Grimshaw (1981; 1982), Montalbetti (1981)). Moreover, it provides an invariant theory of subcategorization for configurational and nonconfigurational grammars alike (section 4).

The *predicate argument structure* of a lexical item lists the arguments for which there are selectional restrictions; the *grammatical function assignment* lists the syntactically subcategorized functions (without repetitions) and may include a null symbol ϕ;[1] the *lexical form* pairs arguments with functions. This study employs the notations given in (1), which differ slightly from those in Kaplan and Bresnan (1982) and Bresnan (1980; 1982b). (The notation given here provides a simple formal interpretation of the coherence condition; see section 5.) (1) gives notation for representing properties of the lexical item *seem* as it is used in sentences like *John seems sick to Mary*.

(1) a. predicate argument structure: SEEM⟨1,2⟩
b. grammatical function assignment: {(XCOMP), (OBL_{GO}), (SUBJ)}
c. lexical form: 'SEEM-TO⟨(XCOMP)(OBL_{GO})⟩(SUBJ)'

In this usage *seem* has a dyadic predicate argument structure, whose first argument denotes a state of affairs (e.g. John's being sick) and whose second argument denotes a perceiver of that state of affairs (e.g. Mary). (1b) gives the functions that are subcat-

[1] It is important to note that ϕ is *not* a grammatical function, but a lexical symbol indicating that an argument is semantically filled in the lexicon and is not assigned any function (Bresnan (1980), Halvorsen (1981)). See examples (8) and (9) below.

egorized by *seem*. In (1c) the function XCOMP is assigned to the state-of-affairs argument, and the function (OBL_{GO}) is assigned to the perceiver argument. The notation in (1c) indicates that *seem* subcategorizes three functions but exerts its selectional restrictions only on the XCOMP and OBL_{GO} functions. The SUBJ in (1c) is a "nonlogical" or "nonthematic" function, upon which *seem* imposes no selectional restrictions. Its relation to the XCOMP is determined by the theory of control (section 8). The material inside the angled brackets in (1a,c) is called the *argument list* of the predicate; this lists the predicate's semantic arguments, which give rise to its "selectional restrictions" (see Bresnan (1982b)).

Although example (1) shows a verb that does not impose selectional restrictions on its subject, it is well established that lexical items may also impose semantic selectional restrictions on subjects. In (2) and (3), for example, *admire* selects an animate subject, while *frighten* selects an animate object:

(2) a. *Mary* admires John.
    b. #*Sincerity* admires John.
(3) a. Mary frightens *John*.
    b. #Mary frightens *sincerity*.

It has been argued that certain semantic asymmetries between subjects and nonsubjects justify the suppression of the "logical" subject argument from lexical predicate argument structure (Marantz (1981)). (The term *logical subject* involves a confusion of distinct levels of representation, as pointed out in Bresnan (1982b). In what follows, the term *logical subject argument* will be used to refer to the thematic argument of a predicate that has SUBJ assigned to it in the unmarked active lexical form.) According to this view, the selectional restrictions in (2) would be attributed, not to the verb *admire*, but to the meaning of the verb phrase *admires John*. Nevertheless, many selectional restrictions imposed on the subject ultimately reduce to properties of the lexical item, for subject selection occurs when the only content of the verb phrase is the lexical predicate itself, as in (4).

(4) a. *Mary* will relax.
    b. #*Sincerity* will relax.

The question is not, then, whether subject selection is a property of lexical items, but whether this property is represented by the same predicate argument structure that represents nonsubject selection.

Two rationales have been given for an asymmetrical representation of the logical subject argument and other arguments in predicate argument structure (Marantz (1981, 50–51)). First, the "choice of object (or other argument of a verb) affects the semantic role of the logical subject while choice of logical subject does not affect the semantic role of the object," and second, there exist object idioms but no subject idioms with free arguments. (Arguments similar to Marantz's have been attributed to Chomsky in lectures.) That is, while many idioms express properties which are noncompositionally

derived by combining a verb and an object (these are the so-called "object idioms"), there are claimed to be no idioms that express properties which are noncompositionally derived by combining a verb and a subject (these would be "subject idioms"). Examples like (5a–d) are given by Marantz to illustrate these points.

(5) a. kill an insect
b. kill a conversation
c. kill a bottle (i.e. empty it)
d. kill an audience (i.e. wow them)

The wide range of predicates expressed by the examples in (5), including literal, figurative, and idiomatic senses of *kill*, depends upon the choice of object of the verb.

As we will see, neither of these considerations is well-founded factually, but it must also be recognized that the logic of the argument that is based upon them is faulty. Note that both of these considerations are based upon the compositional semantics of sentences. The assumption is that if the subject is always the *last* argument to be semantically composed with the predicate, one can explain the generalization that the choice of nonsubject arguments does not depend on the choice of the subject argument. By suppressing the subject argument from the verb's predicate argument structure, one prevents the subject from being combined directly with the verb before the verb and its nonsubject arguments have been assembled into a predicate; the subject can then be semantically composed only with a completely formed predicate. But the issue of whether or not the subject argument has a special role in the semantic composition of the sentence is logically independent of the issue of whether or not a subject argument position should appear in lexical predicate argument structure. For example, in Dowty's (to appear) theory, the subject is always the last argument to be semantically composed with the predicate; yet the lexical function that expresses the meaning of a transitive verb in his theory contains variables for both the subject and the object arguments. In short, one could capture the subject/nonsubject generalizations without affecting the lexical representation of predicate argument structure, simply by giving the subject a distinguished role as final argument in the semantic composition of the sentence.

We will not take this approach, however, because the subject/nonsubject generalization itself is factually ill-founded, as mentioned above. First, there are in fact subject idioms with free nonsubject arguments: for example, *The cat's got x's tongue* ('*x* can't speak'), *What's eating x?* ('What is making *x* so irritable?'), *Time's up* (*for x*) ('The time (for *x*) has expired'), *x's goose is cooked* ('*x* is in trouble and there is no way out'). Second, there are clear cases in which the semantic choice of a nonsubject argument does depend upon choice of the subject. Consider, for example, (6a–c).

(6) a. The ceiling caved in on John.
b. The wall caved in on John.
c. The roof caved in on John.

Example (6c) has a figurative or metaphorical sense ('Everything went wrong for John') as well as the literal sense that part of the structure of a house collapsed on John; but examples (6a,b) are unambiguously literal. Thus, the choice of *the roof* as subject of *caved in* gives rise to a special meaning. Under this special meaning, one can choose as the object of *on* an abstract noun phrase which cannot occur with the literal meaning:

(7) a. #The ceiling caved in on John's dreams.
b. #The wall caved in on John's dreams.
c. The roof caved in on John's dreams.

How do we know that it is not the object of *on* which gives rise to the special meaning and thereby determines the choice of the subject *the roof*? When we omit the *on*-phrase altogether, we find that *The roof caved in* still has the figurative sense 'Everything went wrong' while *The ceiling caved in* and *The wall caved in* lack it. It is not difficult to find other examples which support the same conclusion (e.g. *A truck hit John* vs. *An idea hit John,* pointed out by K. P. Mohanan).

In conclusion, there is no justification for suppressing the subject argument from predicate argument structure. Moreover, theories of lexical representation which omit the subject argument from predicate argument structure so as to give it the special status of final argument in the semantic composition of the sentence are incompatible with the evidence of subject idioms and the subject-determined selection of nonsubject functions.

By separating the predicate argument structure from the grammatical function assignment, as in (1a,b), we open the possibility of finding principles which will enable us to predict the possible function–argument correspondences and so derive, or at least narrowly constrain, the set of lexical forms. Current research in lexical representation suggests that there are universal constraints on the function–argument pairings in lexical forms. One constraint which is discussed elsewhere in Bresnan (1982a) is *function–argument biuniqueness*: in each lexical form, the predicate arguments and the functions they are paired with must be in one-to-one correspondence (Bresnan (1980), Grimshaw (1982)).

Recent research suggests that there are also *semantic constraints on function–argument pairings*: there are semantically restricted functions that can only be paired with arguments of specified semantic types (see Rappaport (1980)). The specified types of arguments may include the thematic relations AG(ent), TH(eme), EXP(eriencer), SO(urce), GO(al), LOC(ation), DIR(ection), BEN(eficiary), INSTR(umental), as well as MNR (manner), MEANS (a secondary TH), CAUSEE (a secondary AG), PART, PATH, QUANT(ity), and PROP(ositional). (Nothing rests on the details of the classification of argument types; see Jackendoff (1976) and Amritavalli (1980) for some discussion.) To enforce the semantic constraints on function–argument pairings, it is convenient to assume that there is a labeling function from the argument types to the predicate arguments of each predicate argument structure. As illustrated in (8), a single argument may bear more than one label (Jackendoff (1972)).

(8) a. SEEM-TO ⟨PROP, GOAL⟩
EXP

b. BUY ⟨AG, TH, SO⟩
GO

c. SELL ⟨AG, TH, GO⟩
SO

A semantically restricted function, then, can be paired only with an argument one of whose labels matches its semantic type. In particular, the oblique functions can be paired only with an argument type whose index they carry: for example, OBL_{AG} must be paired with an AG argument; OBL_{GO}, with a GO; and OBL_{LOC}, with a LOC. The open and closed complement functions XCOMP and COMP can be paired only with what may be broadly referred to as "propositional" arguments (those labeled PROP). COMP and XCOMP differ in their syntactic encoding properties (section 4) and in their control properties (sections 8, 9). The null symbol ϕ is not a function; it appears only in argument positions, where it signifies that the argument is lexically interpreted and no function is assigned (footnote 1). For example, in (9) ϕ indicates that the AG is semantically filled in the mediopassive lexical form of *read*, and in (10) it indicates that the TH is filled in the intransitivized lexical form of the same verb:

(9) a. The novel reads easily.
b. READ ⟨ ϕ (SUBJ)⟩
AG TH

(10) a. John reads frequently.
b. READ ⟨(SUBJ) ϕ ⟩
AG TH

We can formulate these restrictions as follows. Let G be a semantically restricted function and R its associated argument type. A *designator of G* is any functional designator in the sense of Kaplan and Bresnan (1982) which mentions the function G: for example, (↑ SUBJ) is a designator of SUBJ. Now we will require that in every lexical entry, every lexical designator of G must be assigned to an argument of type R, either by appearing in an argument list at an argument of that type or by being equated with another lexical designator that does so. (The latter condition allows an extraposed COMP function, for example, to be equated with a SUBJ.) Note that these restrictions specify the *only* possibilities for lexical items to refer to semantically restricted functions; this will have consequences for the theory of control (section 8), for it restricts the set of lexically induced functional controllers ((20)) and the set of lexically induced functional anaphors ((35)). Because the open functions have broader syntactic properties, their designators may also appear in certain other contexts (see sections 4, 8, 9).

Unlike the semantically restricted functions, the semantically unrestricted functions SUBJ, OBJ, OBJ2 may be paired with any argument type or remain unpaired with an argument (as in the case of "nonlogical" subjects). There are nevertheless important constraints on their assignment to lexical forms. In particular, there appears to be a

hierarchy for the assignment of SUBJ, OBJ, and OBJ2 to predicate argument structures: in the unmarked case, OBJ2 is assigned only if OBJ has been assigned, and OBJ only if SUBJ has been (cf. Bresnan (1980) and Rappaport (1980)).

Lexical rules can alter the assignment of functions to predicate argument structures, and only lexical rules can do so, given the principle of direct syntactic encoding (Bresnan (1982b)). Lexical rules will not be motivated in this study; see elsewhere in Bresnan (1982a) and Roberts (1981), Rappaport (1980), Neidle (in preparation).

3. Categories

Constituent structure categories are decomposed into a *type*, or level of structure, and a *feature matrix* of universal categorial features as in Bresnan (1975; 1976a). The features and types of major categories assumed in this study are shown in figures 2 and 3, respectively, with their common notations. As we will see (section 4), the categorial features "predicative" and "transitive" of figure 2 can be defined in terms of the functional primitives SUBJ and OBJ. Hence, we can eliminate primitive categorial features from our theory altogether. For simplicity in representing c-structure trees in what follows, the convention will be adopted that when a nonsentential category of type 1 is exhaustively dominated by a category of type 2 or exhaustively dominates a category of type 0, the type 1 category will be suppressed.

The categories of type 0 are called *lexical categories*; the categories of types 1 and 2 are called *projections*; projections of the highest type (i.e. type 2 in figure 3) are called *maximal projections*; and lexical categories and their projections are all considered *major categories*. S and $\bar{S}$ are considered to be major categories which are projections of no lexical category (Hornstein (1977), McCloskey (1979)). There are also *minor categories* of null or degenerate type, including DET and COMP (Bresnan (1975; 1976a)) and certain uses of P(reposition) (see section 4). Not all languages have instantiations of all categories. For example, Warlpiri does not instantiate the category A; instead, the category N functions as both A and N (Nash (1980), Simpson (in preparation)). Nor do all languages distinguish all types in every category; for example, many VSO languages do not instantiate V′ or V″. In short, the instantiation of both types and features of categories is a source of variation among languages.

	"predicative"	"transitive"	
V	+	+	verbal
P	±	+	pre- or postpositional
N	±	−	nominal
A	+	−	adjectival
S	−		sentential

Figure 2
Feature matrices of major categories

type	0	1	2
category	V	V′	V″ (VP)
	P	P′	P″ (PP)
	N	N′	N″ (NP)
	A	A′	A″ (AP)
		S	$\bar{S}$

Figure 3
Major category types

C-structure rules are context-free rewriting rules (or recursive transition networks) defined over the vocabulary of major and minor categories. The set of possible rules is narrowly constrained, as discussed in the next section. Natural classes of categories and rule schemata can be designated by the X-bar (X-prime) notation as in Bresnan (1975; 1976a). For example, $\underset{[+t]}{X} = \{V, P\}$ and $\underset{[+p]}{X''} = \{VP, PP, NP, AP\}$.

4. Syntactic Encoding of Grammatical Functions

Grammatical functions are assigned to the particular c-structure rules and inflectional features of each language. The possible syntactic encodings of functions into structure are highly constrained. In *configurational encoding* a basic form of c-structure rule is, for any categorial feature matrix X, $X^{n+1} \rightarrow C_1 \ldots X^n \ldots C_m$, where $n \geq 0$ and C_i is either a minor category or a maximal projection. (Falk (1980a) suggests that this basic rule form is further decomposed into separate specifications of dominance and precedence relations.) For this rule form, the basic principle of configurational encoding is to associate a function-assigning equation $(\uparrow \text{G}) = \downarrow$ with each C_i if and only if C_i is a maximal projection, and to associate the equation $\uparrow = \downarrow$ elsewhere. A major category bearing the equation $\uparrow = \downarrow$ is called the *head*. Hence, according to the basic principle of configurational encoding, all and only the maximal projections of categories can bear functions, every phrase has a unique head, and the functional features of a phrase are identified with those of its head. It follows immediately that in predominantly configurational languages, only maximal projections of categories will appear to be subcategorized for (section 1); this is a widely recognized characteristic of subcategorization.

The X-bar theory must be elaborated to permit other rule forms such as those for coordinate structures and for exocentric constructions (Bloomfield (1933)). The syntax and semantics of coordination in the present theory are discussed in Andrews (1981), Peterson (1981a), and Halvorsen (1981). Among the exocentric constructions of English are the sentence (e.g. *John walks slowly*) and the gerund phrase (e.g. *John's walking*

slowly); in these the head is VP, a maximal projection of V, while the dominating category is not a projection of V. Examples are given in (11) and (12). (Here I am adapting the analysis of the English S proposed by Falk (1980b).) In the gerund construction of rule (12), the case-marking constraints are omitted.

(11) $S \rightarrow$ NP VP
$(\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow$

(12) $NP \rightarrow$ NP 's VP
$(\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow$

Because it is the projection of no lexical category, S is an exocentric category in all languages. S can be headed by VP (as in English and other SOV languages), by V (as in VSO languages), by nominals (as in Warlpiri (Simpson (in preparation))), and by adjectival phrases (as in Russian (Neidle (in preparation))).

Exocentric rules fall under the rule form $X^n \rightarrow C_1 \ldots C_m$, where $n > 0$ and C_i is either a minor category (e.g. *'s*) or a maximal projection of a lexical category. The principle of function assignment for rules of this form is to associate the equation $\uparrow = \downarrow$ with a single major category (the head) and to associate a function assignment equation $(\uparrow \text{G}) = \downarrow$ with all other maximal projections. Consequently, for both exocentric and endocentric categories, the head is a major category annotated with $\uparrow = \downarrow$, while all other major categories (and only these) are annotated with $(\uparrow \text{G}) = \downarrow$. Thus, it remains true that in configurational encoding, only the maximal projections of categories can bear functions, every phrase has a unique head, and the functional features of a phrase are identified with those of its head.

In *nonconfigurational encoding,* the basic form of c-structure rule is $C \rightarrow X^*$, where C is a major nonlexical category and X is a lexical or nonlexical category. The basic principle of nonconfigurational encoding is to associate pairs of function-assigning and feature-assigning equations of the form given in (13a) with arbitrary X:

(13) a. $\left\{ \begin{array}{l} (\downarrow \text{F}) = \text{v} \\ (\uparrow \text{G}) = \downarrow \end{array} \right\}$

b. $\uparrow = \downarrow$

In (13a), G is a function selected by the value v of the feature F. For example, taking F to be CASE and v to be NOM or ACC, $(\downarrow \text{CASE}) = \text{NOM}$ could be associated with $(\uparrow \text{SUBJ}) = \downarrow$, and $(\downarrow \text{CASE}) = \text{ACC}$ with $(\uparrow \text{OBJ}) = \downarrow$. These pairs of schemata are arbitrarily associated with categories in c-structure rules; see Mohanan (1982a) for a detailed example. The schema $\uparrow = \downarrow$ in (13b) is also associated arbitrarily with categories in c-structure rules. (The association with $\uparrow = \downarrow$ could also depend upon some inflectional feature.) The *head* of C is defined to be any major category which is annotated with $\uparrow = \downarrow$ and which has a PRED. It follows from the consistency condition (section 5 and Kaplan and Bresnan (1982)) that the head is unique.

In nonconfigurational encoding as in configurational encoding, every phrase has a unique head, and the features of a phrase are identified with those of its head. However,

in nonconfigurational encoding, functions need not be assigned to maximal projections; instead, they may be assigned to submaximal projections, or even to single lexical categories, as in the example given in figure 4. This example is based on the analysis of Warlpiri given in Simpson and Bresnan (in preparation). Although alternative assignments of functions to categories are possible in figure 4, the conditions on well-formedness of f-structures (section 5) eliminate "incorrect" assignments, admitting only those that yield consistent, coherent, and complete f-structures.

In sum, the fundamental difference between configurational and nonconfigurational syntactic encoding lies in the surface realizations of grammatical functions. In configurational encoding, functions are identified by the category and by the order of maximal constituents within the immediately dominating phrase, while in nonconfigurational encoding, functions are identified by the case and other inflectional features of unordered, possibly submaximal, constituents. It follows from this theory that in predominantly nonconfigurational languages, only the case marking of constituents will appear to be subcategorized for, and not their phrase structure configurational properties.

Languages may employ both types of encoding, as observed by Mohanan (1981a). In English, the OBJ and OBJ2 are configurationally encoded. Rule (14) illustrates this.

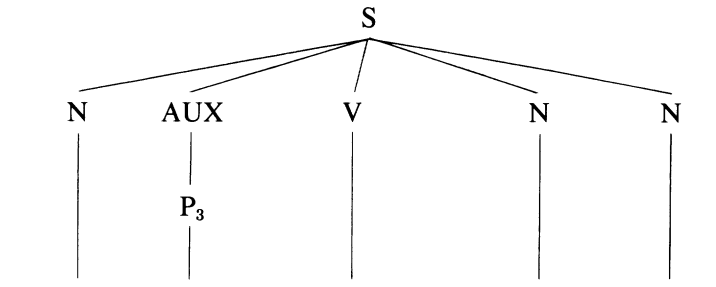
$$(14)\quad V' \rightarrow \underset{\uparrow = \downarrow}{V} \left(\underset{(\uparrow \text{ OBJ}) = \downarrow}{\text{NP}}\right) \left(\underset{(\uparrow \text{ OBJ2}) = \downarrow}{\text{NP}}\right) \underset{\left\{\begin{array}{c}(\uparrow \text{ OBL}_\theta) = \downarrow \\ (\downarrow \text{ PCASE}) = \text{OBL}_\theta\end{array}\right\}}{\text{PP*}}$$

In (14), the function names OBJ and OBJ2 are assigned, respectively, to the first and second NPs immediately dominated by VP. However, the OBL_θ function names are assigned on the basis of the features of the unordered PPs. The pair of equation schemata assigned to PP* in (14) actually abbreviates the disjunction of a finite set of pairs, one for each value of θ: $\{(\uparrow \text{OBL}_{\text{AG}}) = \downarrow, (\downarrow \text{PCASE}) = \text{OBL}_{\text{AG}}\}$ or . . . or $\{(\uparrow \text{OBL}_{\text{GO}}) = \downarrow, (\downarrow \text{PCASE}) = \text{OBL}_{\text{GO}}\}$. The PCASE features are borne by prepositions, which can serve as case markers in English (Bresnan (1979)). For example, the preposition *to* carries the lexical information $(\uparrow \text{PCASE}) = \text{OBJ}_{\text{GO}}$ and the preposition *by* carries the equation $(\uparrow \text{PCASE}) = \text{OBL}_{\text{AG}}$. Since the value of the PCASE feature is always identical with the function name of the PP in rule (14), we can abbreviate the disjunction of these pairs of equations by substituting equals for equals. Thus, we replace the symbol OBL_θ in $(\uparrow \text{OBL}_\theta) = \downarrow$ by the designator $(\downarrow \text{PCASE})$, whose value is OBL_θ: the result is the single equation $(\uparrow (\downarrow \text{PCASE})) = \downarrow$ that appears in rule (15).

$$(15)\quad V' \rightarrow \underset{\uparrow = \downarrow}{V} \left(\underset{(\uparrow \text{ OBJ}) = \downarrow}{\text{NP}}\right) \left(\underset{(\uparrow \text{ OBJ2}) = \downarrow}{\text{NP}}\right) \underset{(\uparrow (\downarrow \text{ PCASE})) = \downarrow}{\text{PP*}}$$

We see, then, that despite its superficial dissimilarity from the schemata in (14), the English PP schema in (15), which is discussed in Kaplan and Bresnan (1982), is an instance of nonconfigurational encoding. (It should be remembered that the rules just discussed are not the only sources of PP in English; there are PP adjuncts and complements as well (Bresnan (1979; 1980)). The introduction of various types of PPs under

c-structure



'The small child is chasing the dog.'

syntactic encoding of functions

$$\left\{\begin{matrix}(\uparrow \text{SUBJ})=\downarrow \\ (\downarrow \text{CASE})=\text{ERG}\end{matrix}\right\}, \left\{\begin{matrix}(\uparrow \text{OBJ})=\downarrow \\ (\downarrow \text{CASE})=\text{ABS}\end{matrix}\right\}, \text{etc.}$$

f-structure

$$\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix}
\text{CASE} & \text{ERG} \\
\text{PRED} & \text{'CHILD'} \\
\text{ADJ} & \begin{bmatrix}
\text{SUBJ} & \begin{bmatrix}\text{PRED} & \text{'PRO'}\end{bmatrix} \\
\text{PRED} & \text{'SMALL}\langle(\text{SUBJ})\rangle\text{'} \\
\text{CASE} & \text{ERG}
\end{bmatrix}
\end{bmatrix} \\
\text{PRED} & \text{'CHASE}\langle(\text{SUBJ})(\text{OBJ})\rangle\text{'} \\
\text{TENSE} & \text{NONPAST} \\
\text{ASPECT} & \text{PRES-IMPERF} \\
\text{OBJ} & \begin{bmatrix}
\text{CASE} & \text{ABS} \\
\text{PRED} & \text{'DOG'}
\end{bmatrix}
\end{bmatrix}$$

Any order of constituents is possible (so long as AUX is second); the grammatical relations of the sentence remain invariant. The functions are syntactically encoded by case. (Simpson and Bresnan 1982)

Figure 4
Nonconfigurational encoding in Warlpiri

different nodes, such as S, VP (V″), and V′, may give rise to some ordering constraints among prepositional phrases.)

A second example of mixed configurational and nonconfigurational encoding in one grammar is the rule S → X AUX X* of Warlpiri, where the minor category AUX has a fixed position in the configuration of S (Hale (1979)); in addition, there is evidence that Warlpiri has optional phrasal constituency in nominals as well as infinitival phrase constituents (Nash (1980), Simpson and Bresnan (in preparation), Simpson (in preparation)).

Fundamental to X-bar theory are the principles of syntactic encoding which relate possible functions to categorial features. For example, it was pointed out in Bresnan (1975; 1976a) that the "verbal" ([+V]) categories V and P take direct object NPs while the "nominal" ([+N]) categories A and N do not. Subsequently, Jackendoff (1977) made the important observation that the categories of natural language are better defined in terms of their relations to the subject and object functions than in terms of the categorial features [±N], [±V] of Chomsky (1970). This insight has a very natural expression in lexical-functional theory. The "intransitive" (i.e. [−t]) categories are those which do not permit (↑ OBJ) = ↓ to be annotated to any symbol within the phrase structure rules that expand the intransitive category. This means that adjectival phrases and noun phrases may not contain phrasal direct objects, while verb phrases and pre- or postpositional phrases may.[2] The "predicative" (i.e. [+p]) categories are those which do not permit (↑ SUBJ) = ↓ to be annotated to any symbol within the phrase structure rules that expand the predicative category. This means that VP, AP, and predicative NP and PP cannot contain phrasal subjects, while S and nonpredicative NP and PP may. (We have already seen that S and NP may contain phrasal subjects; a possible example in English of a PP which contains a phrasal subject is the absolutive PP construction involving *with*: *With John impossible to talk to, Mary left the room, With John away, Mary was happy*.) Thus, we see that we can eliminate primitive categorial features from our theory altogether.

Not only can the basic phrase structure categories of natural language be defined in terms of the primitive functions SUBJ and OBJ, but the function of a category is a fairly good predictor of what are likely to be the categorial and case features of the category. For example, knowing that a category C is a SUBJ, OBJ, or COMP, one can predict that C is likely to be (respectively) a nominative NP, an accusative NP, or an S̄, although S̄ can also have the SUBJ function, and PP, the OBJ function (Grimshaw (1981)), and although nominative NPs can be objects and accusative NPs, subjects (Andrews (1982)), Mohanan (1982a)). Where function and category diverge from their typical correlations, it is the function that correctly predicts the direction in which linguistic rules generalize in such domains as subcategorization, agreement, and control (Grimshaw (1981), Neidle (1982; in preparation), Mohanan (1982a), Simpson (in preparation), Simpson and Bresnan

[2] It is clear that As and Ns may have oblique NP objects (Maling (to appear)). In Russian, adjectives appear to have NP objects; whether or not they must be analyzed as direct objects deserves further investigation.

(in preparation)). To express the typical, or unmarked, relations between function and category, however, we postulate substantive constraints on the pairing of functions with categorial and case features. For example, we may say that NOM case is the unmarked feature encoding the SUBJ function. Similarly, we may postulate substantive constraints on the pairing of functions with categories. In particular, the predicative/nonpredicative categorial feature is correlated with the open/closed function distinction.

There are two logically possible ways of pairing the predicative/nonpredicative categories with the open/closed functions: we could say that the predicative categories must have closed functions, and the nonpredicative categories, open functions; or the opposite. It turns out that other properties of our theory dictate which correlation is correct: *a category is predicative if and only if it can be assigned an open function.*[3] This principle of association is motivated by very general conditions on f-structures; only by the function–category associations asserted in this principle can it be ensured in all cases that the f-structures of categories will be consistent and complete (section 5) with respect to their subject functions. For predicative categories lack structural subjects; the assignment of an open function to them induces functional control, which obligatorily provides a subject from a category-external source (section 8.2). If these categories were assigned closed functions, a subject could not always be ensured. In contrast, nonpredicative categories may have structural subjects; the assignment of a closed function to them precludes functional control and thereby prevents the assignment of a subject from an external source, which could create inconsistent f-structures. As we will see below (sections 9.2, 9.6), this principle of function–category association has a number of explanatory consequences.

We digress briefly to point out that the analysis of PPs given here differs in two ways from the analysis presented in Kaplan and Bresnan (1982). First, wherever possible, the PCASE values used here are drawn from a universal set of feature values {OBL_{GO}, OBL_{AG}, OBL_{INST}, etc.} rather than from the English-particular set {TO, BY, WITH, etc.}. This is motivated by the need for a universal format for representing grammatical relations (see Bresnan (1981), Pinker (1982), and Halvorsen (1981)). The choice of which prepositions mark which oblique functions appears to depend upon the meaning of the prepositions in their predicative uses. In *This road is to London*, the preposition *to* designates a semantic relation between the SUBJ (*this road*) and the prepositional OBJ (*London*), in which the object is the goal of some motion. Similarly, in *This work is by a famous sculptor*, the preposition *by* designates a semantic relation between the SUBJ (*this work*) and the prepositional OBJ (*a famous sculptor*), in which the object is the agent of some action. We can therefore assume that each case-marking preposition is selected from the set of semantically related predicative prepositions. That is, the preposition that serves as a formal marker of the OBL_θ function is chosen from the set of prepositions whose lexical forms associate their prepositional OBJ with the argument type θ. This

[3] Fassi Fehri (1981) has proposed the further restriction that only S can be assigned the closed COMP function.

assumption supports the "semantic bootstrapping" theory regarding acquisition of grammars described in Pinker (1982) and also accounts for the fact that the case-marking uses of prepositions appear to be related to their predicative uses cross-linguistically. There do, however, appear to be some idiosyncratically marked oblique objects. To account for these, we will extend our notation to OBL_{FORM}, where FORM is a specific preposition name. (For example, in *John relies on Mary,* we could say that *on Mary* is an OBL_{ON}.) Future research may reveal a semantic motivation for these apparently idiosyncratic markings.

The second difference between this analysis of PP and the analysis of Kaplan and Bresnan (1982) is that the oblique PP is treated here as an exocentric constituent in which the NP actually functions as the head of the PP and the P is a minor category that contributes only the PCASE feature. Thus, it is assumed that in structures of the form $[_{\text{PP}}$ P NP], either P is the head and NP is an OBJ (yielding the endocentric, predicative PP) or else P is a minor category and NP is the head (yielding the exocentric, oblique PP). The motivation for this is to eliminate compound names (e.g. TO OBJ in Kaplan and Bresnan (1982)) for oblique objects. Andrews (1982) has shown that Icelandic verbs which subcategorize for oblique prepositional objects may also govern the case marking of the NP within the PP; if the oblique prepositional object had a compound name, the case-marking restriction on this NP would violate the functional locality principle (for example, no verb could specify "$(\uparrow$ TO OBJ CASE$)$ = (ACC)").[4]

In conclusion, this theory of syntactic encoding has three main results. First, it provides a characterization of the syntactic notion *head of a phrase* which holds universally for endocentric and exocentric constituents and for configurational and nonconfigurational structure types. Second, it unifies the theories of lexical subcategorization for configurational and nonconfigurational languages. And third, it identifies the basic parameter of syntactic variation for configurational and nonconfigurational languages (see Simpson and Bresnan (in preparation)).

5. Representation of Grammatical Relations

The lexical and syntactic encodings of grammatical functions determine a mapping, or a set of associations, between the word and phrase configurations of a language and its semantic predicate argument structures. The grammar of the language assigns to each sentence a set of these associations between surface form and predicate argument structures; these are the *grammatical relations* of the sentence.

Grammatical relations are formally represented by f-structures, which are pairs of

[4] Andrews (1982) presents an analysis of PPs that differs from the one assumed here. Given our revision, case marking by a preposition can still be distinguished in f-structure from case marking by a nominal inflection, in that the former is expressed by PCASE and the latter by CASE. An example of the need for this distinction is the use of prepositional and nominal instrumentals in Malayalam (Mohanan (1982a)); to express $\text{OBL}_{\text{INSTR}}$, the instrumental case and the instrumental postposition are used in free variation; to express the OBL_{AG} in passives, only the instrumental case is used; and to express the causativized SUBJ, only the postposition is used.

fnames and *fvalues*. (This differs from the terminology used in Kaplan and Bresnan (1982), where *fnames* and *fvalues* are referred to as *attributes* and *values*.)

(16) $\begin{bmatrix} \text{fname}_1 & \text{fvalue}_1 \\ \vdots & \vdots \\ \text{fname}_n & \text{fvalue}_n \end{bmatrix}$

The *uniqueness* condition on f-structures (also called *consistency*) requires that every fname have a unique value. An *fname* is a symbol denoting one of a universal set of features and functions; an *fvalue* is a symbol (denoting one of a universal set of feature values), a semantic form, an f-structure, or a set of fvalues. Semantic forms differ from symbols in that they are uniquely instantiated when their designators are instantiated (Kaplan and Bresnan (1982)). Consequently, two otherwise identical f-structures with separately instantiated semantic forms are formally distinct. For linguistic motivation for this property, see Grimshaw (1982) and Montalbetti (1981). There are two kinds of semantic forms: those with argument lists—called *lexical forms*—and those without.

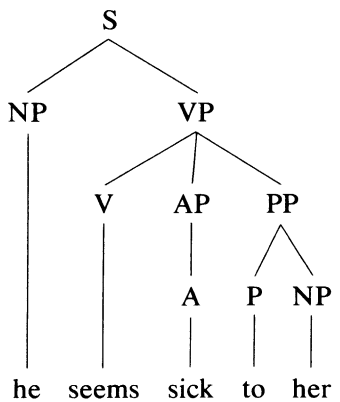
Certain f-structures are distinguished as *clause nuclei*; these contain an fname PRED whose fvalue is a lexical form (i.e. a predicate argument structure paired with a grammatical function assignment). For example, the f-structure shown in figure 5 has two clause nuclei, the f-structure labeled f_1 and its subsidiary f-structure labeled f_2. (The f-structure of figure 4 contains three clause nuclei.)

The clause nucleus is the domain of lexical subcategorization in the sense that it makes locally available to each lexical form the grammatical functions that are subcategorized by that form. Lexical subcategorization is enforced by the requirement that every f-structure be both *coherent* and *complete*. An f-structure is locally coherent if and only if all of the subcategorizable functions that it contains are subcategorized by its PRED; an f-structure is then (globally) coherent if and only if it and all of its subsidiary f-structures are locally coherent. Similarly, an f-structure is locally complete if and only if it contains values for all of the functions subcategorized by its PRED; and an f-structure is then (globally) complete if and only if it and all of its subsidiary f-structures are locally complete. This treatment of coherence and completeness permits a very simple formal interpretation: to be locally complete and coherent, an f-structure must contain subcategorizable functions $G_1, \ldots, G_n$ if and only if it contains a PRED whose value is a lexical form with the grammatical function assignment $\{G_1, \ldots, G_n\}$.[5]

The level of f-structure differs crucially from representations of constituent structure

[5] Note that in this formulation of the coherence condition, *local* is interpreted to mean 'in the same f-structure'. There is an alternative formulation of coherence, based on the functional locality principle (section 1), which interprets *local* to mean 'in the f-structure that immediately contains or is contained in this f-structure': an f-structure is locally coherent if and only if the values of all of the subcategorizable functions that it contains are subcategorized by a PRED; by the functional locality principle, only a local PRED (in the second sense) can subcategorize these function values (cf. Kaplan and Bresnan (1982)). The two formulations have empirically distinguishable consequences; for example, the formulation adopted in the text above is inconsistent with the use of compound function names such as "TO OBJ". Further research is required to determine which interpretation of coherence is optimal. Note also that, to give a uniform definition of *head*, we have adopted the instantiation procedure of Kaplan and Bresnan (1982, footnote 6).

a. *c-structure*



c. *f-structure*

$$
f_1\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix} \text{GEND} & \text{MASC} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \\ \text{CASE} & \text{NOM} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\
\text{TENSE} & \text{PRES} \\
\text{PRED} & \text{'SEEM-TO}\langle(\text{XCOMP})(\text{OBL}_{\text{GO}})\rangle(\text{SUBJ})\text{'} \\
\text{OBL}_{\text{GO}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{GO}} \\ \text{CASE} & \text{ACC} \\ \text{GEND} & \text{FEM} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\
\text{XCOMP} & f_2\begin{bmatrix} \text{SUBJ} & \\ \text{PRED} & \text{'SICK}\langle(\text{SUBJ})\rangle\text{'} \end{bmatrix}
\end{bmatrix}
$$

f_1 and f_2 are clause nuclei

b. *c-structure annotated with functional schemata*

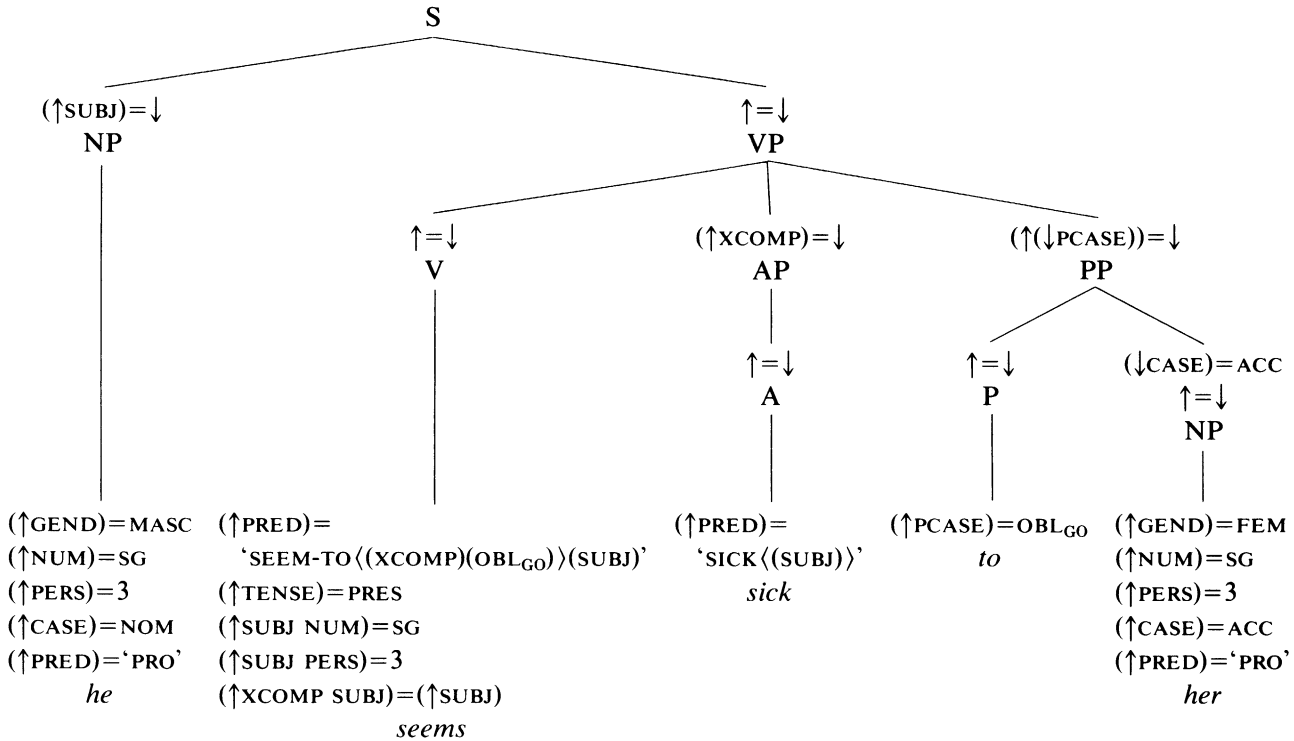


Figure 5
Clause nuclei

in that it represents grammatical relations universally, locally with respect to lexical subcategorization, and compositionally with respect to semantic interpretation (Bresnan et al. (in preparation), Simpson and Bresnan (in preparation)). See Halvorsen (1981) for a compositional procedure for the semantic interpretation of f-structures, which accounts for quantification, adverb scope, control, the distribution of intensional objects, and the interpretation of idiom chunks and grammatical expletives.

6. Representation of Constituency

The representation of grammatical relations by f-structures rather than by c-structures permits a massive simplification in the representation of the constituency relations of natural language. These can now be represented by a context-free phrase structure grammar (or a recursive transition network), defining a single level of c-structure representation over a small set of universal features and types.[6] The form of c-structure rules is largely predictable from the theory of syntactic encoding, and the categorial features are defined in terms of the primitives SUBJ and OBJ (section 4). Since there is only a single level of c-structure to represent the syntactic constituency relations among the words of a sentence, syntactic operations that restructure the dominance or precedence relations among the words and their affixes are eliminated. Consequently, only fully inflected and morphologically complete words are lexically inserted into phrase structures, as Selkirk (1981), Lieber (1980), Lapointe (1980), Nash (1980), and Mohanan (1982c) have argued on independent grounds. A further consequence is that any "restructuring" rules that alter word structure (as proposed by Chomsky (1977a), Hornstein and Weinberg (1981)) must be presyntactic rules of morphology, as justified by Bresnan (1982b), Rothstein (1981), Peterson (1981b). In addition, it follows that there are no structure-dependent operations that map c-structures into c-structures, although there are structure-dependent operations that map c-structures into f-structures, and there may be restricted kinds of operations that map c-structure rules into c-structure rules in a function-preserving fashion (cf. Kaplan and Bresnan (1982), Bresnan (1982b), Gazdar (in press)). Thus, syntactic transformations are eliminated.

Finally, the use of null c-structure can be virtually eliminated, restricting both the expressive and the generative power of grammars (Kaplan and Bresnan (1982)). The formal theory of grammatical representation permits a nonterminal node to exhaustively dominate the empty terminal string *e* only in cases of *constituent control*, the long-distance structure-dependent relation that characterizes "*wh*-movement" and similar constructions, as illustrated in figure 6.

[6] By themselves, context-free grammars are not sufficient to represent the constituency relations of natural language. Bresnan et al. (in preparation) show that there is no context-free phrase structure grammar that strongly generates just the set of empirically motivated c-structure representations for Dutch. However, the correct representation of constituency relations can be strongly generated by letting the lexical subcategorization conditions on f-structures filter the associated c-structures.

7. Government

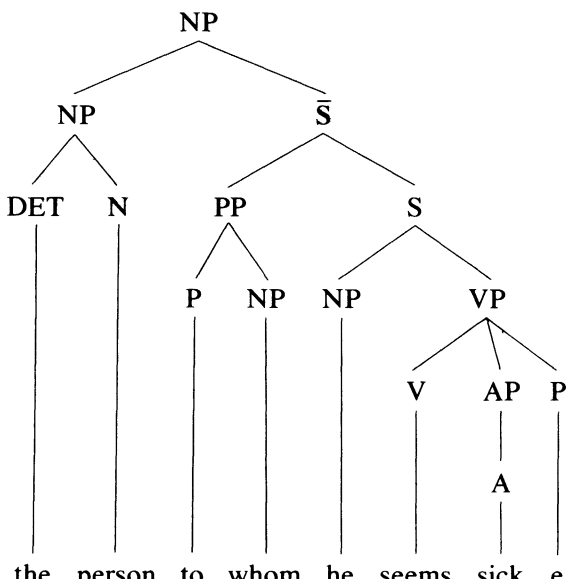
Government refers to the ability of lexical items to determine the features of other constituents. The term *agreement* is often used in preference to *government* when an inflectional morpheme of a word and an inflectional morpheme of a governed constituent have mutually constraining features. Prepositions govern the case of their objects (Andrews (1982)); verbs and adjectives govern the case, person, or number marking of their subjects, objects, and oblique objects (Andrews (1982), Mohanan (1982a,b), Maling (to appear), Craig (1977), Roberts (1981), Robertson (1976)); complementizers and verbs may govern the finiteness, mood, and constituent control features of their clauses (Zaenen (1981), McCloskey (1979)).

The fundamental problem of government is to explain the observed structural relations between governing morphemes and the constituents they govern. The theory of syntax developed in the preceding sections already provides a solution to this problem. Inflectional affixes are lexical items which must be combined with other lexical items prior to lexical insertion (Selkirk (1981), Lieber (1980), Bresnan (1982b), Mohanan (1982c)). Syntactic information may be encoded in any lexical item in the form of functional features; inflectional affixation sums the feature set of the inflectional morpheme and the feature set of the lexical category that it inflects (Bresnan (1982b)). After lexical insertion, the features of a word become the features of the phrase of which the word is either the head or a minor category (as defined in section 4). Thus, referring to figure 6b, we see that the features of the preposition become features of the PP that dominates it, the features of the determiner become features of the NP, and the features of the verb become features of the $\bar{S}$.

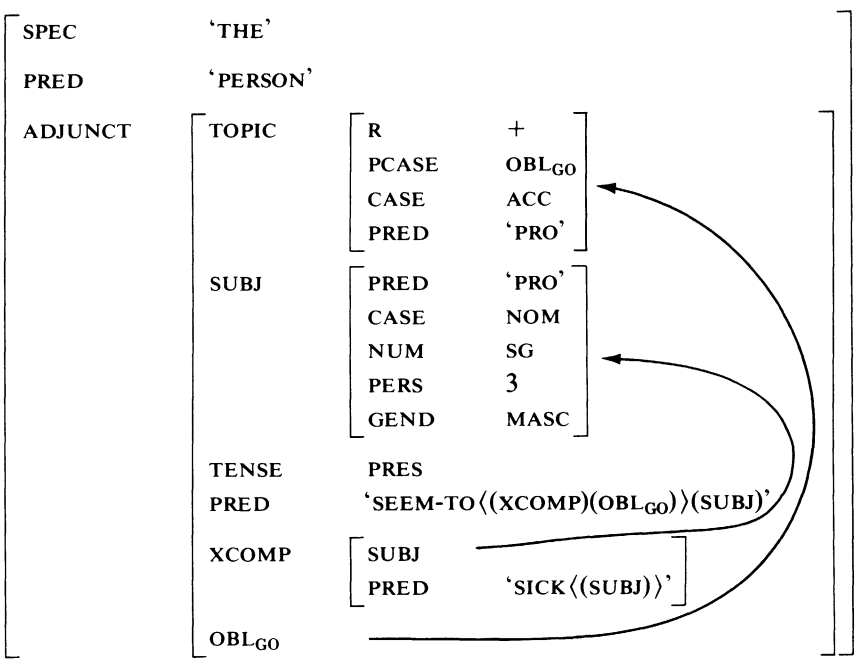
Consider in particular how a verb bearing the present tense inflectional affix *-s* governs (i.e. agrees with) features of the subject. As in Kaplan and Bresnan (1982) and in Bresnan (1982b), *-s* encodes the information that the SUBJ person is 3 and the SUBJ number is singular. When *-s* is affixed to *seem,* this information becomes part of the feature set of *seems*. When *seems* is lexically inserted into V in figure 6b, the feature set is propagated up the tree: by the $\uparrow = \downarrow$ equations, the features of the (f-structure of the) V become features of the (f-structure of the) VP, and the features of (the f-structure of) VP become features of (the f-structure of) S and ultimately of $\bar{S}$. The uniqueness principle guarantees that any subject that the clause contains will have the features required by the verbal affix and vice versa. Now suppose, hypothetically, that a morpheme similar to *-s* had been affixed to *whom,* the prepositional object of *to*. At lexical insertion, the features of this hypothetical *whom-s* bearing the information about the number and person of the SUBJ would again be propagated up the tree—but only as far as the PP dominating *whom*. Because this PP is neither a head nor a minor category in $\bar{S}$ or VP, its features will not propagate to $\bar{S}$. Hence, the syntactic information borne by the affix will remain within the PP without affecting the SUBJ of the clause.

Our theory of syntax thus provides an explanation for the fact that governing mor-

a. *c-structure*



c. *f-structure*



b. *c-structure annotated with functional schemata*

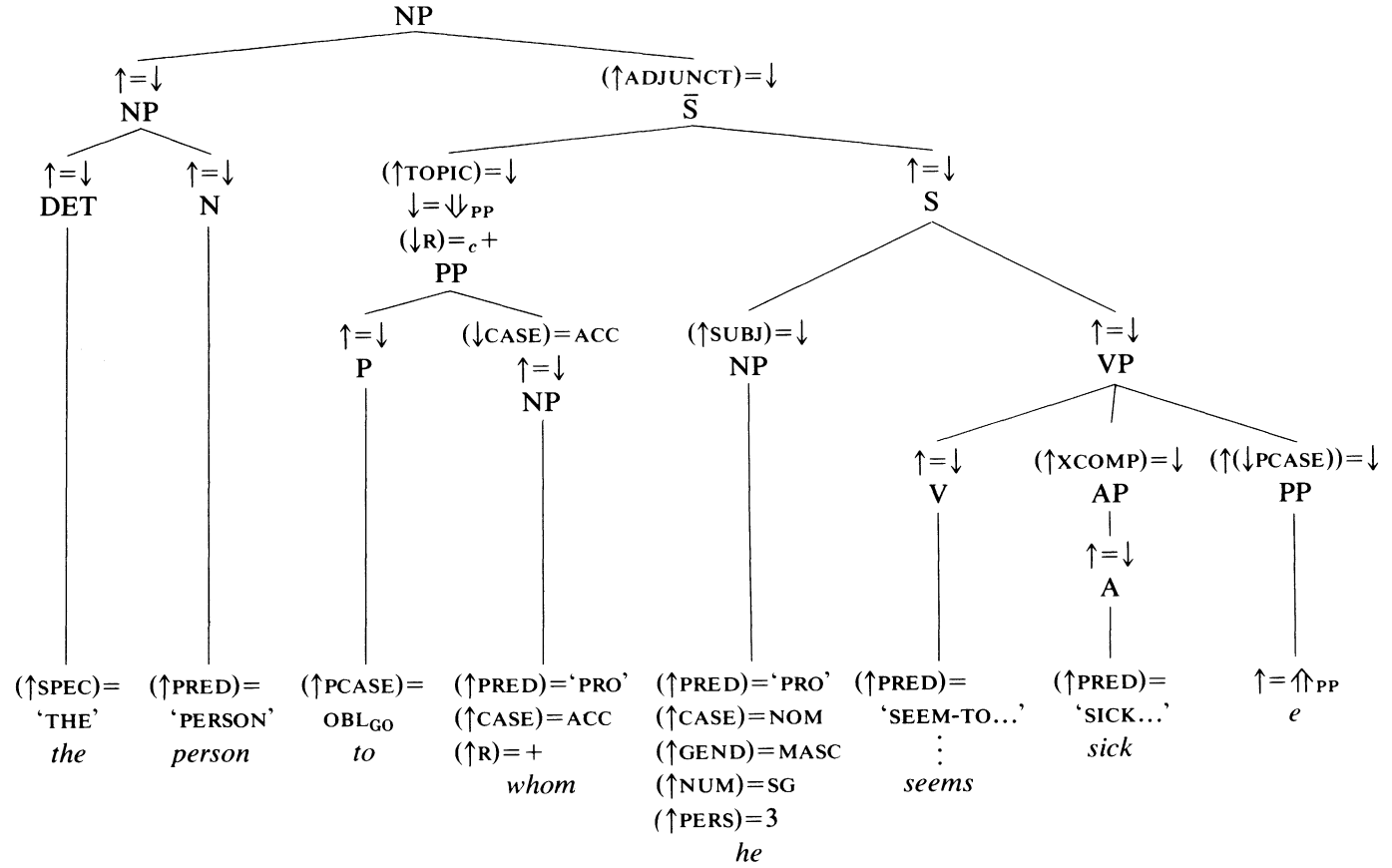


Figure 6
Constituent control

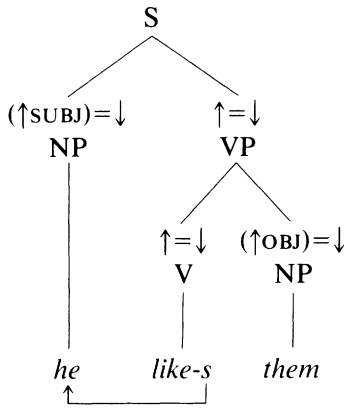
phemes appear either in the heads (or heads (of heads . . .)) or in minor categories of the phrase whose constituents they govern. There is no need to stipulate what the domains of government are, as much recent work does (Chomsky (1981)): the possible governing relations follow from the lexical theory of inflectional affixation (section 6) and the theory of syntactic encoding (section 4)—both of which are independently motivated subtheories.

This theory has a further consequence of interest. Because government relations follow from the functional relations of categories and not from a set of stipulated structural configurations, the same kinds of government relations can appear in languages with very different c-structure configurations as well as in different c-structures within the same language. Figure 7 illustrates several different configurations in which verbal affixes govern (i.e. agree with) the person and number or case features of the SUBJ (and OBJ): number and person agreement in English (figure 7a); case and person agreement in the Mayan language Jacaltec (figure 7b) (see Craig (1977, 108)); animacy and person agreement in the Athapaskan language Navajo (figure 7c). In this last example, the same affix *-yi-* determines the number and person of the SUBJ and OBJ (Roberts (1981)). Figure 8 illustrates two configurations in Dutch in which the verb *glimlachen* 'to smile' governs the oblique PP object, determining the choice of the preposition *naar* 'at'; for justification of these structures, see Evers (1975) and Bresnan et al. (in preparation). In both structures 8a,b, V_3 governs the preposition, but this similarity is not captured at the level of c-structure representation; rather, it is expressed in the *functional* relation between the verb *glimlachen* and its oblique object, which holds regardless of the very different structural configurations that express it. This result follows from the uniqueness principle: the functions in each functional structure must be unique; hence, the XCOMP within the XCOMP in figure 8b, though it is described in discontinuous c-structure positions (namely, at both VP_3 and V_3), forms a single functional unit containing both the OBL_θ and the PRED of *glimlachen*. See Bresnan et al. (in preparation) for details.

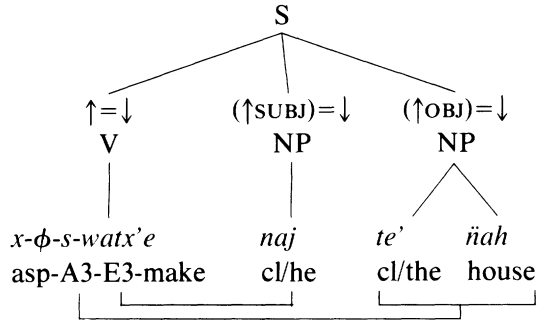
We see, then, that a word may govern the functions in the f-structure that immediately contains the word's features. No set of c-structure categories provides an invariant universal characterization of this domain of government. However, given the syntactic encoding of functions into the c-structure rules of a particular language or language type, one can predict in which c-structure configurations government relations will hold in that language or type.

These results—the explanation of the distribution of governing morphemes, the uniform cross-configurational characterization of government relations, and the predictability of the actual structural configurations in which government relations can appear—depend on the assumption that verbs govern their subjects. But this assumption is in fact a necessary consequence of the theory of syntax adopted here. Recall from the theory of lexical encoding that lexical items exert selectional restrictions on a subset of their subcategorized functions (section 2), and that verbs impose selectional restrictions on their subjects. Thus, verbs are subcategorized for their subjects. But then it follows from the completeness and coherence conditions (section 5) that f-structures

a. *English (number and person agreement)*

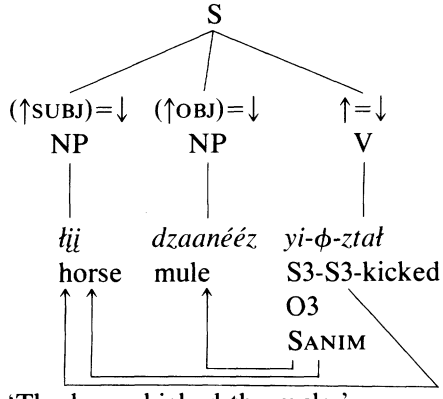


b. *Jacaltec (case and person agreement)*



'He made the house.'

c. *Navajo (animacy and person agreement)*



'The horse kicked the mule.'

Figure 7
Verbal inflections governing (agreeing with) properties of SUBJ

which contain the PREDS of verbs must also contain SUBJS (unless, of course, a particular verb in question is subjectless, as in cases of "spontaneous demotion" (Comrie (1977), Sridhar (1979))). Therefore, verbs govern their subjects.

By the same reasoning, we can establish the more general consequence that *lexical items govern all of their subcategorized functions*. Since we have assumed that the only functions that can be referred to in lexical entries are the subcategorizable functions (section 1), it follows that the subcategorizable functions and the governable functions are one and the same. This is an extremely strong substantive constraint on government which further research must test and may perhaps modify.

To summarize, we see that several major results follow from the theory of syntax proposed here: first, that governing morphemes universally appear either in the heads (or heads (of heads . . .)) or in minor categories of the phrases whose constituents they

a. *c-structure 1*

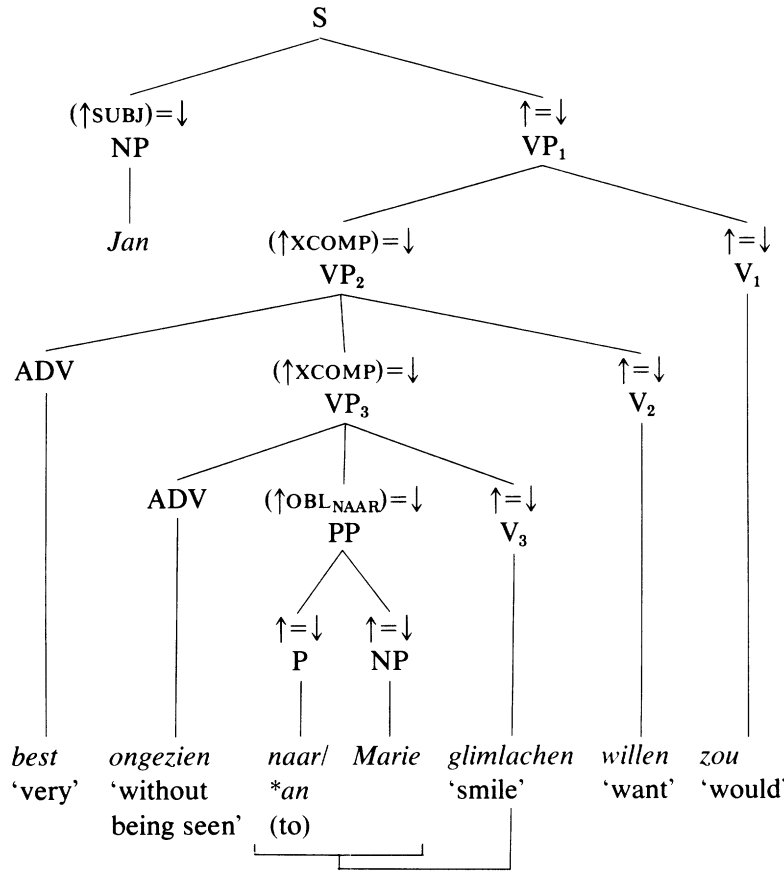
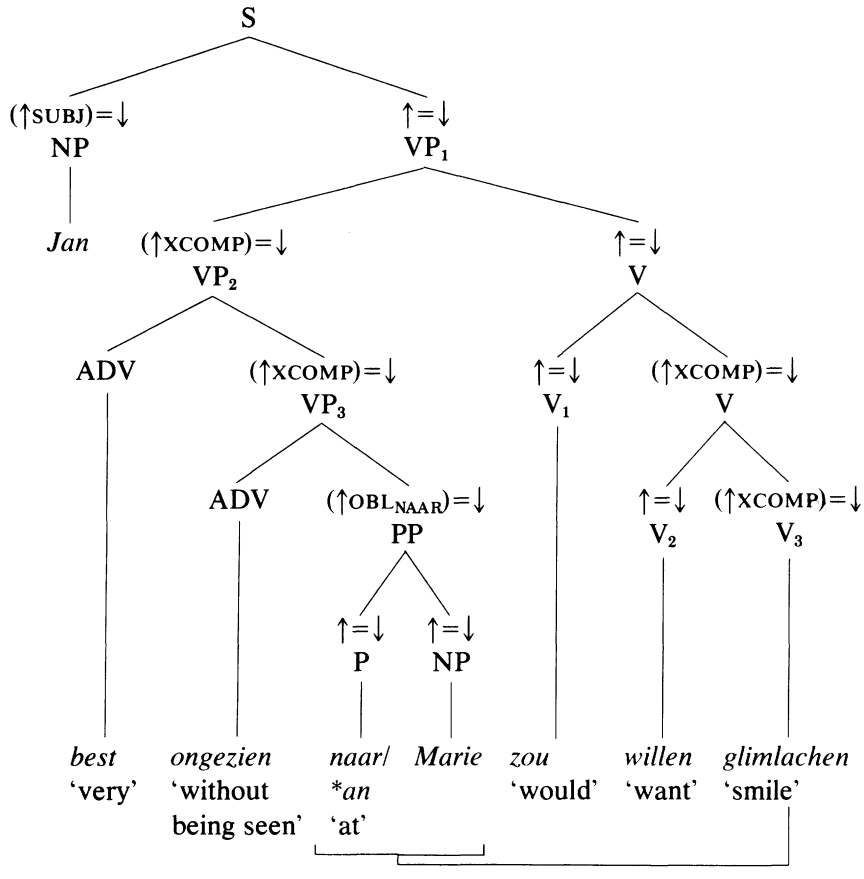
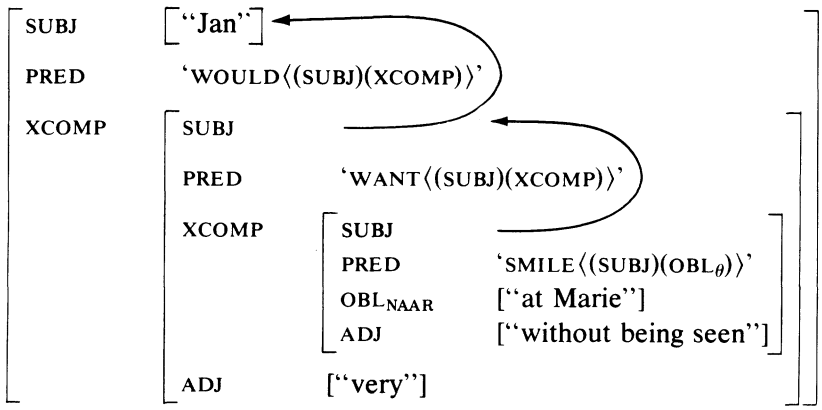


Figure 8
Verb governing oblique object in Dutch

b. *c-structure 2*



c. *f-structure*



govern; second, that similar government relations are instantiated in configurationally dissimilar structures; and third, that the types of structural configurations which instantiate government relations in a particular language (type) are predictable from the syntactic encoding of functions in that language (type).

This theory contrasts with what may be called *configurational* theories of government, which stipulate a set of possible governing relations in terms of phrase structure configurations. For example, it is assumed in the theory of government and binding (Chomsky (1981)) that government is a structural relation between a major type 0 category and an NP which is dominated by a projection of that category and that V and INFL (the category of the tense and subject-agreement morpheme) are distinct major type 0 categories having separate syntactic projections (VP and S, respectively). Verbs govern their objects, but not their subjects, which lie outside of the VP (i.e. outside of the projections of V). However, subjects are governed by the finite inflectional morpheme in INFL (or by certain superordinate verbs and prepositions that are exceptionally permitted to govern the subjects of a subordinate clause).

While it is easy to stipulate sets of possible government relations in terms of various c-structure configurations, this does not solve the fundamental problem of government, which is to *explain* the observed structural relations between governing morphemes and the constituents they govern. Why, for example, are the case, number, and person of the subject of a sentence recorded in the affix of its verb and not, say, in the affix of a prepositional object of the verb? Why is it that verbs and complementizers govern the finiteness of their clauses, while the determiners of direct objects do not? Why do the same kinds of government relations appear in unrelated languages with radically different constituent structures? Our theory provides principled answers to these questions.

8. Control

Control refers to a relation of referential dependence between an unexpressed subject (the *controlled* element) and an expressed or unexpressed constituent (the *controller*); the referential properties of the controlled element, including possibly the property of having no reference at all, are determined by those of the controller. The term *controlled clause* is also applied to the clause whose subject is controlled. To illustrate, two control relations appear in examples (17a,b).

(17) a. At the moment, the goal of the police is to try to prevent a riot.
b. At the moment, the goal is to try to prevent a riot.

In (17a), the unexpressed subject of *try* is controlled by *the police,* and the unexpressed subject of *prevent* is controlled by the (unexpressed) subject of *try*. In (17b), the unexpressed SUBJ of *try* lacks an antecedent, but this is often viewed as a degenerate control relation, called *arbitrary control.*

8.1. Theoretical Considerations

Theories differ according to what are considered the major generalizations to be derived, and what are taken as basic assumptions. For example, Chomsky's (1981) theory of control is designed to obtain the following generalizations as theorems.

(18) Only subjects are controlled.
(19) Only nonfinite clauses have controlled subjects.

To derive these results, it is assumed in Chomsky's theory that government is a structural relation, as summarized in section 7. It is further stipulated that all NPs must be governed, except for the controlled NP ('PRO'), which must not be governed. Verbs govern their objects, but not their subjects, which lie outside of the VP. However, subjects are governed by the finite inflectional morpheme in INFL (or by certain superordinate verbs and prepositions that are exceptionally permitted to govern the subjects of a subordinate clause). Therefore, the controlled NP 'PRO' can occur only in the ungoverned position of subject of a nonfinite clause. The main idea of this theory is illustrated in figure 9. Because phrase structure configurations are used to express government and control relations, this can be referred to as a *configurational* theory of control.

It is evident that these assumptions are incompatible with the lexical–functional theory of syntax. In this theory, government is a functional, not a phrase-structural, relation (section 7); verbs must govern their subjects (section 7); bound inflectional morphemes must be affixed prior to lexical insertion into phrasal structure (section 6); and null structure appears only in cases of constituent control (section 6), that is, only in cases of unbounded, structure-dependent syntactic binding relations (cf. Bresnan and Grimshaw (1978), Kaplan and Bresnan (1982)).

We have already seen that these properties of the theory adopted here lead to several explanatory results in the theories of subcategorization, inflection, and government (section 7). The same results do not follow from the assumptions of the configurational theory of control without further stipulations. Since in that theory only the inflectional affix of the verb and not the verb itself governs the subject, and since government is a structural relation that associates each governing category with a distinct phrasal domain, it follows that the phrase structure representation in terms of which government relations are stated will be distinct from the representation of the surface constituency relations of the morphemes in a sentence. Therefore, a separate component of "affix hopping" or "morphological spelling" or "realization" rules must be postulated to map between these two phrase structure representations. As in figure 9, one can stipulate that such rules must place the morpheme dominated by INFL on the morpheme dominated by V, but any other stipulation is equally possible. Thus, in the configurational theory of control, it is only an accident that the inflectional affixes that govern the subject in figures 7a–c are *verbal* affixes.

Similar considerations show that the invariance of government relations across such

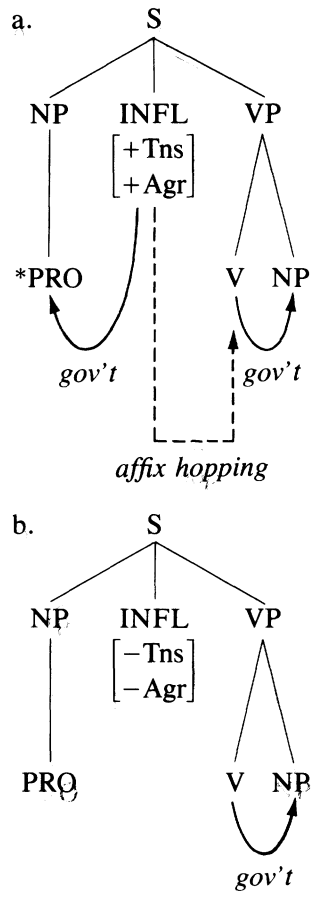


Figure 9
A configurational theory of control (Chomsky 1981)

configurationally different structures as those in figures 7 and 8 does not follow from the configurational theory without additional stipulations. Indeed, if government were a structural relation between phrasal nodes, and if such structural relations determined the distribution of 'PRO', one would predict that in languages that lack VPs, subjects could not be controlled (because in such languages verbs would always govern their subjects). However, research on Irish, a VSO language (McCloskey (1980; 1981)), and Malayalam, a verb-final language having "flat" clausal structure (Mohanan (1982a,b)), falsifies this prediction. To maintain the configurational theory, one could draw the procrustean conclusion that there is an underlying phrasal structure in terms of which government is invariantly defined for all languages and that it resembles English surface structure prior to affix hopping as in figure 9 (Chomsky (1981)). With this move, however, the claim that government is a phrase-structural relation is emptied of its most interesting and falsifiable content. The stipulation of a further level of abstract phrase structure

representation complicates the theory with the addition of yet another component of realization rules to map the English-like abstract phrase structure onto the internally motivated phrase structures of each language.

The contrasting properties of the lexical–functional theory of syntax have strong independent justification, in addition to the above explanatory results. First, careful research on the syntax of Irish (McCloskey (1980; 1981)) and Malayalam (Mohanan (1982a,b)), taking into account the various options for configurational theories of control, has yielded clear evidence that verbs govern their subjects; and in Icelandic (Andrews (1982)) and in Russian (Neidle (1982)), controlled subjects show evidence of case marking governed by the verb or by the nonfiniteness of the clause. Second, the research of Selkirk (1981), Nash (1980), Lapointe (1980), Lieber (1980), and Mohanan (1982c) supports the more constrained lexicalist theory of inflectional morphology over the transformationalist theory. Third, the constraints on null structure are supported by research into null anaphora in Malayalam (Mohanan (1981b)), Navajo (Hale (1979)), and English (Grimshaw (1979), Bresnan (1980), Levin (1982)).

It would therefore be a mistake to suppose that the configurational theory of control, in deriving generalizations (18) and (19) from other assumptions, is more "explanatory" than other theories. In fact, we have just seen that its assumptions lead to the loss of major generalizations which follow instead from the theory described here.

An important further difference between the configurational theory of control and the one presented here is the treatment of the controlled element. As a structuralist theory, the configurational theory of control cannot make direct reference to grammatical functions such as SUBJ. Instead, it takes the property of not being governed to be a universal defining property of the controlled element. Thus, to derive generalization (19), the configurational theory first stipulates that 'PRO' cannot be governed and then that the NP immediately dominated by nonfinite S (i.e. the subject position of a nonfinite clause) is an NP position that is not governed. The latter property is just as much a stipulation as the former, because there is no reason *in principle* why the type 0 category of the nonfinite S—the [−agr] or [−tns] INFL in figure 9b—should not be just as much a governor as the [+agr, +tns] INFL in figure 9a; but if it were, 'PRO' could never appear in subject position. Similarly, there is no reason *in principle* why verbs are always taken to govern their objects. The category V could bear a syntactic diacritic feature exactly as the category INFL does, which would make it an optional governor of its object; but if it were, 'PRO' could appear in nonsubject positions.

In contrast, the theory of control adopted here can make direct reference to grammatical functions. Instead of stipulating that not being governed is the defining property of the controlled element 'PRO', our theory assumes that the property of being a SUBJ is a defining property of the functionally controlled element. (The qualification "functionally controlled" is explained below.) Similarly, instead of stipulating that the only domain in which the controlled element can appear is nonfinite S, our theory assumes that only the open functions XCOMP and XADJ can denote functionally controlled clauses. We turn now to the precise expression of these assumptions within the formal system

of grammatical representation. We will then see that a number of major generalizations follow from the present theory of control relations.

Recall that control has been characterized as a relation of referential dependence between an unexpressed subject and an expressed or unexpressed controller. The characterization of control relations in terms of referential dependence captures their semantic similarity, but masks an important underlying grammatical difference. Where the referential dependence is accompanied by the complete identity of all functional features of the controller and the controlled element, we have *functional control* (also called *grammatical control* elsewhere in Bresnan (1982a)). Where the referential dependence implies no identity of grammatical features, we have *anaphoric control.* That is, functional control entails identity of f-structures of the controller and controlled element, while anaphoric control entails mere "identity of reference" (i.e. only referential dependence). The next two sections articulate the theory of functional and anaphoric control. The reader is referred to figure 14 for a schematic summary of the theory.

8.2. Functional Control

In *functional control relations,* the controlled element is the SUBJ function and the controlled clauses are designated by the open grammatical functions XCOMP and XADJ. Thus, the term *controlled clause* refers to a clause at the level of f-structure: a *clause nucleus* (section 5). The control relation is expressed by a control equation, a functional schema which equates the f-structures of the controller and the controlled element. Functional control relations are either *lexically induced* or *constructionally induced,* depending on whether the control equation is part of a lexical entry or a c-structure rule annotation. The range of possible controllers depends upon whether the functional control relation is lexically or constructionally induced.

In lexically induced functional control, the control equation is part of a lexical entry. The control relation is defined in terms of functions that are subcategorized by the lexical item that induces functional control. Thus, the controlled clause (that is, the function whose subject is controlled) is the XCOMP (the predicative or open complement). The controller is specified by a control equation of the form $(\uparrow \text{G}) = (\uparrow \text{XCOMP SUBJ})$, which is added to the lexical entry of the item. From the severe constraints on the lexical encoding of semantically restricted functions (stated in section 2), it follows that these functions cannot be lexically induced functional controllers. Hence, only SUBJ, OBJ, and OBJ2 are possible functional controllers, or values of G, in the above control equation. The unmarked choice of controller is predictable from the following universal rule:

(20) *Lexical Rule of Functional Control*

Let L be a lexical form and F_L its grammatical function assignment. If XCOMP $\in \text{F}_\text{L}$, add to the lexical entry of L:
$(\uparrow \text{OBJ2}) = (\uparrow \text{XCOMP SUBJ})$ if OBJ2 $\in \text{F}_\text{L}$;
otherwise:
$(\uparrow \text{OBJ}) = (\uparrow \text{XCOMP SUBJ})$ if OBJ $\in \text{F}_\text{L}$;
otherwise:
$(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$

That is, the XCOMP of a lexical form is functionally controlled by the OBJ2 if there is one, otherwise by the OBJ if there is one, otherwise by the SUBJ.

This rule of unmarked lexical control is interpreted as a redundancy rule; that is, the rule obligatorily expands an eligible lexical entry which lacks a control equation, but it blocks if the otherwise eligible lexical entry already has a control equation. For example, rule (20) specifies that the controllers of the predicative complements of *appear* and *regard* in (21) and (22) are the SUBJ and OBJ, respectively, but the rule is blocked by the lexically marked control equation for *strike* in (23).

(21) a. John seems sick to Mary.
b. (↑ PRED) = 'SEEM⟨(XCOMP)(OBL_θ)⟩(SUBJ)'

(22) a. John regards Mary as friendly.
b. (↑ PRED) = 'REGARD⟨(SUBJ)(OBJ)(XCOMP)⟩'

(23) a. John strikes Mary as friendly.
b. (↑ PRED) = 'STRIKE⟨(SUBJ)(OBJ)(XCOMP)⟩'
(↑ SUBJ) = (↑ XCOMP SUBJ)

An example of lexically induced functional control by OBJ2 is given in (24).

(24) Tom will serve you the fish raw.

The adjective in (24) is a *state predicate,* an optional predicative complement to a verb, which describes a state of one of the verb's arguments at the time of the action denoted by the verb. There are strong lexical restrictions on the introduction of a state predicate as complement to a verb. For example, the predicate must describe an objective state (cf. **Tom will serve you the fish tasty*), the verb's argument must be a THEME (cf. **Tom will give the fish a sauce raw*), and the verb must denote an action rather than a state (cf. **I envy Tom the fish raw*). Within these lexical restrictions, any of SUBJ, OBJ, or OBJ2 are possible functional controllers:

(25) a. The package arrived unopened.
b. John will serve the fish to you raw.
c. I sent you the letter sealed.

Our theory predicts that only SUBJ, OBJ, and OBJ2 are possible controllers in cases of lexically induced functional control; hence, the ungrammaticality of examples like (26a,b)—noted by Williams (1980)—is explained.

(26) a. I presented it to John dead.
b. *I presented John with it dead.

The verb *present* in (26) allows its THEME argument to be expressed either as OBJ or as OBL_θ; however, as predicted by our theory, only the object can control the state complement. These state complements are a special type of XCOMP; they differ from XADJS to be described directly.

In constructionally induced functional control, the control equation is part of a c-structure rule annotation. The controlled clause is the XADJ (the predicative or open

adjunct), and the controller is specified by a functional schema of the form ($\uparrow$ G) = ($\downarrow$ SUBJ), which is added to the functional annotations of the adjunct. Because the control equation is syntactically, rather than lexically, specified, it is not constrained by the restrictions on lexical encoding of functions; this means that a wider range of controllers is available in principle. The set of possible controller functions Γ appears to be a parameter of variation across languages. In Malayalam, Γ = {SUBJ} (Mohanan (1981b)), while in Russian (Neidle (1982)) and in English, Γ = {SUBJ, OBJ, OBJ2, OBL_θ}. The rule of constructional control is given in (27).

(27) *Constructional Rule of Functional Control*
If ($\uparrow$ XADJ) = $\downarrow$ is a syntactically encoded functional annotation, conjoin to it the disjunction of the schemata {($\uparrow$ G) = ($\downarrow$ SUBJ) | G $\in \Gamma$}.

Rule (27) is interpreted as a syntactic redundancy rule: if a predicative adjunct lacks a controller, rule (27) obligatorily specifies the set of possible controllers; but if an adjunct construction is marked for a particular controller, the rule's application is blocked.

To illustrate the effect of rule (27), let us consider examples of predicative adjuncts in English. (This discussion disregards other kinds of adjuncts, such as relative or appositional adjuncts.) In (28), either the SUBJ *Mary* or the OBJ *John* can be the controller of *drunk as usual,* although there may be a slight preference to interpret *Mary* as controller:

(28) Mary passed John in the hall yesterday drunk as usual.

The following example brings out the interpretation in which the OBJ *John* is the controller, because we assume that Mary would probably not describe herself as "drunk as usual":

(29) Mary said she passed John in the hall yesterday drunk as usual.

Note that in contrast to the cases of lexically controlled state complements, the controller of these predicative adjuncts is not restricted to one thematic argument of the verb. Moreover, in (30) we see that an OBL_θ can control the adjunct, for it is possible for *drunk as usual* to be predicated of the OBL_{AG} *Mary* as well as the SUBJ *John:*

(30) John was passed by Mary in the hall yesterday drunk as usual.

Again, a reportive example such as (31) brings out the interpretation in which the subject is not the controller, for once again we assume that John would probably not describe himself as "drunk as usual":

(31) John said he was passed by Mary in the hall yesterday drunk as usual.

Furthermore, these adjuncts modify a much wider range of lexical predicates than the complements do and can describe nonobjective states. For example, in contrast to **John will serve you the fish tasty,* we find *Mary caught a glimpse of the fish lying on the table, tasty and fragrant with herbs* and even *John will serve you the fish, tasty and*

fragrant with herbs. The distinction between these XADJs and the state predicate XCOMPS previously discussed corresponds to Halliday's (1967) distinction between "conditional attributes" and "depictive attributes". Halliday observes that multiple conditional attributes can occur in a simple sentence, while only one depictive attribute may. This is exactly what the analysis in terms of XCOMPS and XADJS predicts (Bresnan (1980), Kaplan and Bresnan (1982)).

While functionally controlled XADJS generally have a range of possible controllers, including OBL_θ as we have just seen, there is one construction in English in which functional control of the XADJ is restricted. This is the clause-initial position of the adjectival phrase adjuncts shown in (32) and (33).

(32) a. Sure of winning, Mary entered the competition yesterday.
b. #Sure of winning, the competition was entered by Mary yesterday.
(33) a. #Sure of winning, the competition excited Mary yesterday.
b. Sure of winning, Mary was excited by the competition yesterday.

In all of (32a,b) and (33a,b), *sure of winning* is controlled by the SUBJ of the sentence; where the SUBJ denotes an inanimate entity, the result is semantically anomalous (signaled by #). Thus, we will assume that SUBJ-control is a marked property of the clause-initial XADJ construction shown in (34). ((34) is one instance of a more general rule introducing S-initial adjuncts of various categories.)

(34) $\text{S} \rightarrow \left(\begin{Bmatrix} (\uparrow \text{XADJ}) = \downarrow \\ (\uparrow \text{SUBJ}) = (\downarrow \text{SUBJ}) \end{Bmatrix} \right) \quad \begin{matrix} \text{AP} \\ (\uparrow \text{SUBJ}) = \downarrow \end{matrix} \quad \begin{matrix} \text{VP} \\ \uparrow = \downarrow \end{matrix}$

Since the functional controller of the adjunct has been marked in the clause-initial construction given in (34), the rule of constructionally induced functional control (27) is blocked from applying there.

Because functional control relations have been formalized by equations identifying the controlled SUBJ with the controller, the f-structures of controller and controlled will be "merged" into one and the same f-structure; as a result, the controlled SUBJ will have all of the grammatical features (CASE, PERSON, NUMBER, etc.) of the controller. The referential dependence of the controlled SUBJ upon the controller follows from the semantic interpretation of the functional control relation by "quantifying in" (Halvorsen (1981)).

8.3. Anaphoric Control

Anaphoric control relations arise from the presence of a functional anaphor ('PRO') which is not expressed in c-structure. The functional anaphor is created by an optional functional schema of the form $(\uparrow \text{G PRED}) = \text{'PRO'}$ for any function G. The possible occurrences of these schemata are quite limited by independent principles. 'PRO' is a semantic form, and all semantic forms are introduced in the lexicon (Halvorsen (1981)). Hence, the optional schema $((\uparrow \text{G PRED}) = \text{'PRO'})$ must belong to some lexical entry. That cannot

be the entry for a null NP, for our theory does not permit the use of null c-structure to represent local grammatical relations (section 6). Therefore, since ((↑ G PRED) = 'PRO') lacks its own lexical entry as a null category, it must be introduced as part of the lexical entry of a lexical form that governs G (section 1) and G must be a subcategorizable function. Moreover, for any lexical entry L to which an equation ((↑ G PRED) = 'PRO') is added, the completeness and coherence conditions restrict G to the particular functions subcategorized by the lexical form of L. The constraints on lexical encoding of functions (section 2) further restrict G to be one of the set of semantically unrestricted functions {SUBJ, OBJ, OBJ2}. Finally, a language-particular parameter may restrict G to a subset Δ of the set of semantically unrestricted functions. We will further assume that only [±FIN] lexical items permit the functional anaphor, and that the value of the feature [αFIN] is a language-particular parameter. These properties are incorporated in rule (35).

(35) *Rule of Functional Anaphora*

For all lexical entries L, for all G ∈ Δ, assign the optional pair of equations {((↑ G PRED) = 'PRO'), (↑ FIN) $=_c$ α} to L.

By fixing the parameters α = − and Δ = {SUBJ}, we derive the rule of functional anaphora for English. Thus, the functional anaphor 'PRO' arises in English only as the subject of a nonfinite verb (infinitive or gerund).

An *anaphor* is a grammatical element which may be assigned an antecedent by the rules of sentence grammar. Semantically, anaphors are either coreferential with or referentially dependent upon their antecedents. 'PRO' is an anaphor similar (but not identical) in its interpretation to the definite pronouns *her, he, they,* etc. The so-called control relation between 'PRO' and its controller is actually an anaphoric relation between 'PRO' and its antecedent; we designate this relation *anaphoric control* to distinguish it from functional control. Note that the term *anaphor* as used here applies both to those pronouns that are obligatorily assigned antecedents within the sentence, such as the reflexive pronoun *herself,* and to those pronouns that are only optionally assigned antecedents within the sentence, such as the definite pronoun *her.* The former can be distinguished as *bound anaphors.*

In contrast, in the government and binding theory, *anaphors* are distinguished from *pronominals* in that anaphors "have no capacity for 'inherent reference'" (Chomsky (1981, 188)). The semantic content of this notion of "inherent reference" is unclear, but what appears to be meant by it is that anaphors lack the independent capacity to refer to specific extrasentential referents. For example, in *We thought Mary should have the operation. *However, herself is the only one that can make that decision,* the anaphor *herself* cannot refer by itself to the extrasentential referent Mary; but in *We thought Mary should have the operation. However, she herself is the only one that can make that decision,* the pronominal *she* can refer to Mary. Anaphors and pronominals are distinguished by the "binding conditions" of the government and binding theory, which assert (A) that an anaphor must be bound within its minimal governing category (NP or S), and (B) that a pronominal must be free within its minimal governing category

(Chomsky (1981, 188ff)). Thus, *herself* in *Mary voted for herself* is an anaphor, while *her* in *Mary voted for her* is a pronominal. As Chomsky (1981, 191) writes, 'PRO' is considered to be both a pronominal and an anaphor:

> Consider next pronominals without a phonetic matrix, i.e., PRO. Note that PRO is like overt pronouns in that it never has an antecedent within its clause or NP. PRO also resembles anaphors in that it has no intrinsic referential content but is either assigned reference by an antecedent or is indefinite in interpretation, lacking specific reference. It is reasonable, then, to regard PRO as a pronominal anaphor. If so, it is subject to both the binding conditions (A) and (B). Then PRO is bound and free in its governing category, a contradiction if PRO has a governing category. Therefore PRO has no governing category and is therefore ungoverned. We therefore derive the principle (20) [that PRO is ungoverned], which . . . is the essential property of PRO.

Thus, the motivation for the claim that 'PRO' is an anaphor is that when it does not derive its reference from an antecedent within the sentence, it lacks definite reference, as in *It is unclear what to do*. (This is the so-called arbitrary control.) The claim that 'PRO' is an anaphor (as well as a pronominal) in turn motivates the crucial assumption that it is ungoverned, from which follow its most essential properties for the government and binding theory. In particular, it follows that 'PRO' can be the subject of a nonfinite clause (in which INFL does not govern the subject), but not an object or a subject of a finite clause (in which INFL does govern the subject).

However, the assumption that 'PRO' is an anaphor in the sense of the government and binding theory presents a serious problem: unlike anaphors, but like pronominals, 'PRO' does have the capacity to refer independently to specific extrasentential referents. In English, many examples arise in the context of indirect discourse, as (36) illustrates.

(36) a. Mary was happy and excited. To have involved herself in the group was a risky action. But it was proving that she could change her life.
b. Tom felt sheepish. Pinching those elephants was foolish. He shouldn't have done it.
c. She sighed and looked around the empty room. It was unclear what to do with herself now that Molly was gone.

In these examples, the PRO subjects of the infinitives are respectively understood as referring to Mary, Tom, and the referent of *she*. Examples of definite reference of 'PRO' also occur in other discourse contexts such as (37).

(37) Frankly, I'm worried about Mary. What has she gotten herself into? Don't get me wrong: I think it was fine to join the group. But getting herself photographed with those starving wolves was dangerous.

Here, too, the PRO subjects of the infinitive and gerund refer specifically to Mary. Like expressed pronouns, 'PRO' can also be understood as referring to a specific referent which is presented in the nonverbal context of an utterance. Consider a situation in which two observers have witnessed a young man commit suicide by leaping from a

fortieth-floor window. One observer can say to the other, "I think that killing himself was a terrible mistake", speaking specifically of the young man and meaning by this that his killing himself was a terrible mistake.

In general, the reference of reflexive pronouns is determined by their antecedents. In the above examples, the reflexive pronouns refer to specific individuals, and their antecedents are 'PRO'. Furthermore, in these examples, 'PRO' does not derive its reference from an antecedent within the sentence. Thus, we are led to the conclusion that 'PRO' does have the capacity to refer independently to specific extrasentential referents. (This argument is due to P.-K. Halvorsen.) Thompson (1973) argues convincingly that the interpretation of 'PRO' as definite or generic is predictable from temporal or aspectual and contextual properties of the sentence in which it occurs; in contexts of specific temporal reference such as (36) and (37), 'PRO' may have a definite interpretation, but in contexts of nonspecific temporal reference such as (38), it has the generic interpretation.

(38) a. To involve oneself in that kind of group is risky.
b. Pinching elephants is foolish.

The fact that 'PRO' can have definite reference in the previous examples shows that it cannot be an anaphor in the sense of the government and binding theory. However, if 'PRO' is not an anaphor in this sense, then there is no motivation for the stipulation that it must be ungoverned; 'PRO' should occur in ungoverned positions, such as object position, like other nonanaphors in English. Even if we assumed that 'PRO' is ambiguous, being either [+anaphoric, +pronominal] or [−anaphoric, +pronominal], the government and binding theory would be unable to account for the fact that 'PRO' has the same syntactic distribution in English (namely, subject of nonfinite clause) when it is an anaphor as when it is a nonanaphor. Thus, "the essential property" of 'PRO' for the government and binding theory—its being ungoverned—remains unexplained.[7]

Note that the government and binding theory cannot adopt our definition of anaphors as grammatical elements that may be assigned antecedents by the rules of sentence grammar. That definition would admit pronouns as anaphors, because pronouns can, and sometimes must, be assigned sentential antecedents (as in *Louise craned her neck*); by Chomsky's reasoning given above, it would then follow (contrary to fact) that pronouns could occur only in ungoverned positions. The alternative of defining anaphors as grammatical elements that *must* be assigned sentential antecedents would also fail, because 'PRO' need not have such an antecedent; for example, it lacks an antecedent in *It is unclear what to do* and *I think that killing himself was a terrible mistake*.

In the theory proposed here, 'PRO' is not a bound anaphor, but a pronominal element. It must be distinguished from the expressed definite pronouns, however, because it has special restrictions on its anaphoric relations (to be discussed below). Let us therefore

[7] The implications of this analysis of 'PRO' for the government and binding theory were pointed out by K. P. Mohanan (personal communication).

assume that there is some feature—call it U (for *unexpressed morphologically*)—which separates 'PRO' from other pronouns. The functional structures for 'PRO' and *she* will therefore resemble (39a,b).

(39) a.
$$\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{U} & + \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{U} & - \\ \text{GEND} & \text{FEM} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \\ \text{CASE} & \text{NOM} \end{bmatrix}$$

Given that 'PRO' is an anaphor which need not be bound, how can we explain the contrast between (40a) and (40b)?

(40) a. Mary wished to vote.
b. Mary wished for her to vote.

The government and binding theory attributes the contrast to the analysis of 'PRO' as a bound anaphor and *her* as a pronominal. Assuming that the complementizer *for* exceptionally governs *her* in (40b), the minimal governing category of the pronominal *her* is extended to include both the matrix clause and the complement clause. The minimal governing category now includes *Mary;* thus, *her* must be free in this domain and disjoint from *Mary*. In contrast, the 'PRO', as a (bound) anaphor, must be bound within the sentence or assigned the *arb* (indefinite) interpretation. Presumably, lexical properties of *wish* force the bound interpretation.

The present theory provides a natural alternative to the above analysis. Note that in these examples, the [+U] anaphor ('PRO') is bound to the SUBJ *Mary,* while the [−U] anaphor (*her*) cannot be bound to the SUBJ. This is reminiscent of the phenomenon known in many languages as *obviation,* by which certain pronouns exclude coreference with specified types of antecedents (Hale (1978)). We therefore formulate the *Obviation Principle* given in (41). This formulation should not be regarded as a maximally general statement of obviation across languages, but rather as an alternative hypothesis to the claim of the government and binding theory (Chomsky (1981)) that (40b) is an instance of disjoint reference (i.e. of a pronoun being free within its minimal governing category). In fact, (41) can be derived from a general theory of pronominal obviation, but this topic is beyond the scope of our present concerns. See Simpson and Bresnan (in preparation) and the references cited there.

(41) *Obviation Principle*

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 + or − U, respectively.

In English, an obviative clause is any clause that can be marked by the complementizer *for*.[8] Since the infinitival complement to *wish* can be marked by the complementizer *for*, it is an obviative clause, and given our analysis of pronominals, principle (41) immediately explains the contrast between (40a) and (40b).

Our hypothesis that a type of pronominal obviation is involved in examples like (40a,b) makes a strikingly different prediction from the government and binding theory of pronominals and anaphors. In our theory, obviation is a functional relation which holds between matrix subjects and complement subject pronominals. In the government and binding theory, by contrast, the property of being free within a minimal governing category is a structural relation which prohibits *all* NPs within the matrix clause—subject and nonsubject NPs alike—from binding the pronominal subject of the *for–to* complement. The government and binding theory therefore predicts no difference between matrix subject and nonsubject antecedents of a complement subject pronominal; the complement subject should be disjoint from both. The Obviation Principle in our theory predicts that disjoint reference applies only with subject and not with nonsubject antecedents in the matrix. The evidence given below crucially chooses between the two theories. (Examples (42d) and (43c) contain nonobviative clauses; they are included as minimal contrasts.)

(42) a. For her to lose after all that effort would really surprise Louise.
[*her* = *Louise*, possibly]

b. It would really surprise Louise for her to lose after all that effort.
[*her* = *Louise*, possibly]

c. Louise would really be surprised for her to lose after all that effort.
[*her* ≠ *Louise*]

d. Louise would really be surprised if she lost after all that effort.
[*she* = *Louise*, possibly]

(43) a. Louise signaled to Ted for him to follow her.
[*him* = *Ted* and *her* = *Louise*, possibly]

b. Louise signaled to Ted for her to follow him.
[*her* ≠ *Louise* and *him* = *Ted*, possibly]

c. Louise signaled to Ted that she would follow him.
[*she* = *Louise* and *him* = *Ted*, possibly]

[8] We conjecture that the subjunctive complementizer in Romance languages is also an obviative clause marker. In other words, we propose that examples like *Marie veut qu'elle parte*, wherein *elle* must be disjoint from *Marie*, are instances of subject–subject obviation parallel to *Mary would like for her to leave* in English. Since the subjunctive clause of French is finite, the [+U] 'PRO' does not arise in it, but the [−U] pronouns are nevertheless subject to the principle. Simpson and Bresnan (in preparation) show that complementizers can be obviative markers in Warlpiri.

It is possible that the Obviation Principle (41) should be reformulated to apply only to [−U] pronouns, for the obligatory assignment of an antecedent to the [+U] 'PRO' in examples like (40a) could be attributed to lexical properties of the matrix verb, which could induce "thematic binding" of [+U] 'PRO' (Kisala (in preparation)). Until further research decides this question, the present formulation of the Obviation Principle will be tentatively maintained.

Examples (42) and (43) show that, contrary to the predictions of the government and binding theory, the complement subject pronoun is disjoint only from the matrix subject, and not from nonsubject arguments of the matrix. (It might be supposed that the extraposed clause in (42b) lies outside the minimal governing category of *Louise,* as would be the case if the extraposed clause were dominated by $\bar{S}$ rather than by VP. However, there is evidence that the extraposed clause is dominated by VP. Note that the extraposed clause in *It would really surprise her for Louise to lose after all that effort* must be c-commanded[9] by the direct object in order to account for the noncoreference of *her* and *Louise* (assuming Reinhart's (1976) formulation of the noncoreference condition); but this example is exactly parallel in structure to (42b). Hence, the extraposed clause in (42b) does lie within the same minimal governing category as the matrix object.)

The present theory also predicts the contrast in the following examples.

(44) a. To be able to shave oneself at five years of age would really surprise one's father.
b. It would really surprise one's father to be able to shave oneself at five years of age.
c. One's father would really be surprised to be able to shave *oneself/himself at five years of age.

In examples (44a,b), the matrix subject is not a possible antecedent for the complement 'PRO', so binding is not obligatory by the Obviation Principle. By contrast, the matrix subject in (44c) *is* a possible antecedent, and the binding of 'PRO' is obligatory, as is shown by the fact that the reflexive pronoun must agree with *one's father.*[10]

It must be noted that the Obviation Principle (41) does not provide the only conditions under which anaphoric control of 'PRO' by a matrix argument may be obligatory. 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 (45), which indicate that 'PRO' is obligatorily bound to the GOAL argument of *signal.*

(45) a. Louise signaled to Ted to shave himself.
b. *Louise signaled to Ted to shave oneself.
c. *Louise signaled to Ted to shave herself.

The Obviation Principle is inapplicable to (45) even though *Louise* is a subject, because *Louise* is thematically excluded as a possible antecedent A of the complement 'PRO'.

'PRO' is further distinguished from expressed definite pronouns by a universal condition on anaphoric control which is due to K. P. Mohanan (1981b). We formulate this condition as (46).

[9] A c-structure node N_1 c-commands a c-structure node N_2 if N_1 does not dominate N_2 and every node that dominates N_1 dominates N_2.

[10] If, however, we adopt the alternative formulation of the Obviation Principle suggested in footnote 8, the contrast between (42a,b) and (42c) would have to be attributed to lexical differences between the verb *surprise* and the deverbal adjective *surprised.*

(46) *Universal Condition on Anaphoric Control*

If A is a grammatically assigned antecedent of P, where the value of P is [PRED 'PRO', U +], then A must f-command P.

F-command is a relation on f-structures defined as follows:

(47) *F-command*

For any occurrences of the 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.

To illustrate the effect of the Universal Condition on Anaphoric Control, figure 10 shows the f-structure for the sentence *People who know John often discuss working too hard.* We see in this figure that the SUBJ α_1 f-commands the SUBJ β, while OBJ α_2 does not f-command SUBJ β. Hence, it follows from the Universal Condition on Anaphoric Control that $\langle \alpha_1, \beta \rangle$ but not $\langle \alpha_2, \beta \rangle$ is a possible control relation. This explains why the PRO subject of *working* in (48a) can be interpreted as *people* but not as *John;* (48b) shows that the universal condition applies only to the unexpressed 'PRO' and not to definite personal pronouns.

(48) a. People who know John often discuss working too hard.
b. People who know John often discuss his working too hard.

Similarly, the Universal Condition on Anaphoric Control explains why, in (49a), *Mr. Jones* does not anaphorically control the PRO subject of *contradicting,* even though *Mr. Jones* can be a grammatically assigned antecedent of the pronoun *his* in (49b).

(49) a. *Contradicting himself will demonstrate that Mr. Jones is a liar.
b. His contradicting himself will demonstrate that Mr. Jones is a liar.

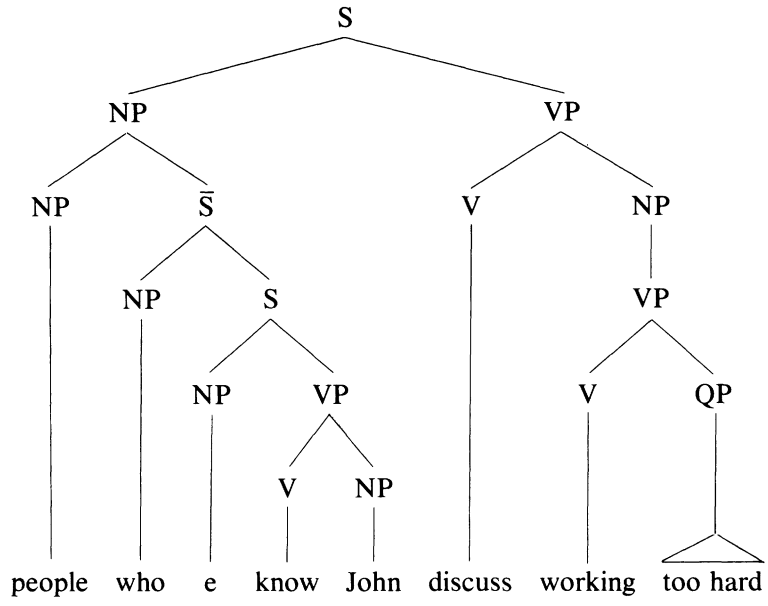
Example (49a) is ill-formed when the PRO subject of *contradicting* is taken to be *Mr. Jones*. The reason is that in the f-structure for (49a), shown in figure 11, the SUBJ *Mr. Jones* does not f-command 'PRO'. In contrast, *Mr. Jones* can anaphorically control 'PRO' in (50).

(50) Contradicting himself will discredit Mr. Jones.

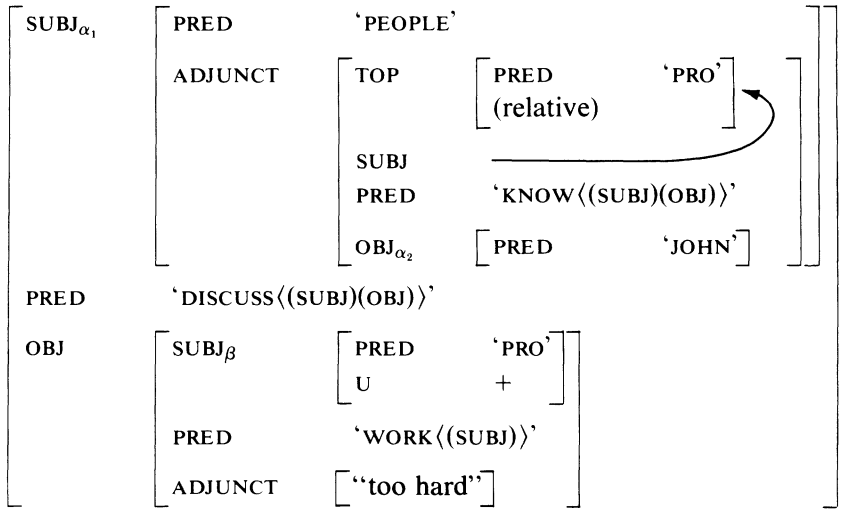
In (50), the object of the verb *discredit* does f-command the 'PRO' complement subject; see figure 12.

These examples also show very clearly that f-command, and not c-command, is the determining factor in the universal condition (46) on anaphoric control relations, for in neither of the examples *Contradicting himself will discredit Mr. Jones* and **Contradicting himself will demonstrate that Mr. Jones is a liar* does *Mr. Jones* c-command the complement subject position (see figures 11 and 12). Similarly, in *Improving himself seems important to John,* the OBL_{GO} *John* f-commands, but fails to c-command, the SUBJ of *improving himself;* see figure 13.

c-structure



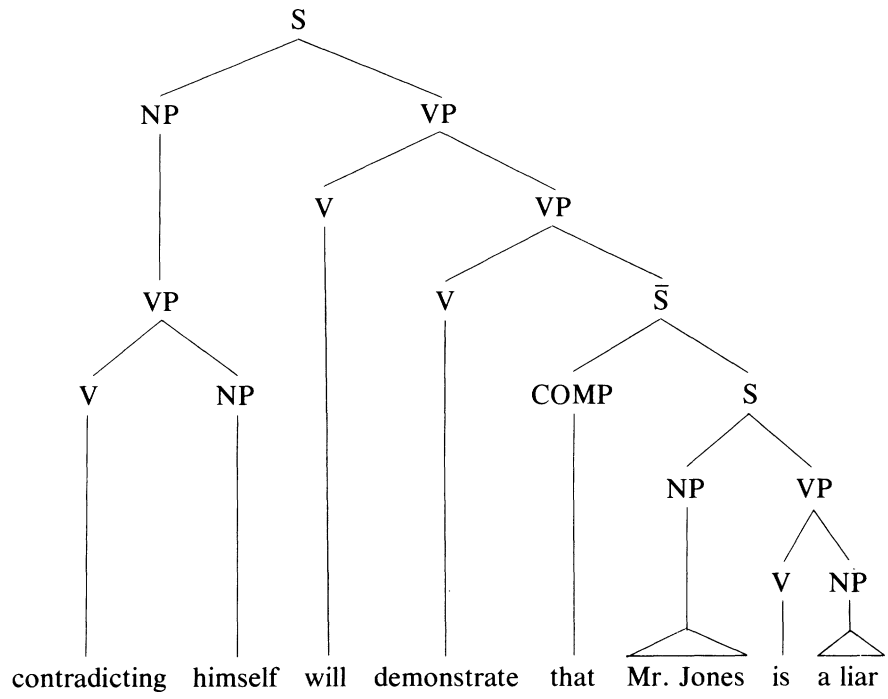
f-structure



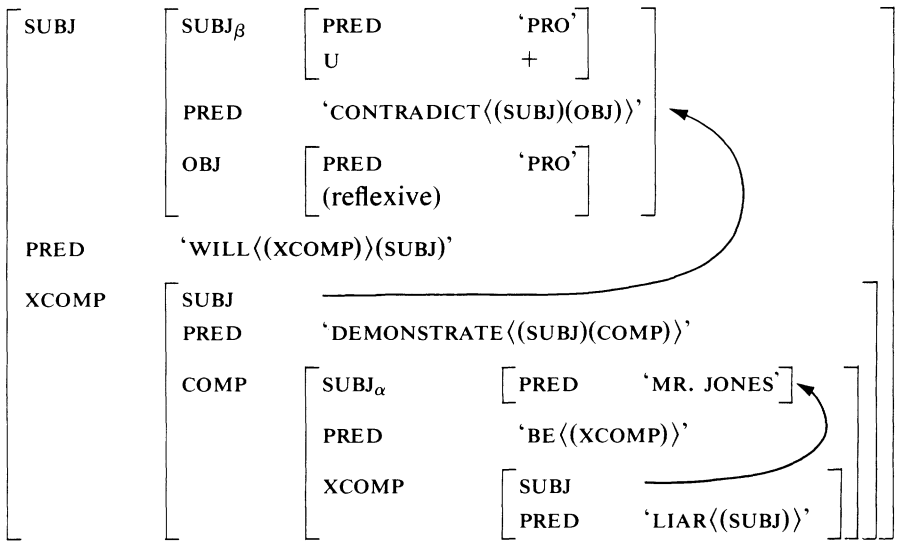
SUBJ$_{\alpha_1}$ f-commands SUBJ$_\beta$
OBJ$_{\alpha_2}$ does not f-command SUBJ$_\beta$

Figure 10
The F-Command Condition on anaphoric control

c-structure



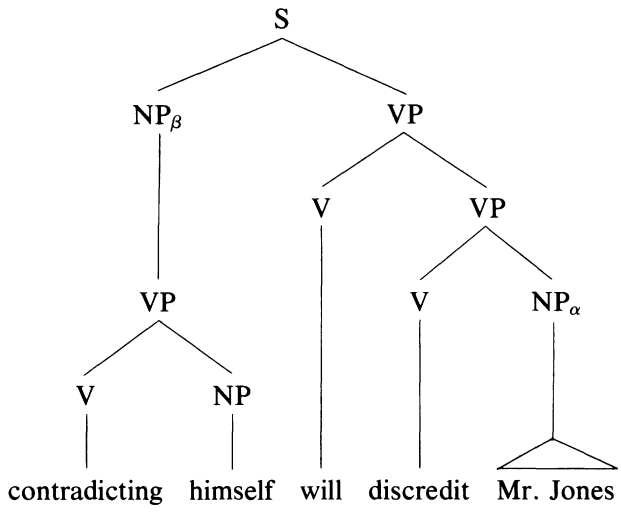
f-structure



SUBJ$_\alpha$ does not f-command SUBJ$_\beta$

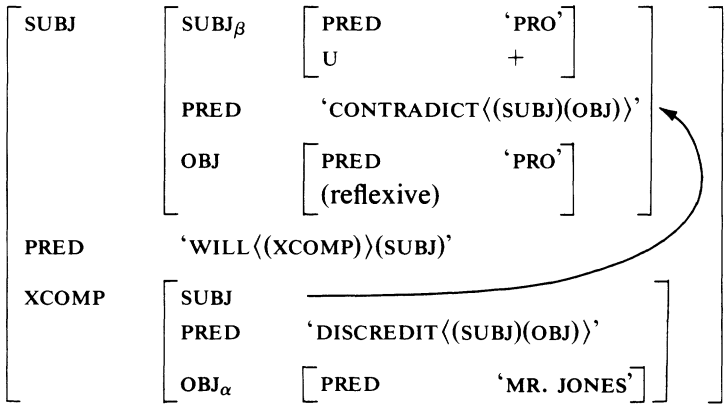
Figure 11
F-Command Condition on anaphoric control

c-structure



NP$_\alpha$ does not c-command NP$_\beta$

f-structure

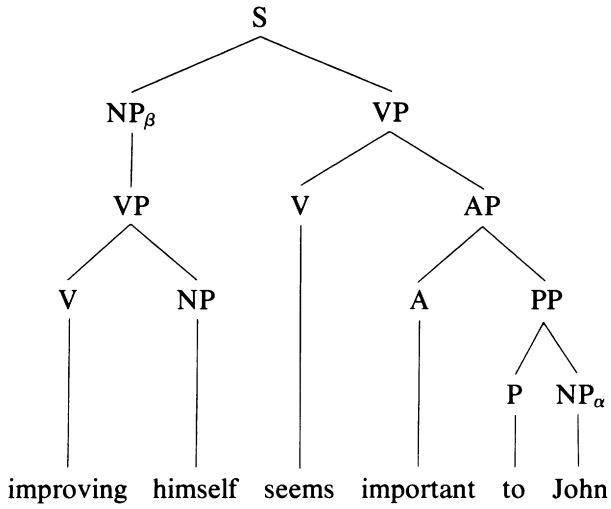


OBJ$_\alpha$ f-commands SUBJ$_\beta$

Figure 12
F-command vs. c-command

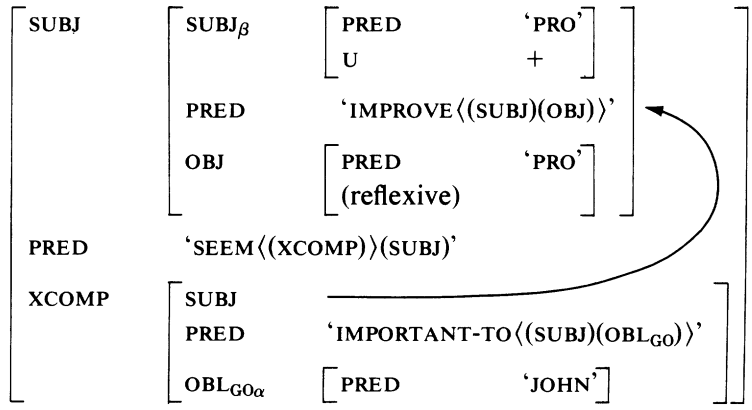
Because 'PRO' can refer to discourse antecedents, as we have seen, the possibility arises that both 'PRO' and another referential phrase in the sentence may refer to the same antecedent in discourse. When this happens, 'PRO' is not anaphorically controlled by the other phrase (because the other phrase has not been grammatically assigned as the antecedent to 'PRO'). In just these cases, apparent counterexamples to the F-Command Condition may appear. As an initial example of this situation, consider (51).

c-structure



NP_α does not c-command NP_β

f-structure



$OBL_{GO\alpha}$ f-commands $SUBJ_\beta$

Figure 13
F-command vs. c-command

(51) Contradicting himself will demonstrate that he is a liar.

Example (51) is parallel in structure to (49a), yet *he* and 'PRO' can corefer, where *Mr. Jones* and 'PRO' could not. The reason is that *he,* more easily than *Mr. Jones,* can be taken as referring to an implied discourse antecedent; *he* could even have 'PRO' as its antecedent. However, there are certain conditions under which a nonpronominal noun

phrase like *Mr. Jones* could be coreferential with 'PRO'. In general, unstressed definite noun phrases can refer to discourse antecedents much more easily than can stressed noun phrases; unstressed phrases are usually taken as designating old information, and stressed phrases, new information. The contrast between (52a) and (52b) illustrates this phenomenon.

(52) a. John was unhappy. ??The Nobel Prize had prevented the discovery of a vaccine by John.
b. John was unhappy. The Nobel Prize had prevented John's discovery of a vaccine.

In these examples, we see that the possessive *John's* in (52b) can refer to a discourse antecedent much more easily than can the prepositional object *by John* in (52a). The prepositional object, unlike the genitive prenominal NP, is in a stressed position. On the basis of this difference, we therefore expect the contrast in (53).

(53) a. Winning the Nobel Prize had prevented John's discovery of a vaccine.
b. ??Winning the Nobel Prize had prevented the discovery of a vaccine by John.

In (53a), but much less so in (53b), *John* and the PRO subject of *winning* may be coreferential with an implied discourse antecedent. Hence, in (53b) 'PRO' can easily refer to *John* only if 'PRO' is assigned *John* as an antecedent within the sentence; but this is ruled out by the Universal Condition on Anaphoric Control. *John* is properly contained within the object of the main clause, and so clearly fails to f-command the complement subject. Therefore, in (53b) it is very difficult to interpret the PRO subject of *winning* as *John*.

We see, then, that the analysis of 'PRO' as an unexpressed pronoun which is subject to the Obviation Principle (41) and the Universal Condition on Anaphoric Control (46) better expresses the generalizations that govern its behavior than the theory that it is an ungoverned pronominal anaphor. It is of interest that nothing in our theory of anaphoric control predicts that 'PRO' cannot occur in so-called governed positions, such as subject of a finite sentence or object or second object of a verb. In this respect, the theory presented here differs crucially from Chomsky's (1981) government and binding theory, which was designed to derive that generalization as a theorem. Our theory predicts that 'PRO' may have any of the semantically unrestricted functions SUBJ, OBJ, OBJ2. It would therefore be of crucial theoretical interest to know whether there are languages in which this occurs. Let us now suppose that such a language is found. Note that a possible response for the government and binding theory would be that the elements in question are not true instances of 'PRO'. Of course, in bringing evidence to bear on the choice between alternative theories of control of 'PRO', one must have clear criteria for identifying an instance of 'PRO' other than its conformity to one of the theories in question. What, then, are the independent criteria for identifying 'PRO'? Taking into account the above discussion, we can propose the following properties as collectively criterial for 'PRO':

(a) 'PRO' is an unexpressed pronoun.
(b) 'PRO' may have either definite or indefinite reference, depending upon the context.
(c) 'PRO' is not assigned an antecedent within its minimal clause.
(d) 'PRO' may be obligatorily bound to certain thematically or grammatically specified antecedents.
(e) An antecedent grammatically assigned to 'PRO' must be superordinate to 'PRO' in the clause structure representing grammatical relations (in our theory, the antecedent must f-command 'PRO').
(f) 'PRO' has a restricted set of grammatical functions (either {SUBJ} alone as in the government and binding theory, or a subset of {SUBJ, OBJ, OBJ2} as in the lexical–functional theory).

It might be questioned whether (e) should be admitted as a theory-independent property of 'PRO', since there is no alternative to the F-Command Condition. (Although it has been suggested in passing by Chomsky (1981) that the controller must be an argument of the matrix of the controlled complement, this requirement cannot replace the F-Command Condition (cf. footnote 16 and the discussion of examples (141a–c)).) In fact, a condition similar to the F-Command Condition was embodied in the earliest generative analyses of PRO phenomena, which hypothesized a cyclic Equi NP Deletion transformation. The condition of strict cyclicity limited the application of cyclic rules so that no rule could analyze material solely contained within previous cycles, a constraint that had essentially the same effect as (e).

Malayalam provides crucial evidence regarding the existence of governed 'PRO's as defined by the above criteria (Mohanan (1981b)). First, Malayalam has unexpressed pronouns. Second, these pronouns may have either generic or definite reference, depending on context. Third, they cannot have antecedents within their minimal clauses. Fourth, they may be obligatorily bound to certain functionally or thematically specified antecedents. An example given by Mohanan (personal communication) is (54).

(54) ellaawarkkum [r̄aawile kuḷik'k'un̲n̲aṯ̲ə] iṣṭamaaṇə
all-dat morning bathe-pres-it liking-be-pres
'Everyone likes bathing in the morning.'

In example (54), the PRO subject of 'bathing' is obligatorily bound to the matrix subject 'everyone'; in contrast, the expressed pronoun subject in (55) is not obligatorily bound to 'everyone'.

(55) ellaawarkkum [awar r̄aawile kuḷik'k'un̲n̲aṯ̲ə] iṣṭamaaṇə
all-dat they-nom morning bathe-pres-it liking-be-pres
'Everyone likes their bathing in the morning.'

Fifth, the Malayalam unexpressed pronouns, unlike the expressed pronouns, must be f-commanded by their grammatically assigned antecedents within the sentence. Example (56), from Mohanan (1981b), illustrates this point.

(56) [[jooṇ meeřiye ummawecca] kaařyam] awal/ϕ ařootum paraññilla
John-nom Mary-acc kiss-past thing she-nom/PRO anyone say-past
'She$_i$/*PRO$_i$ did not tell anyone that John had kissed Mary$_i$.'

Sixth, and crucially, these unexpressed pronouns can be SUBJ, OBJ, and OBJ2, but not OBL_θ or ADJ. Examples of a PRO OBJ having definite and indefinite interpretations are given in (57) and (58) (from K. P. Mohanan (personal communication)).

(57) wakkiilanmaar caṯiccu
lawyer-nom-pl cheat-past
'The lawyers cheated [definite pronoun: him, her, them, me, etc.].'

(58) wakkiilanmaar caṯik'k'um
lawyer-nom-pl cheat-fut
'Lawyers will cheat one.'

An example illustrating the F-Command Condition on an OBJ 'PRO' is given in (59) (from K. P. Mohanan (personal communication)).

(59) [jooṇ ṉuḷḷiyappool] meeriyuṭe acchan kařaññu
[John-nom pinch-past then] Mary-poss father cry-past
'Mary$_i$'s father$_j$ cried when John pinched him$_j$/*her$_i$.'

For further discussion of the F-Command Condition and evidence that the C-Command Condition does not hold, see Mohanan (1981b). Finally, an example illustrating that 'PRO' cannot be OBL_θ in Malayalam is given in (60) (from Mohanan (1981b)).

(60) řaajaawə waaḷ eṭuṯṯu. aḍḍeeham maṉṯriye waḍhiccu
king-nom sword-nom take-past he-nom minister-acc kill-past
'The king took the sword. He killed the minister (*with it).'

As indicated in the gloss of (60), it is not possible to interpret the second sentence as having a definite 'PRO' as instrumental. (Further examples are given in Mohanan (1981b).) In Malayalam, the criterial properties of 'PRO' are shared by the unexpressed subjects, objects, and second objects of finite clauses (Mohanan (1981b and personal communication)).

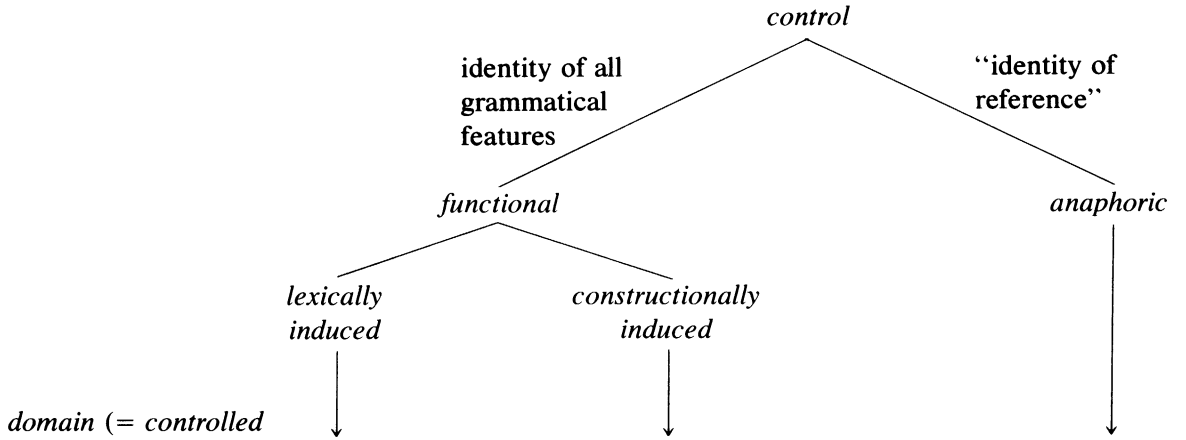
Thus, the evidence from Malayalam crucially favors the present theory of anaphoric control over the government and binding theory of 'PRO' as an ungoverned pronominal anaphor. For Malayalam, the parameters α and Δ of the universal Rule of Functional Anaphora (35) are fixed as $\alpha = +$ and $\Delta = \{\text{SUBJ, OBJ, OBJ2}\}$.

Our theory of functional and anaphoric control is summarized in figure 14.

9. Consequences of the Theory of Control

9.1. Interpretation, Function, and Agreement

From this theory of functional and anaphoric control, there follow a number of interesting generalizations. First, from the conditions on functional control (section 8.2), it follows immediately that anaphoric and arbitrary control of subject arise only in closed functions



domain (= controlled clause):	XCOMP	XADJ	closed functions
controller:	SUBJ, OBJ, OBJ2	$\Gamma^* \subseteq$ closed functions	f-commanding functions
controlled element:	SUBJ	SUBJ	$\Delta^* \subseteq$ {SUBJ, OBJ, OBJ2}
representation:	*lexical equations:* (↑G)=(↑XCOMP SUBJ) for G a controller	*c-structure rule annotations:* (↑G)=(↓SUBJ) for G a controller	*lexical equations:* (↑G PRED)='PRO' (↑G U)=+ for G a controlled element
rules:	Lexical Rule of Functional Control (20)	Constructional Rule of Functional Control (27)	Rule of Functional Anaphora (35), Obviation Principle (41)

*language-particular parameter

Figure 14
The theory of control

(and, depending on the finiteness parameter, in matrix clauses). Thus, the indefinite or antecedentless interpretation of the controlled subject arises only in SUBJ, OBJ, OBL_θ, or COMP clauses (or in main clauses, as in Malayalam); it never occurs in the predicative complement (XCOMP) or open adjunct (XADJ).

An alternative hypothesis concerning the *arb* (arbitrary) interpretation of 'PRO' is Chomsky's proposal that in the structure . . . V . . . [$_{\bar{S}}$ COMP PRO . . .] . . ., where V strictly governs $\bar{S}$, (i) if COMP (lementizer) is not null and V has no controller, then PRO is assigned the index *arb;* otherwise (ii) PRO is assigned the index *i* of a controller (Chomsky (1980b)). In other words, arbitrary control arises only in complement clauses with nonnull complementizers, such as *It is unclear what to do, It is difficult to leave.* (It is assumed in this theory that a nonnull complementizer *for* occurs in the latter example; this complementizer, through its exceptional government of the subject, is what is supposed to sanction the presence of a lexical NP subject of the infinitive in examples like *It is difficult for us to leave* (cf. section 7)).

The fact that arbitrary control does arise in complement clauses with nonnull complementizers follows as a special case of our theory, for by the principles of syntactic encoding (section 4) $\bar{S}$-complement clauses can only be assigned closed functions such as SUBJ, OBJ, OBL_θ, or COMP. In contrast to Chomsky's proposal, however, our theory predicts that arbitrary control can arise in clauses that lack nonnull complementizers, so long as these clauses have closed functions. There is evidence from Italian that confirms our theory. Manzini (1980) has observed that arbitrary control occurs in infinitival clauses in Italian, despite the fact that these infinitival clauses lack nonnull complementizers. An example is given in (61).

(61) E' difficile andarsene.
'It is difficult to leave.'

In (61), the infinitive *andarsene* functions as either the SUBJ of *difficile* or an extraposed COMP, which is bound to the SUBJ. In either case, it has a closed function, and given the absence of an anaphoric controller, arbitrary control arises. The fact that these infinitival clauses never have nonnull complementizers is irrelevant to the determination of control.

Another consequence of our theory of control is that when a lexical item agrees with (or governs) grammatical features of its SUBJ, it must also agree with (or govern) functional controllers of its SUBJ; but it need not agree with *anaphoric* controllers of its SUBJ. English has only vestigial systems of case marking and agreement, and so provides little illustration of this consequence. However, research on Modern Arabic (Fassi Fehri (1981)), Icelandic (Andrews (1982; to appear)), and Russian (Neidle (1982; in preparation)) bears out the predictions of the theory. For example, Andrews and Neidle show that there are predicative adjuncts in both Russian and Icelandic that agree in case with the subject of the clause in which they occur. In controlled clauses, these adjuncts show two patterns of case marking: in one pattern, the adjunct necessarily agrees in case with the controller of the clause; in the other, the adjunct may appear in a case (nominative in Icelandic, dative in Russian) distinct from that of the controller, the case of the subject

of the controlled clause. It turns out that agreement with the controller is necessary just where there is independent syntactic evidence of functional control. No special rules of agreement are required to account for these facts; this is exactly the result that our theory predicts. Similarly, Fassi Fehri's careful syntactic analysis of Modern Arabic shows that predicative adjectives exhibit different patterns of agreement depending upon whether they are anaphorically or functionally controlled. (Note that while agreement with a functional controller is necessary, agreement with an anaphoric controller is not prohibited; however, it must be effected by a further rule.)

These results of the theory of control reveal a surprising interrelation among semantic interpretation, syntactic function, and morphological agreement. Theories which give a uniform syntactic representation to all control relations—for example, theories which treat the controlled element everywhere as a pronominal anaphor 'PRO'—leave this deep relation among interpretation, function, and agreement unexplained.

9.2. NP Subjects and Split Antecedents

Our theory implies that lexical NPs cannot appear as the subjects of functionally controllable clauses (sections 4, 8.2). Moreover, given the consistency condition, it follows that functionally controllable clauses cannot have split antecedents. (A pronoun that refers to more than one noun phrase is said to have *split antecedents;* for example, in *Tom told Mary that they should leave, Tom* and *Mary* can be split antecedents of *they*.) Functional control by split antecedents is impossible, for the f-structures of each of the functional controllers would be merged with that of the controlled subject, resulting in a clash of features. In contrast, anaphorically controllable clauses may have lexical NP subjects, and (provided that the conditions on anaphoric control permit) they may also have split antecedents, since nonreflexive anaphors in general allow this. Thus, our theory predicts the following correlation: if a controlled clause can have a lexical NP subject, then (subject to conditions on anaphoric control) it can have split antecedents.

In English, both participle phrases and adjective phrases can function as predicative adjuncts, controlled by some element of the clause:

(62) a. Angry at John, Mary left.
b. Being angry at John, Mary left.

Unlike the adjectival adjuncts (which, as we have seen in section 8.2, are functionally controlled XADJS), the participial adjuncts have optional NP subjects, as shown by the contrast between (63a) and (63b).[11]

[11] Jane Simpson has pointed out that Stump (1981) discusses absolutives which might seem to be counterexamples. However, his examples fall into several classes that appear to have special properties. One class includes examples which could be analyzed as verbal participles, such as *Her hair braided, Mary began to put on make-up*. A second class includes such examples as *The children asleep, Bill watched TV*; it is well known that *asleep* belongs to a marked class of predicates, including *awake, aglow,* which have mixed adjectival and nonadjectival properties. For example, in contrast to most adjectives, *asleep* cannot occur prenominally and does allow modification by *right* in standard dialects of English: contrast *an angry child, *an asleep child* and **%The child became right angry, The child fell right asleep*. A third class consists of true adjectives which

(63) a. *John angry, Mary left.
b. John being angry, Mary left.

Thus, if the participles in (62b) and (63b) are instances of the same construction, it cannot be the functionally controlled XADJ. Rather, it must be a closed ADJ having an internal structure similar to the one shown in (64):

(64) $S \rightarrow \left(\begin{array}{c} NP \\ (\uparrow \text{SUBJ}) = \downarrow \end{array} \right) \quad \begin{array}{c} VP \\ \uparrow = \downarrow \\ (\downarrow \text{PART}) = \text{ing} \end{array}$

Given this analysis, our theory then predicts that, unlike the open adjectival adjuncts, the closed participial adjuncts will show anaphoric and arbitrary control:

(65) a. *Hatless, it is unnecessary to bow.
b. Having taken off your hat, it is unnecessary to bow.

(Some speakers of English reject examples like (65b); it is unclear whether this is because they class the participial adjuncts (62b) with the phrasal XADJS (62a) rather than with the absolute (sentential) ADJS (63b), or because they wish to avoid the stylistic stigma of "dangling participles". Our theory further predicts that, since the PRO subject of the participial adjuncts is a nonreflexive anaphor, these adjuncts may have split antecedents (conditions on anaphoric control permitting):

(66) a. *Mary lost track of John because, angry at each other, he had gone one way and she the other.
b. Mary lost track of John because, having been angry at each other, he had gone one way and she the other.

(Again, there are speakers who react to (66) as to (65).)

Note that the converse of this generalization does not hold. It is true that if the subject position of a clause is functionally controlled, then it cannot be replaced by a lexical NP. But it is not true that if the subject position of a clause cannot be replaced by a lexical NP, the subject is functionally controlled. The reason is that other factors than functional control may prevent lexical expression of the subject of a clause. For example, the subject of an infinitival $\bar{S}$ in English can be lexically expressed if and only if the complementizer *for* is present, as illustrated in (67).

(67) a. Louise gestured/said for me to follow her.
b. *Louise gestured/said me to follow her.
c. *Louise gestured/said for to follow her.
d. Louise gestured/said to follow her.

---

modify body parts or possessions of the subject: an example is *Her face scarlet, Mary left the room.* These examples appear to be semantically restricted in ways that require further study: contrast the previous examples with ??*Her lobster scarlet, Mary left the kitchen* and ?*Her dog sick, Mary left the room.* In general, the existence of exceptions to categorial generalizations, including the generalization that adjectives and nouns lack direct objects (Maling (to appear)), supports a markedness interpretation of the category definitions.

However, since *wh*-phrases in English exclude the presence of *for,* the subjects of *wh*-infinitival clauses cannot be lexically expressed:

(68) a. Whether to grant equal rights to women is under debate.
b. *Whether men to grant equal rights to women is under debate.

Nevertheless, it is clear that the *wh*-infinitival clause of (68a) cannot be functionally controlled, because it bears the closed function SUBJ and has the indefinite interpretation of arbitrary control. Similarly, there exist lexically governed cases of obligatory anaphoric control in nominals (Kisala (in preparation)).

9.3. Functional Restrictions on Controllers

A further generalization of control follows from this theory. Recall that in lexically induced functional control relations, OBL_θ cannot be a controller. Since the oblique functions are marked by prepositions in English (Bresnan (1979)), it follows that prepositional objects cannot be lexically induced functional controllers in English. The following examples were previously given to illustrate this point.

(69) a. I presented *it* to John *dead.*
b. *I presented John with *it dead.*

There are apparent counterexamples to this generalization, such as (70a,b):

(70) a. Louise signaled to Ted to follow her.
b. Mary relies on John to dress himself.

However, these are either cases of anaphoric control or cases of Verb–Preposition Incorporation (V–P Incorporation; see Bresnan (1982b)). For example, the fact that we find *Louise signaled to Ted for him to follow her,* in which the subject is a lexically expressed NP, indicates that the complement of *signal* must be anaphorically, not functionally, controlled in (70a). As for *rely on,* it is subject to the rule of V–P Incorporation, which produces the two lexical forms shown in (71).

(71) a. $[\text{rely}]_V$, $(\uparrow \text{PRED}) = \text{`RELY-ON}\langle(\text{SUBJ})(\text{OBL}_{\text{ON}})\rangle\text{'}$
b. $[\text{rely on}]_V$, $(\uparrow \text{PRED}) = \text{`RELY-ON}\langle(\text{SUBJ})(\text{OBJ})\rangle\text{'}$

The verb with lexical form (71a) is inserted into the structure shown in figure 15a; the verb with lexical form (71b) is inserted into the structure shown in figure 15b.

In the structure shown in figure 15a, *on John* is a constituent, while in the structure shown in figure 15b, it is not. Moreover, *John* is an OBL_{ON} in figure 15a, but an OBJ in figure 15b. The former structure accounts for the possibility of *It is on John that Mary relies,* in which the constituent *on John* is clefted; the latter structure accounts for *John is relied on by Mary,* in which *John* as an OBJ has passivized (Bresnan (1982b)). The theory of control adopted here predicts that functional control of a complement to *rely on* should be possible only when the preposition is incorporated into the verb as in figure 15b. Hence, if the functionally controlled complement is present, *on John* must fail to

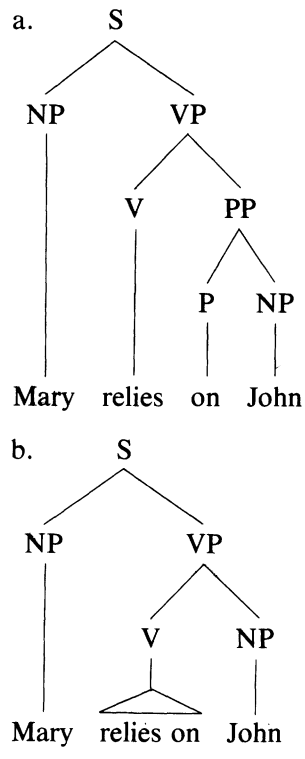


Figure 15
Verb–Preposition Incorporation

form a constituent. This explains the contrast between (72a) and (72b) and the possibility of (72c).

(72) a. *It is on John that Mary relies to dress himself.
b. It is John that Mary relies on to dress himself.
c. John is relied on by Mary to dress himself.

It is interesting to compare this explanation with an alternative proposal, due to Williams (1980). Williams proposes the *C-Command Condition,* which states that in obligatory control relations, the controller must c-command 'PRO'. Let us first observe that Williams's "obligatory control" corresponds to our functional control. That is, the central properties that Williams takes to be characteristic of obligatory control follow from our theory of functional control. First, Williams stipulates that in obligatory control there must be an antecedent. In our theory, this property follows from the completeness condition (section 5); if the functional controller is omitted from a sentence in which functional control obtains, the f-structure value of both the controller and the controlled subject will be missing, producing an incomplete f-structure. Second, Williams stipulates that in obligatory control, lexical NPs cannot appear in the position of the controlled

subject. As we have seen, this property follows from our theory of categories and functional control (sections 4, 8.2). Next, in Williams's obligatory control the antecedent must be thematically or grammatically uniquely determined. In our theory, this property follows directly from the lexical rule of Functional Control (20). In cases of constructionally induced functional control, as we have seen in examples (28)–(31), the choice of grammatical controller is not uniquely determined; it remains true, however, that any function which is chosen as a functional controller must be unique (i.e. "split antecedents" are impossible). Williams suggests as a further property that the antecedent must precede the controlled subject position, but indicates that this property may be accidental to English. This property is *not* implied by our theory. In fact, the property does not hold even for English, as examples (32) and (33) show. The final property of Williams's obligatory control is that the antecedent must c-command the controlled subject position. This property of course does not follow from our theory, since the restrictions on the controller are functional, not structural. Nevertheless, because of the way functions are encoded in English constituent structure, certain effects of the C-Command Condition already follow from our theory.

Since NP objects of prepositions will fail to c-command nodes outside of the PP, the C-Command Condition could also account for the above examples. However, it is easy to distinguish the two explanations empirically: the theory adopted here defines control in terms of relations between functions; the alternative theory defines it in terms of relations between constituent structure positions. Now it happens that in English, the OBL_θ function is syntactically realized as PP and the OBJ function as NP immediately dominated by VP, but this is not universally the case. There are languages in which oblique objects are syntactically realized as case-marked NPs and languages in which direct objects are syntactically realized as prepositional objects. In such languages, the predictions of the C-Command Condition diverge from those of the present theory of control.

In Icelandic, oblique objects can be realized as case-marked NPs which are not dominated by the node PP (Andrews (1982), Levin and Simpson (1981)). If our theory is correct, the oblique objects in Icelandic cannot be lexically induced functional controllers; but if the c-command hypothesis is correct, these oblique NPs can be obligatory controllers, since they c-command the controlled subject. The evidence given in (73) and (74), cited by Levin and Simpson (1981), confirms our theory. These are examples of *state predication,* discussed in section 8, in which a complement is predicated of a theme of the main verb.

(73) Hann rændi *matnum hráum* frá mér.
he-nom stole the-meat-dat raw-dat from me-dat
'He stole the meat from me raw.'

(74) *Hann rændi mig *matnum hráum.*
he-nom robbed me-acc the-meat-dat raw-dat
('He robbed me of the meat raw.')

Levin and Simpson argue on independent grounds that in (73) *matnum* 'meat' is a direct

object of the verb, while in (74) it is an oblique object of the verb. For example, the direct object passivizes, while the oblique object does not. Thus, it is the function of the NP, and not the c-command relation, that correctly predicts its possible control relations.

Spanish provides a case in which an NP does not c-command constituents of the verb phrase but does have the OBJ function. In Spanish, animate direct objects are realized as objects of the preposition *a*. According to the c-command hypothesis, these animate objects should not be possible obligatory controllers; according to our theory, they should be. The evidence below (provided by M. Montalbetti, personal communication) favors the functional theory.

(75) Juan la encontro a *ella borracha*.
Juan CL-acc met her drunk-fem
'Juan met her drunk.'

(76) *Juan le hablo a *ella borracha*.
Juan CL-dat spoke to her drunk-fem
('Juan spoke to her drunk.')

In (75), the NP *ella* 'her' does not c-command *borracha* 'drunk', but control is nevertheless possible. The presence of the doubled accusative clitic *la* shows that *ella* is the direct object of the verb 'meet' in (75). Example (76) illustrates that the OBL_{GO} fails to be a possible controller, as both theories would predict.

This evidence from Icelandic and Spanish shows that it is the function and not the c-structure position that determines the eligibility of an "obligatory" controller in functional control relations. Furthermore, we can see even in English that the C-Command Condition cannot be a necessary property of obligatory controllers. We have already seen examples of constructionally induced functional control, such as (77), which show an obligatory control relation between an NP and a predicate that it fails to c-command. (In this respect, AP adjuncts differ from participial adjuncts.)

(77) John said he was passed by *Mary* in the hall yesterday *drunk*.

Note that a controller of the predicative AP in (77) is required in the sentence, and that a lexical NP in place of the controlled subject is not possible in the AP construction:

(78) *It was raining outside drunk.

(79) *John was passed in the hall, Mary drunk.

Hence, the C-Command Condition is both too weak and too strong. Where objects and oblique objects happen to be syntactically encoded as NPs and PPs, respectively, the C-Command Condition will appear to hold, but the underlying restrictions on "obligatory" controllers are functional, not structural.

9.4. Visser's Generalization

Another consequence of our theory is that the controllers of lexically induced functional control relations must change under lexical operations on function assignments. Con-

sider, for example, the verb *keep* as used in *John kept laughing;* its lexical entry includes the lexical form shown in (80a).

(80) a. 'KEEP⟨(SUBJ)(XCOMP)⟩'
b. (↑ SUBJ) = (↑ XCOMP SUBJ)

The lexical rule of Functional Control obligatorily adds the unmarked control equation shown in (80b) to the lexical entry for *keep*. Now, like other intransitive verbs in English, *keep* undergoes a lexical rule of Causativization. Causativization adds a new AGENT argument, which is assigned the function SUBJ, and the SUBJ of the intransitive verb becomes the OBJ of the causativized verb by the rule (OBJ) → (SUBJ) (cf. Mohanan (1981a; 1982a)). Because Causativization replaces lexical occurrences of (SUBJ), or (↑ SUBJ), by (OBJ), the control equation changes as shown in (81).

(81) a. '$\text{KEEP}_{\text{caus}}$⟨(SUBJ)(OBJ)(XCOMP)⟩'
b. (↑ OBJ) = (↑ XCOMP SUBJ)

The result is to shift control to the object of the causativized verb, as in *Mary kept John laughing*.

The transfer of control in the examples *John kept laughing* and *Mary kept John laughing* could also be obtained if Causativization simply asserted that the THEME argument is the controller of predicative complements (cf. Anderson (1977)). However, this alternative formulation fails to account for the fact that *nonthematic* subjects can also become objects under Causativization; compare *Tabs started being kept on celebrities* and *What started tabs being kept on celebrities?*

Predictably, control is also transferred under the lexical operation of Passivization: (OBJ) → (SUBJ). For example, Passivization of the transitive active lexical entry given in (81) produces the intransitive lexical entry shown in (82). (For detailed formulation and discussion of the rule, see Bresnan (1982b).)

(82) a. '$\text{KEEP}_{\text{caus}}$⟨($\text{OBL}_{\text{AG}}$)(SUBJ)(XCOMP)⟩'
b. (↑ SUBJ) = (↑ XCOMP SUBJ)

Accordingly, we find that in *John was kept laughing by Mary,* control has transferred to the passive subject.

The theory of control predicts, however, that Passivization of a verb whose SUBJ is a (lexically induced) functional controller should be impossible, for Passivization shifts the semantically unrestricted function (SUBJ) to the semantically restricted function (OBL_{θ}) or to ϕ, and these cannot be functional controllers. This explains *Visser's generalization* (pointed out in Bresnan (1976b))—the observation that verbs whose complements are predicated of their subjects do not passivize (Visser (1963–1973, part III.2, 2118)). Consider, for example, the contrast between (83) and (84).

(83) a. His friends regard him as pompous.
b. Aunt Mary made the boys good little housekeepers.

c. Her friends had failed her in some unclear way.
d. The vision struck him blind.
e. Frank persuaded Mary to leave.

(84) a. He strikes his friends as pompous.
b. The boys made Aunt Mary good little housekeepers.
c. Max failed her as a husband.
d. The vision struck him as a beautiful revelation.
e. Mary promised Frank to leave.

The verbs in (83) have objects as functional controllers; those in (84) have subjects as functional controllers. Only the former type has corresponding well-formed passives:

(85) a. He is regarded by his friends as pompous.
b. The boys were made good little housekeepers by Aunt Mary.
c. She had been failed by her friends in some unclear way.
d. He was struck blind by the vision.
e. Mary was persuaded to leave by Frank.

(86) a. *His friends are struck (by him) as pompous.
b. *Aunt Mary was made good little housekeepers (by the boys).
c. *She was failed (by Max) as a husband.
d. *He was struck (by the vision) as a beautiful revelation.
e. *Frank was promised to leave (by Mary).

Note that the examples in (86) are ill-formed whether or not the agentive *by*-phrase is expressed.

Where the complement subject is not functionally controlled, Passivization should be possible. This accounts for the contrast between (87) and (88).

(87) a. John promised Mary to be on time.
b. *Mary was promised by John to be on time.

(88) a. John promised Mary that he would be on time.
b. Mary was promised by John that he would be on time.

The following observations explain the contrast in more detail. The lexical form for *promise* in (88) is 'PROMISE⟨(SUBJ)(OBJ)(COMP)⟩'; the COMP (closed complement) is not a functionally controllable clause. In contrast, the lexical form for *promise* in (87) has an open XCOMP, which must be grammatically controlled. However, this is a marked control relation, for even though *promise* has an object, it evinces subject control.

(89) 'PROMISE⟨(SUBJ)(OBJ)(XCOMP)⟩'
$(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$

In (89), Passivization would substitute (OBJ_{AG}) or ϕ for ($\uparrow$ SUBJ), yielding an impossible functional control relation, since only the semantically unrestricted functions can be functional controllers.

Certain apparent exceptions to the conditions on grammatical control relations are actually instances of anaphoric control. Examples are given in (90).

(90) a. Mary was never promised to be allowed to leave.
b. It was never promised to Mary to be allowed to leave.
c. To be allowed to leave was never promised to Mary.

(90a) appears to contradict Visser's generalization, in that a subject control verb (*promise*) passivizes, but (90b,c) show that *to be allowed to leave* behaves like a closed COMP, which cannot be functionally controlled. Unlike XCOMPS, closed COMPS may undergo *It* Extraposition ((90b)) and may appear as the subjects of finite clauses ((90c)). Thus, (90) is analogous to (91).

(91) a. Mary was never promised that she would be allowed to leave.
b. It was never promised to Mary that she would be allowed to leave.
c. That she would be allowed to leave was never promised to Mary.

In contrast, the infinitival complement in (92) does not behave like a COMP.

(92) a. John promised Mary to be on time.
b. *It was promised to Mary to be on time.
c. *To be on time was promised to Mary.

The reason that *to be on time* in (92) is interpretable as an XCOMP but not as a COMP may be that there is a subtle difference between the meaning of *promise . . . to* and *promise . . . that*. To promise . . . to . . . is to commit oneself to someone to act in some way. In *X promised Y to Z,* X is an agent, Y is a goal or beneficiary, and Z is an action. If the action Z is not in X's power to perform, the result is odd: cf. *#John promised Mary to be tall, #John promised Mary to be allowed to feed himself.* But to promise . . . that . . . is to commit oneself to a (possibly abstract) transfer of some benefit to someone. In *X promised Y that Z,* X is an agent and perhaps source as well, Y is a goal, and Z is a theme. This sense of *promise* is comparable to its sense in *John promised Mary an apple,* which means 'John promised Mary that she could have an apple'. With *promise . . . that,* if the theme Z is not transferable to the goal, the result is odd. This would account for the difference between (90a), *Mary was never promised to be allowed to leave,* and (87b), **Mary was promised by John to be on time; to be allowed to leave* can be interpreted as an abstractly transferable benefit ("permission to leave"), but *to be on time* cannot. In short, the semantic content of the complement biases the choice of predicates *promise-to* or *promise-that*. Thematic constraints on anaphoric control imply that a theme will be anaphorically controlled by its goal, experiencer, or possessor (Bresnan (1979)); hence, this analysis of *promise-that* subsumes the interpretation of (90) under the same rule that interprets (93).[12]

[12] Although the passives *Mary was never promised permission to leave* and *Mary was never promised to be allowed to leave* are both judged grammatical by many speakers, the actives *Fred promised Mary permission to leave* and *Fred promised Mary to be allowed to leave* differ, the latter being less acceptable for many. It is unclear exactly why this should be so. Perhaps the infinitival marker *to* creates a strong lexical bias toward the *promise-to* reading, which is incompatible with the content of the infinitival phrase (cf. Ford, Bresnan, and Kaplan (1982)).

(93) a. No one ever promised Mary permission to leave.
b. Mary was never promised permission to leave.
c. Permission to leave was never promised to Mary.

The unexpected deviance of some examples of the type in (87) had occasionally been noticed in the literature of transformational grammar (Chomsky (1965, 229), Jenkins (1972, 200ff)). However, its significance has only recently been appreciated, since it was an assumption in transformational grammar so long and widely maintained as to seem a truism, that however passives may differ from actives in scope relations, discourse function, or stylistic effect, the *grammatical relations* relevant to the semantic interpretation of passive sentences are identical to those of corresponding active sentences.[13] Indeed, this is the fundamental reason for postulating a Passive transformation. Visser's generalization shows that this truism is false. In the lexical analysis of passivization required by our theory, grammatical relations do vary under passivization while the basic predicate argument structure remains invariant (Bresnan (1982b), Roberts (1982), Mohanan (1982a), Halvorsen (1981)). This offers a straightforward explanation of Visser's generalization, as we have just seen.

Since the significance of Visser's generalization was first pointed out (Bresnan (1976b)), there have been several attempts to provide an account for it that is consistent with transformational theories of passivization (Wasow (1977)), Anderson (1977), Chomsky (1977b; 1980b), Williams (1980)), but none of them are satisfactory. The first attempt in transformational grammar to explain the deviance of examples like (87) appears to be that of Jenkins (1972, 200ff), who proposed a constraint stating that the object of *by* cannot be coreferential with an implicit or expressed subject of a complement. If this were true, it would itself require explanation, but the following examples show that this *by*-phrase constraint does not express the correct generalization:

(94) a. John had been promised by Mary that she would meet him at the station.
b. John expects a promise by Mary to remain faithful to him.
c. An attempt by the gang of four to advance themselves now would be foolhardy.

In all of these examples, the object of *by* is or can be understood as coreferential with the subject of the complement. In (94a) the complement subject *she* is expressed, while in (94b,c) the subjects of the infinitival complements are implicit. Moreover, the *by*-phrase constraint fails to account for the persisting deviance of examples (86a–e) and (87b) when the *by*-phrases are omitted. In contrast, our theory correctly distinguishes examples like (94a–c) from the deviant example (87b). In (94a), the *that*-complement is a (closed) COMP, which cannot be functionally controlled. In (94b,c), the nominals *promise* and *attempt* cannot have undergone Passivization, since derived nominals lack both SUBJ

[13] Recall that we use the term *grammatical relations* in a theory-neutral way to refer to the mapping between semantic predicate argument structure and syntactic constituent structure. Thus, *grammatical relations* are to be distinguished from *grammatical functions*.

and OBJ (Rappaport (1980)).[14] The anaphoric relations between the objects of *by* and the implicit subjects of the infinitives are similar to those in (95).

(95) a. John awaits a promise from Mary to remain faithful to him.
b. An attempt on the part of the gang of four to advance themselves now would be foolhardy.

Anaphoric control in nominals is analyzed in Kisala (in preparation).

Anderson (1977) also attempts to explain Visser's generalization. He formulates a thematic control rule which states that a complement is attributed to the theme of its clause (Anderson (1977, 375–376)) and implies that the passivized subject in (87b) is not thematic. However, this account provides no insight into the contrast between *Mary was promised to be allowed to leave* and **Mary was promised to be on time,* for in neither example can *Mary* be considered the theme. Moreover, Anderson's hypothesis, like Jenkins's, fails to explain why the infinitival clause can be predicated of a nontheme in (94b,c). Finally, as mentioned above in the discussion of *Tabs started being kept on celebrities,* complements *can* be predicated of nonthematic subjects and objects. As Borkin (1974) has observed, in many idiolects of English, examples like (96a) are found alongside the construction (96b).

(96) a. There struck me as being too few women in positions of power.
b. Too few women struck me as being in positions of power.

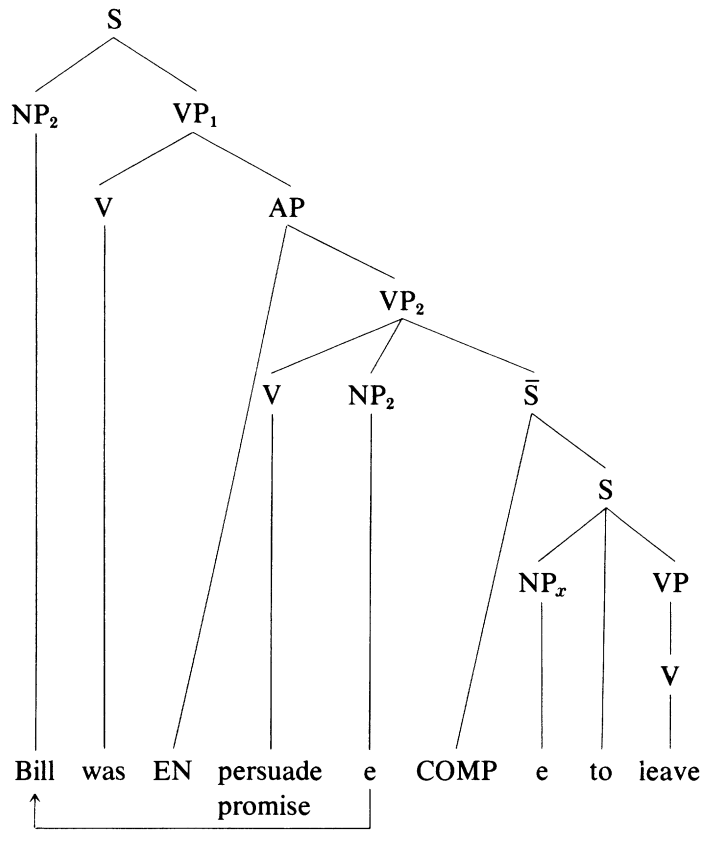
Our theory explains why these examples fail to passivize:

(97) a. *I was struck (by there) as being too few women in positions of power.
b. *I was struck (by too few women) as being in positions of power.

However, Anderson's hypothesis offers no explanation for the contrast between (96a) and (97a) and for the parallel behavior of (96a,b) and (97a,b), for in neither (96a) nor (97a) is the complement predicated of a theme.

Still another recent attempt to explain Visser's generalization in a way consistent with transformational theory is found in Chomsky (1977b; 1980b). Chomsky proposes that passive sentences have the structure shown in figure 16 (see Chomsky (1977b, 13)). As indicated in the figure, an NP Movement transformation has shifted the object *Bill* of the verbs *promise, persuade* into position as subject of *be;* NP Movement leaves a trace NP_2 which is coindexed with the moved NP_2. The NP_x in the $\bar{S}$-complement to *promise* and *persuade* is a 'PRO' which must be coindexed with a subject or object NP

[14] According to Rappaport's analysis, passivization appears to apply in derived nominals because (1) nominalized verbs inherit the predicate argument structures of their base verbs, but not the grammatical function assignments of those verbs; (2) the grammatical function assignments of nominals include only POSS(essor) and the semantically restricted functions. Thus, in *the destruction of the city by the Romans* the *of*-phrase is an OBL_{TH} and the *by*-phrase is an OBL_{AG}. Rappaport's analysis explains why both types of phrases appear with nominalizations of intransitive verbs: *the arrival of John, a loud sneeze by Bill. John* is the THEME of the nominal *arrival* and the AGENT of the nominal *sneeze*. It also explains why the nominalizations corresponding to *Mary destroyed John* and *Mary surprised John* differ in form: *Mary's destruction of John, *Mary's surprise of John, John's surprise at Mary. John* is the TH of *destroy*, but the EXP of *surprise*.



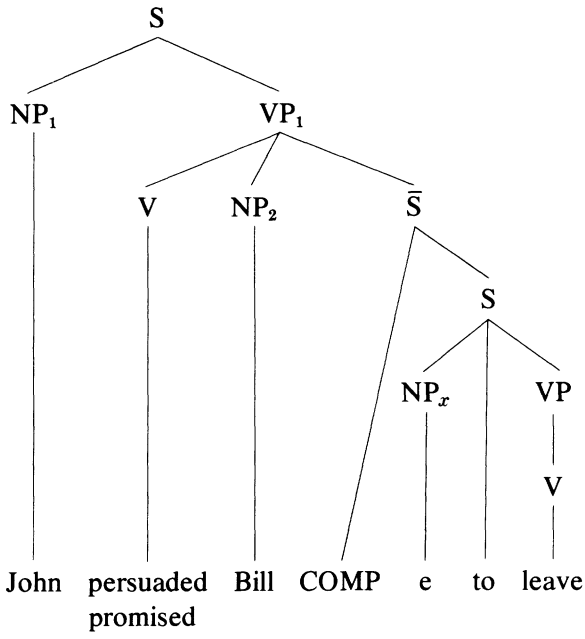
In this structure, *promise* has no subject, so control of its complement's subject is impossible; the object of *persuade* controls its complement's subject.

Figure 16
A configurational account of Visser's generalization (Chomsky 1977b)

of these verbs. Chomsky (1977b, 13–14) writes:

> The rule of control for *persuade* applies within VP_2, assigning to x of NP_x the value 2, as before, *persuade* being the main verb of VP_2 with NP_2 as its object. But the rule of control for *promise* is inapplicable since *promise* in [figure 16] is not the main verb of VP_1 with the surface subject *Bill;* compare in contrast [figure 17], where *promise* [is] the main verb of the VP sister to the matrix subject.

Since NP_x cannot be assigned subject control for *promise* in figure 16, the passive sentence *Bill was promised to leave* is ungrammatical. Chomsky concludes, "More generally, it follows that a verb with subject control (*promise, strike,* etc.) cannot appear in the passive" (p. 14). This account is technically inadequate in certain ways that are



Control of complement's subject is possible for both *promise* and *persuade* in this structure.

Figure 17
Examples of control in active structures

corrected in Chomsky (1980b). However, the basis for Chomsky's explanation of Visser's generalization remains that "there is no subject under passive, hence no way for control to be assigned [to passivized verbs of obligatory subject control]" (Chomsky (1980b)), and this account is carried over in Chomsky's subsequent government and binding theory (Chomsky (1981, 75–76)).

Thus, for Chomsky, the apparent subject of a passive verb—e.g. *Bill* in figure 16—is actually only the subject of *be;* it is related to the deep object of the passivized main verb only by the trace left by NP Movement. Hence, when a passive verb is embedded in the complement to any verb that assigns control to its complement subject, what is assigned control is the PRO subject NP_x of *be*, as figure 18 illustrates. The binding relation between the subject of *be* and the object of *elect* is accomplished by the transformation of NP Movement. The rule of control for *try* will thus assign both occurrences of x in figure 18 the value 1.

The key assumption in Chomsky's explanation for Visser's generalization is that passivized VPs by themselves have no subject. From this assumption it follows straightforwardly that such VPs could not be direct complements to verbs that assign complement subject control; this situation is schematized in figure 19. Here, neither NP_1 nor

NP_2 can control the subject of VP_2, because VP_2, by hypothesis, *has* no subject. Thus, the ungrammaticality of (98a,b) is consistent with Chomsky's assumption.

(98) a. *Bill tried elected.
b. *Mary forced Bill examined by a doctor.

The ill-formed example **Bill tried elected* contrasts with the well-formed example *Bill was elected* because *try,* unlike *be,* assigns a thematic role to its subject, precluding NP Movement into the subject position. (Formally, this would be ruled out by the Theta Criterion (Chomsky (1981)).)

However, examples like (98a,b) could be ill-formed simply because *try* and *force* require infinitival complements in their intended uses here, while *be* clearly does not. To control for this possibility, we should find verbs that resemble *try* and *force* in having thematic subjects and objects, but differ from these verbs in permitting noninfinitival phrasal complements to be predicated of their arguments. Sample verbs that would fit

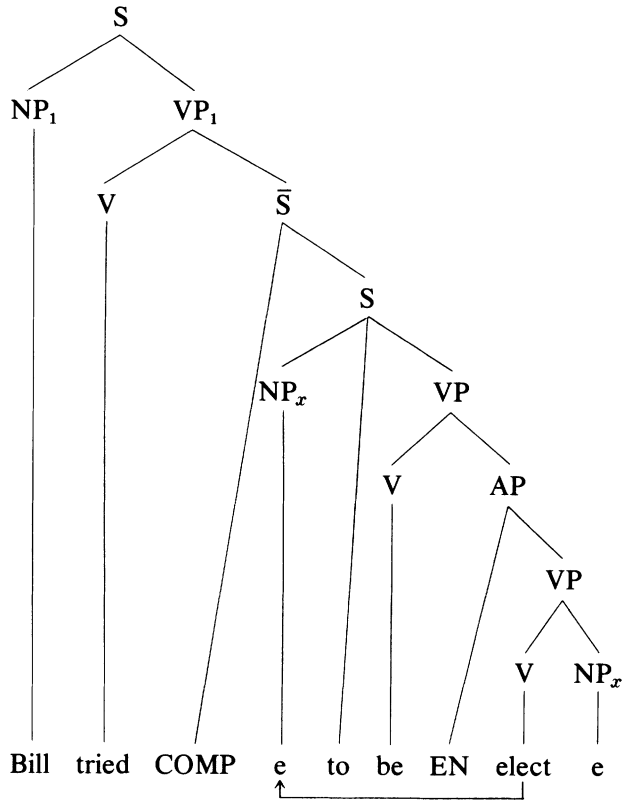


Figure 18
Passivization by NP Movement in the complement

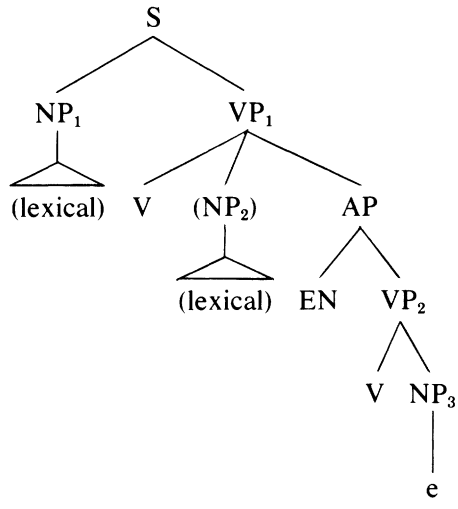


Figure 19
An impossible configuration

into the main verb position of structures like figure 19 with ordinary AP complements would be *stay, keep, feel,* and *get* in examples like the following.

(99) a. The room stayed empty.
b. She kept the room empty.
c. John felt glum.
d. That got John glum again.

These verbs have dyadic or triadic predicate argument structures, assigning thematic roles to their subjects and to their objects (in addition to their predicative complements):

(100) a. *The tables have stayed turned on us.
(= Our roles have stayed reversed.)
b. The tables have been turned on us.
(= Our roles have been reversed.)
(101) a. *The reporters will keep the cat out of the bag.
(= The reporters will keep the secret exposed.)
b. The reporters will consider the cat out of the bag.
(= The reporters will consider the secret exposed.)
(102) a. *Our goose feels cooked.
(= We feel in trouble with no way out.)
b. Our goose is cooked.
(= We are in trouble with no way out.)
(103) a. *Trying to help his friends got excessive advantage taken of John.
(= Trying to help his friends got John taken advantage of excessively.)
b. I consider excessive advantage to have been taken of John by his friends.
(= I consider John taken advantage of excessively by his friends.)

(Example (103b) is ill-formed without *to have been* because *consider* takes AP complements, and a thematic condition on adjectival passives rules out semantically empty subjects (see Bresnan (1982b) for a detailed account). In contrast, the addition of *to be* in (103a) does not improve the example.)

Of course, the selectional restrictions of the AP-complement verbs are not *identical* to those of the infinitival-complement verbs; the latter often require animate arguments where the former require only themes. Contrast *The situation stayed/felt unbearable* with *#The situation tried to be unbearable,* and *She kept/got the room warm* with *#She forced the room to be warm.* Nevertheless, in their imposition of selectional restrictions on their subjects and objects, these verbs behave like *try* and *force* rather than *seem* and *expect*:

(104) a. Our goose seems to be cooked. (idiomatic)
b. Our goose tries to be cooked. (unidiomatic)
(105) a. I expected excessive advantage to be taken of John by his friends.
b. *I forced excessive advantage to be taken of John by his friends.

If Chomsky's explanation of Visser's generalization is correct, then passivized VPs should not appear as direct complements to the verbs *stay, keep, feel,* and *get* as used above. NP Movement from the complement object position is impossible; the passivized VP itself lacks a subject, and the matrix verb assigns a thematic role to its own subject and object, precluding NP Movement into these positions. Consider then the crucial examples (106a–d).

(106) a. The wall will stay painted black.
b. She kept the wall painted black.
c. John felt betrayed and made to look like a fool by Susan.
d. That got John handed a can of worms.

Indeed, we have already seen passives embedded in some of the examples given above (e.g. *Trying to help his friends got John taken advantage of*). The passivized VPs or APs in each of these well-formed examples show that Chomsky's explanation for Visser's generalization cannot be correct.[15] (The lexical–functional theory of passivization correctly predicts the existence of such examples (see Bresnan (1982b)).)

One might think that Chomsky's proposal could be rescued by assuming that a rule of *To Be* Deletion has applied in (107a–d), deleting the verb *be* from the immediate

[15] It is easy to verify that the VP complements in these examples contain passivized verbs: only Vs and not As in English can take direct NP complements (Wasow (1977), Bresnan (1978)) and *made* in *made to look like a fool* has the verbal sense 'caused' and not the adjectival sense 'created'. In any event, however, adjectival passives pose exactly the same problem for the NP Movement theory as do verbal passives, for Bresnan (1982b) has shown that adjectival passives are formed by conversion from verbal passives and that there is no separate lexical process of adjectival passivization distinct from verbal passivization. This means that the NP Movement approach can express the relation between adjectival and verbal passives only by assuming that both are derived by the same NP Movement rule, as Fiengo (1977) proposes (in violation of the lexicalist hypothesis; see Bresnan (1982b) for a discussion of problems inherent in Fiengo's analysis). Thus, whether they contain adjectival or verbal passives, examples like *John felt betrayed by Mary* are inconsistent with Chomsky's explanation of Visser's generalization.

complement of the main verb. However, there is evidence against an underlying *be* in these examples:

(107) a. *The wall will stay *to be* painted black.
b. *She kept the wall *being* painted black.
c. *John felt *to be* betrayed and *to be* made to look like a fool by Susan.
d. *That got John *to be* handed a can of worms.

Moreover, when the phrasal complement of one of these verbs does contain *be,* as in *John felt eager to be on the list,* it may not delete: **John felt eager on the list.* Under the *To Be* Deletion hypothesis, the rule of *To Be* Deletion would have to apply obligatorily and locally to verbs like *stay, feel, keep,* and *get* and yet be inapplicable to verbs like *try* and *force.* These properties of lexical specificity and strict locality are of course precisely the characteristics of lexical subcategorization. In fact, Baker (1979) has argued from considerations of language acquisition that replacing the *To Be* Deletion analysis with the direct subcategorization hypothesis yields a more explanatory theory of grammar. His arguments indicate that rules such as *To Be* Deletion should be eliminated in principle from grammars.

Since Chomsky (1970; 1980b) recognizes the theoretical deficiencies of the *To Be* Deletion analysis, the only other recourse would seem to be to assume that the verbs *stay, keep, feel,* and *get* in (106) have sentential complements containing *be* in *logical form,* as Chomsky (1980b) has proposed for verbs like *strike.* Accordingly, (106a) would have the logical form sketched in figure 20. Note that the circled portion of this figure was not present in the syntactic structure shown in figure 19; this is structure that has been created in logical form by lexically governed structure-building rules. A rule of control for the verb *stay* could then assign control to the PRO subject of the logical form $\bar{S}$ which was not present in the transformational derivation of the example. Observe, however, that since the PRO subject NP_2 of *be* was not present in the transformational derivation of the example, it remains unbound to the PRO object of *paint.* Whereas the structure-preserving transformation of NP Movement established this crucial connection in figure 18, NP Movement cannot have applied in figure 20, because, by hypothesis, no subject NP position existed for the object NP to be moved into in the transformational derivation of the sentence. Hence, there is no explanation for the fact that in *The wall will stay painted black,* the phrase *painted black* has the passive interpretation. The same argument applies, mutatis mutandis, to each of the other examples of (106). We might think of adding a new rule of NP Coindexing to logical form to rescue this example by binding the syntactic object of *painted* in figure 20 to its logical form subject. However, this would of course make the transformational rule of NP Movement redundant; moreover, it would raise the question of why examples like **John tried examined by a doctor,* **Mary forced John examined by a doctor,* and even **Bill was promised to leave* cannot be rescued in the same way, by NP Coindexing in logical form.

Thus, the *To Be* Insertion analysis is just as deficient as the *To Be* Deletion analysis. In fact, in Chomsky's current government and binding theory, both analyses must be

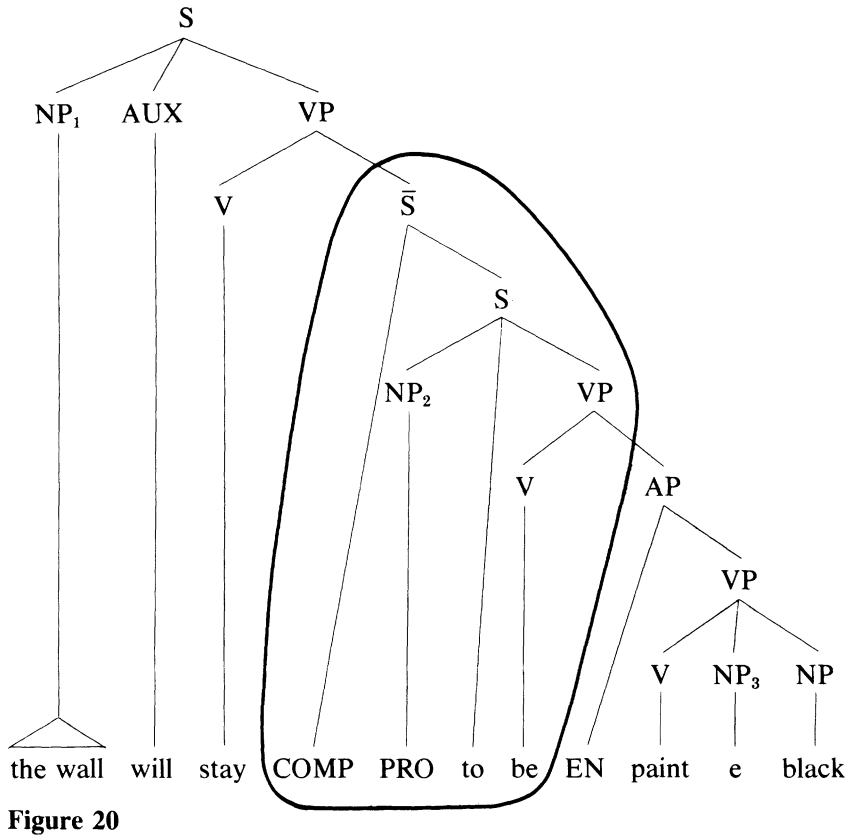


Figure 20
Structure-building in "logical form"

rejected as incompatible with the Projection Principle. However, this leaves the government and binding theory with no explanation at all for Visser's generalization.

Another recent attempt to explain Visser's generalization in a transformational theory of passivization is Wasow's (1977, 352–353). Wasow suggests that the generalization could be explained if passives were transformationally derived from the "logically subjectless" deep structures proposed in Bresnan (1972, 139–143) and in Emonds (1970). Except for the assumption that NP Movement leaves a trace, this analysis of passivization is essentially the one adopted in Chomsky (1977a; 1980b; 1981), and it suffers from the same deficiencies.

Williams (1980) offers an account of Visser's generalization based on his C-Command Condition. He observes that in examples like (108a), *Tom* fails to c-command *smart*.

(108) a. *Mary was struck by Tom as smart.
b. *Mary was struck as smart.

Williams also observes that the ill-formedness of (108a) persists when the *by*-phrase is

omitted ((108b)), and he therefore adds the stipulation that 'PRO' must have an antecedent in such examples. However, we have already seen in section 9.3 that the C-Command Condition provides neither a necessary nor a sufficient characterization of obligatory control and predication relations; furthermore, the need for an antecedent in (108b) is precisely what must be explained.

We see, then, that despite a concerted effort, no adequate explanation for Visser's generalization has yet been found within transformational theories of passivization. The fact that this generalization is among the consequences of our theory counts as very strong support. Before leaving this topic, let us consider a very different approach to the problem.

Bach (1979; 1980) has proposed a novel explanation for Visser's generalization within the framework of Montague grammar as developed by Bach, Dowty, Thomason, and others (Partee (1976)). In this framework, syntactic structures are represented by a categorial grammar, which projects combinatorial information about sentence structure onto basic lexical categories. For example, the transitive verb *hit* is assigned to the lexical category IVP/NP (something that combines with an NP to form an intransitive verb phrase), while the intransitive verb *die* is assigned to the lexical category IVP (intransitive verb phrase, that is, something with which an NP combines to form a sentence). Similarly, the verb *persuade* is assigned to the lexical category (IVP/NP)/to VP (something which takes an infinitival VP to form a phrase which takes an NP to form an intransitive VP), while the verb *promise* is assigned to the lexical category (IVP/to VP)/NP (something that takes an NP to form a phrase that takes an infinitival VP to form an intransitive VP). The motivation for this categorial difference between *persuade* and *promise* is in part semantic: unlike *persuade* and like *try, promise* is a verb whose complement is predicated of its subject. Bach comments on this classification (1979, 518): "Just as *persuade to go* is a complex representative of the category of transitive verb, so *promise Mary* is a complex representative of the same category as *try*." In Bach's terms, *persuade to go* and *hit* are transitive verb phrases and *promise Mary* and *try* are intransitive verb phrases. Bach's explanation for Visser's generalization is then that Passivization applies only to transitive verb phrases and that its effect is to turn them into intransitive verb phrases.

Like the lexical–functional framework, a categorial framework encodes a large amount of surface syntactic information into lexical entries. Because this syntactic information is encoded in the syntactic *categories* of individual lexical items in a categorial framework, however, many cross-categorial lexical generalizations cannot be expressed directly in the lexicon. For example, there is no uniform lexical characterization of the notion *transitive verb,* for *hit* and *persuade* have very different lexical categories (owing to their different combinatorial properties). Hence, a rule such as Passivization that applies to transitive verbs cannot be formulated as a lexical rule. However, the cross-categorial nature of Passivization can be captured by formulating it as a phrasal rule that operates on transitive verb phrases, for at the phrasal level of representation, both *persuade to go* and *hit* are transitive verb phrases.

As evidence in favor of this phrasal analysis of Passivization over a lexical approach, Bach (1979; 1980) cites the existence of infinitivals that may modify transitive verb phrases and passive verb phrases, but not intransitive verb phrases:

(109) a. Mary bought it to drive to work in.
b. It was bought to drive to work in.
c. *The Dean came in to talk to.

However, the assumption that these infinitivals can modify only *transitive* verb phrases leaves examples like (110a–c) unexplained:

(110) a. My car is always available (for you) to drive to work in.
b. I am always on hand for students to talk to.
c. It will serve nicely to fix drinks on.

It is clear that these infinitivals are not extraposed infinitival relative clauses, for possessed NPs and definite personal pronouns resist relative clauses in English: **My car (for you) to drive to work in is always available, *I for students to talk to am always on hand, *It to fix drinks on will serve nicely*. Nor can the verb phrases containing the adjective *available,* the idiomatic prepositional phrase *on hand,* and the intransitive verb *serve* be naturally analyzed as transitive. Moreover, many transitive verb phrases fail to permit these infinitival modifiers (e.g. **Mary ate it to amuse John with*). Thus, it appears that these infinitival modifiers are neither necessary nor sufficient indicators of transitivity. More generally, Halvorsen (1981) has shown that a variety of arguments for a phrasal passive in the recent literature (Bach (1980), Keenan (1980)) dissolve when the lexical analysis of passivization is coupled with a theory of the semantic interpretation of functional structures.

Is there then any evidence at all to choose between Bach's explanation of Visser's generalization and the explanation proposed here? The fundamental difference between the two is that on our account, Passivization fails because the functional control relation between the subject and the complement is destroyed, while on Bach's account, Passivization fails because intransitive verb phrases cannot be passivized. Thus, according to our analysis, the transitivity or intransitivity of the verb undergoing Passivization is irrelevant; the crucial factor is that a (subject or object) AGENT controller of the complement is not syntactically available. Therefore, our analysis would predict that an *intransitive* verb whose subject functionally controlled a predicative complement could not be passivized, even if intransitive verbs otherwise undergo Passivization. Indeed, in the lexical theory of passivization (Bresnan (1982b)), intransitive Passivization can be formulated as a subcase of transitive Passivization, resulting from omission of the optional operation that replaces OBJ by SUBJ:

(111) $(\text{SUBJ}) \rightarrow (\text{OBL}_{\text{AG}})/\phi$
$((\text{OBJ}) \rightarrow (\text{SUBJ}))$

In Bach's account, by contrast, transitivity is the crucial factor; the syntactic availability

of the AGENT controller is not important. In Bach's theory of passivization, the AGENT is existentially quantified over the passivized verb phrase, as in (112); hence, the AGENT is semantically available as a controller, even if it is not syntactically expressed.

(112) John is regarded as happy.
$j^*(\hat{}\lambda x\ \exists y\ [[regard'\ (\hat{}happy')]\ (\hat{P}P(x))](y))$
$= \exists y\ [[[regard'\ (\hat{}happy')]\ (\hat{P}P(\hat{}j))](y)]$

Thus, Bach's account of Visser's generalization would not generalize to the case of intransitive passives; in particular, if intransitive verbs could undergo Passivization, an *intransitive* verb whose subject controlled a predicative complement could be passivized. (More detailed reasoning in support of this conclusion is given below.)

Intransitive Passivization does exist in a number of languages, Icelandic and Norwegian among them. (The following evidence from Icelandic was provided by Joan Maling.) In Icelandic, intransitive agentive verbs undergo Passivization, as in (113)–(114).

(113) Allir keyra hættulega í bænum.
all-nom-masc-pl drive-pl dangerously in town-dat
'Everyone drives dangerously in town.'

(114) Það er keyrt hættulega í bænum.
it be-sg-3 driven-nom-neut-sg dangerously in town-dat
'There is dangerous driving in town.'
[Lit: 'It is driven dangerously in town.']

The morphology of the intransitive passive is the same as that of the transitive passive, except that the passive participle appears in the neuter nominative singular form (the unmarked gender, case, and number). Expression of the agent in intransitive Passivization is marginal. There is evidence that *það* is a syntactic placeholder rather than a subject (Maling and Zaenen (1978)).

In Icelandic, a predicative state complement controlled by the subject can appear with intransitive verbs. Unlike an adverb, this complement agrees in gender, case, and number with the subject:

(115) Allir keyra fullir í bænum.
all-nom-masc-pl drive drunk-nom-masc-pl in town
'Everyone drives drunk in town.'

However, no form of this subject-controlled complement can appear in the passivized form of (115), as (116) shows:

(116) *Það er keyrt fullt/fullum/fullir
it be-sg-3 driven-nom-neut-sg drunk-neut-sg/drunk-dat-pl/drunk-nom-pl
í bænum.
in town
('There is driving drunk in town.')
[Lit: 'It is driven drunk in town.']

The ungrammaticality of (116) is predicted by our theory, but not by Bach's. (See below.)

Similar evidence exists in Norwegian. (The following examples from Norwegian were provided by Per-Kristian Halvorsen.) In Norwegian, as in Icelandic, agentive intransitive verbs may undergo Passivization as transitive verbs do, except that the agent is not syntactically expressed.

(117) De kjørte gjennom byen.
they drove through the-town
'They drove through the town.'

(118) Det ble kjørt gjennom byen.
it was driven through the-town
'There was driving through the town.'
[Lit: 'It was driven through the town.']

In Norwegian, a predicative state complement modifying the subject may appear in the active but not the passive form of the sentence:

(119) De kjørte fulle gjennom byen.
they drove drunk-pl through the-town
'They drove drunk through the town.'

(120) *Det ble kjørt fulle/full/fullt gjennom byen.
it was driven drunk-pl/drunk-sg/drunk-neut-sg through the-town
('There was driving drunk through the town.')
[Lit: 'It was driven drunk through the town.']

Thus, the evidence from intransitive passivization in Icelandic and Norwegian confirms our theory that Visser's generalization follows from a syntactic restriction on subject control rather than a condition on transitive verb phrase passivization.

The following reasoning (suggested by P.-K. Halvorsen) supports this conclusion in detail. In general, Bach handles control of infinitival VPs by forming a lambda abstract of the open sentence consisting of the VP and a variable for the subject. Thus, *John persuaded Mary to go* is built up from *persuade to go* and *Mary, John* being added last. Simplifying inessential details, the translation of *persuade to go* is an abstract: $\lambda x[persuade'(\hat{}go'(x))(x)]$. When this is combined with the translation of *Mary*, the x is bound and *Mary* will be interpreted to be the subject of the complement *go*. Passivization existentially quantifies the subject and abstracts over the object, producing $\lambda z[\exists y\ \lambda x\ [persuade'(\hat{}go'(x))\ (x)]\ (z)(y)]$, which reduces to $\lambda z[\exists y\ persuade'(\hat{}go'(z))(z)(y)]$. What would happen in the case of passive intransitives such as *Det ble kjørt* 'There was driving' [Lit: 'It was driven (intrans)'] in Norwegian? The passive intransitive could be translated as $\exists y[drive'(y)]$ (again simplifying inessential details). This is of the wrong type for an IVP, so let us assume that the translation is $\lambda x \exists y[drive'(y)]$, using vacuous lambda abstraction to perform a type adjustment. When the placeholder subject *det* 'it' is combined with the IVP, the lambda disappears and the translation $\exists y[drive'(y)]$ remains for the sentence *Det ble kjørt*. What if there were a complement? An example such as **Det ble kjørt fulle* [Lit: 'It was driven drunk'] would be built up from *kjørt fulle* and *det*. The active phrase *kjørte fulle* would be translated $\lambda x[drive'(\hat{}drunk'(x))(x)]$, in

a fashion similar to *try to go*. (Note that, like *try*, *drive* can be used intransitively, transitively, or with a complement. For evidence that state predicates like *drunk* are (controlled) complements rather than adjuncts, see section 8 and Neidle (1982).) The translation rule for passive would then give $\lambda z[\exists y \hat{} \lambda x[drive'(\hat{}drunk'(x))(x)](y)]$, which reduces to $\lambda z \exists y[drive'(\hat{}drunk'(y))(y)]$, which again combines with the placeholder subject *det* to yield the well-formed interpretation $\exists y[drive'(\hat{}drunk'(y))(y)]$ 'There was someone driving drunk'.

In conclusion, Visser's generalization follows in its full generality from our theory. Despite proposals by Jenkins (1972), Anderson (1977), Chomsky (1977b; 1980b), Wasow (1977), Williams (1980), and Bach (1979; 1980), other theories leave it unexplained.

9.5. Bach's Generalization

In his study of passivization, Bach (1979) has also observed that where the object of a verb is an obligatory controller, intransitivization is impossible. This is illustrated in (121) and (122). In (121), the subject is the controller, and the object may be omitted; in (122), the object is the controller, and may not be omitted.

(121) a. Louise promised Tom to be on time.
b. Louise promised to be on time.
(122) a. Louise taught Tom to smoke.
b. *Louise taught to smoke.

In fact, however, not all cases of obligatory control by an object rule out intransitivization (a fact not explained by Bach's (1979) analysis):

(123) a. Louise signaled Tom to follow her.
b. Louise signaled to follow her.

What is the difference between (122) and (123)? Note that an NP subject can be expressed in the case of (123) but not (122):

(124) *Louise taught Tom for him to smoke.
(125) Louise signaled Tom for him to follow her.

On the basis of this fact, we may tentatively hypothesize that the object in (122a) is a functional controller of its complement, while the object in (123) is an anaphoric controller of its complement. Thus, we can amend Bach's generalization to state that where the object of a verb is a functional controller, intransitivization is impossible. In this form, the generalization appears quite generally to be true:

(126) a. John called someone.
b. John called.
c. John called someone a fool.
d. *John called a fool.

(127) a. John painted something.
b. John painted.
c. John painted something black.
d. *John painted black.

This generalization is the counterpart for objects of Visser's generalization for subjects: since only subjects and objects can be functional controllers, it follows that any lexical operations that eliminate subjects or objects will be inapplicable where they would destroy functional control relations. Consider the control relation in example (122). Assuming that *teach* has a basic lexical form 'TEACH⟨(SUBJ)(OBJ)(XCOMP)⟩', the lexical rule of Functional Control (section 8.2) obligatorily adds the control relation (↑ OBJ) = (↑ XCOMP SUBJ) to the lexical entry. The lexical rule of Intransitivization replaces (OBJ) by ϕ (Bresnan (1980)); it therefore could not apply to *teach* without destroying the functional control relation, since only the semantically unrestricted functions can be functional controllers, and ϕ is not a function (see footnote 1).

9.6. *Raising*

Raising refers to the type of construction illustrated in (128). The italicized NPs are sometimes referred to as "raised" NPs.

(128) a. John considered *himself* for a long time to be inferior to his brother.
b. *Mary* seems to like witty women.

Two properties characterize these examples of raising constructions: first, the verb has a predicative complement whose understood subject is referentially dependent upon (in fact, identical to) a pre- or postverbal NP; and second, this pre- or postverbal NP is not a "logical", or thematic, argument of the verb. From the first property we see that raising is a type of control relation (section 8). From the second property, we see that it must be a *functional* control relation. To see this, suppose that raising involved anaphoric control. The sentences would then be semantically incoherent, containing a subject or object (the "raised" NP) that cannot be bound to any argument position (Halvorsen (1981)). Thus, raising must be a functional control relation. From our theory of functional control it follows at once that only the SUBJ of a predicative complement can be "raised" and that "raised" NPs can appear only as SUBJ, OBJ, or OBJ2.

Let us now consider the problem of how raising constructions are expressed in c-structure. The government and binding theory of Chomsky (1981) incorporates the *Projection Principle* and *Theta Criterion* as fundamental representational assumptions; together, they imply that all representations of syntactic structure are isomorphic to the representation of thematic structure and reflect the subcategorization properties of lexical items. Given these assumptions, raising constructions must have syntactic structures of the form indicated in (129) and (130). The motivation for this analysis is that in their thematic structures, raising verbs have a propositional argument, and $\bar{S}$ is the syntactic form of such an argument.

(129) [$_S$ John considered [$_{\bar{S}}$ himself for a long time to be inferior to his brother]]
(130) [$_S$ Mary seems [$_{\bar{S}}$[$_{NP}$ e] to like witty women]]

However, this analysis is impossible in our theory, for by the principles of lexical and syntactic encoding (sections 2 and 4), $\bar{S}$-complement clauses can only be assigned closed functions such as COMP(lement), which cannot be functionally controlled. (Moreover, of course, we are constrained from using null structure to represent local dependencies (section 6).) Furthermore, if we assumed that there is no functional control at all in object-raising constructions like (129), we would incorrectly predict that raised objects could not passivize. *John is believed to be smart* would be ill-formed for the same reason that **John is believed is smart* is ill-formed: namely, because *John* is not the object of *believe* in such structures (Bresnan (1982b)). Hence, the c-structures shown in (129) and (130) are inconsistent with our theory. This analysis also poses a theoretical problem for Chomsky's government and binding theory, because the complement subject NPs must be governed and assigned case, the $\bar{S}$-complements lack finite INFL to assign case, and government from a matrix verb normally cannot cross a maximal projection such as $\bar{S}$. Chomsky therefore postulates a rule which deletes the $\bar{S}$, leaving only S, across which the matrix verbs are supposed to be able to govern the NP subject of a complement (Chomsky (1981, 66)). However, $\bar{S}$ Deletion has just the properties of strict locality and lexical specificity that characterize lexical subcategorization. It is evident that the use of such lexically governed "pruning transformations" (the term is due to Ross (1969)), like the use of transformational exception features elsewhere (Lakoff (1970)), weakens the theory of structure-dependent rules (Bresnan (1975)) and circumvents the strong constraint imposed by the Projection Principle. Since such syntactic restructuring processes are prohibited from our much more restrictive theory of constituency relations (section 6) in any event, it might seem preferable simply to hypothesize that the raising verbs directly subcategorize for S rather than $\bar{S}$:

(131) [$_S$ John considered [$_S$ himself for a long time to be inferior to his brother]]
(132) [$_S$ Mary seems [$_S$[$_{NP}$ e] to like witty women]]

But this option, too, is closed to us, for it follows from our theory of lexically and syntactically encoded functions that in configurational structures, only *maximal* projections can be subcategorized for, and S is submaximal (section 3).

Thus, our theory requires that the infinitival complements to raising verbs like *consider* and *seem* be maximal projections other than $\bar{S}$. Since a verb heads these complements in the above examples, the natural choice of category is VP, the maximal projection of V. However, the category VP cannot contain a structural subject (section 4). Hence, the only analysis open to us is the one shown in (133), in which the NP *himself* is generated as the object of *considered* outside the predicative VP complement.

(133) [$_S$ John considered [$_{NP}$ himself] for a long time [$_{VP}$ to be inferior to his brother]]
(134) [$_S$[$_{NP}$ Mary] seems [$_{VP}$ to like witty women]]

Note that our theory requires only that the complement to a raising verb be a maximal projection other than $\bar{S}$. Nothing in our theory requires that it be a VP (as opposed to another phrasal category which is a maximal projection). Hence, in our theory, raising should be possible with phrasal complements other than VP; in fact, both *consider* and *seem* also allow phrasal complements of categories other than VP:

(135) John considers himself [$_{AP}$ inferior to his brother]
(136) Mary seems [$_{AP}$ fond of witty women]

Moreover, for many speakers, *consider* and *seem* in these uses may have nonthematic objects and subjects (respectively); compare *Consider your goose cooked,* which can be idiomatic, and *Close tabs seem likely to be kept on celebrities.*

Now according to our analysis, the raising verbs *consider* and *seem* take phrasal complements; they differ in that *consider* is transitive, having a nonthematic object, while *seem* is intransitive, having a nonthematic subject. The phrasal complements in (133)–(136) have the function XCOMP, and the lexical entries for *consider* and *seem* will therefore include the lexical forms given in (137) and (138).

(137) (↑ PRED) = 'CONSIDER⟨(SUBJ)(XCOMP)⟩(OBJ)'
(138) (↑ PRED) = 'SEEM⟨(XCOMP)⟩(SUBJ)'

The lexical rule of Functional Control (section 8.2) will expand these entries with the control equations shown in (139) and (140).

(139) (↑ PRED) = 'CONSIDER⟨(SUBJ)(XCOMP)⟩(OBJ)'
(↑ OBJ) = (↑ XCOMP SUBJ)
(140) (↑ PRED) = 'SEEM⟨(XCOMP)⟩(SUBJ)'
(↑ SUBJ) = (↑ XCOMP SUBJ)

Note that it is because the OBJ of *consider* and the SUBJ of *seem* are nonthematic, that they can designate expletives such as *there* and idiom chunks; see Bresnan (1982b) for a detailed analysis. In sum, the c-structures and f-structures of our raising constructions must be as shown in figures 21 and 22.

However, the analysis of raising required by our theory is incompatible with the basic representational assumptions of Chomsky's (1981) theory (the Projection Principle and the Theta Criterion). The assumptions of each theory require a phrase structure representation of raising constructions which is inconsistent with the assumptions of the other theory. Evidence bearing on the choice of phrasal representations of these constructions would therefore provide grounds for deciding which theory is more explanatory, in the sense of limiting the hypothesis space of grammar construction to the descriptively optimal grammars. In fact, there is strong evidence favoring the analysis of raising constructions given in figures 21 and 22 over Chomsky's sentential analyses in (129)–(132).

First, because the raised NP in figure 21 is an object of the matrix clause, it f-commands the subject in the matrix clause nucleus. If we had adopted the sentential

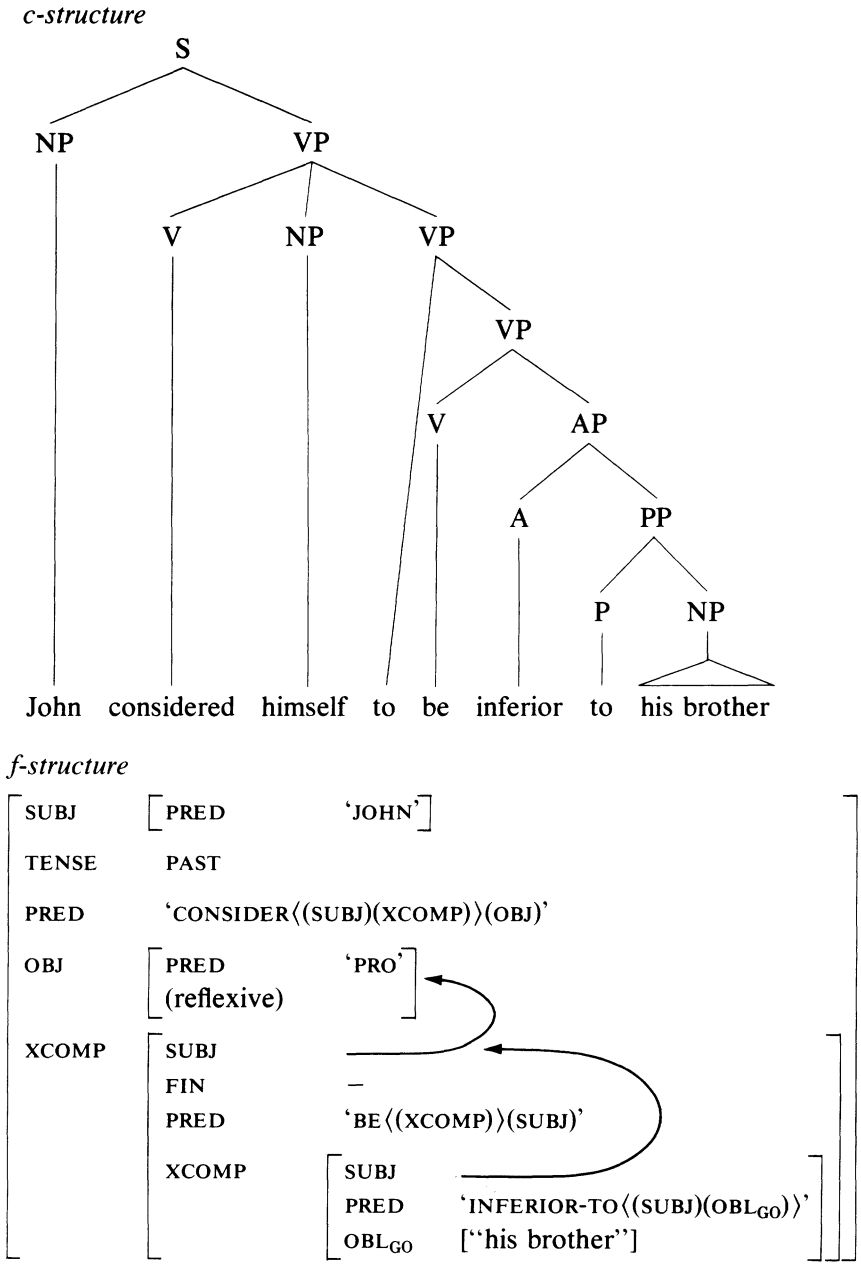


Figure 21
A raising-to-object construction

analysis, *himself* would be contained solely within the COMP(lement), in a position where it could not f-command the subject of the matrix clause nucleus. We therefore predict that anaphoric control of a subject clause by the object in a raising construction should be possible (cf. section 8.3). Examples (141a–c) bear out this prediction.

(141) a. Contradicting himself will prove Mr. Jones to be a liar.
b. *Contradicting himself will prove that Mr. Jones is a liar.
c. Contradicting himself will discredit Mr. Jones.

We see that the object of the raising construction in (141a) behaves like the object of the simple transitive verb in (141c), and not like the subject of the $\bar{S}$-complement in (141b).[16]

Further confirmation for the correctness of our analysis comes from the postposing of a "heavy" postverbal NP to VP-final position (noted in Bresnan (1972)). This rule is bounded, in the sense that it can only postpose the postverbal NP over its sister phrases (as observed in Bresnan (1976a)):

(142) a. I discussed [all those women riding motorcycles nowadays] with my sister.
b. I discussed with my sister [all those women riding motorcycles nowadays].
c. *I discussed [riding motorcycles nowadays] with my sister all those women.

(The boundedness property follows at once if the rule operates on the VP phrase structure rule, reordering the categories in a function-preserving manner.) Note in particular that the postverbal NP *all those women* cannot be shifted outside the clause that immediately contains it ((142c)). In our analysis of object raising constructions, the raised NP is the object of the raising verb, and hence may undergo Heavy NP Shift over material in the matrix VP without violating the boundedness of NP Shift. However, if the NP were dominated by an $\bar{S}$- or S-complement of the raising verb, shifting it over material in the matrix verb phrase *would* violate the boundedness of NP Shift. Both *count* and *consider* are object raising verbs in the author's dialect: cf. *I counted there as being more than 20 women on the payroll, I consider there to be too many fools in this committee.* The following examples therefore confirm our analysis.

(143) a. I'll count [as being dead] from now on any corpse more than 3 days old.
b. I will consider [to be fools] in the weeks ahead all those who drop this course.

These examples are naturally paraphrased as "From now on, I'll count any corpse . . . as being dead" and "In the weeks ahead, I will consider all those . . . to be fools,"

[16] These facts also show that it is insufficient to state the constraint on anaphoric control as a condition that the controller must be an *argument* of the matrix of the controlled complement (cf. Chomsky (1981)). Clearly, *Mr. Jones* is not a semantic argument of *prove* in *Contradicting himself will prove Mr. Jones to be a liar.*

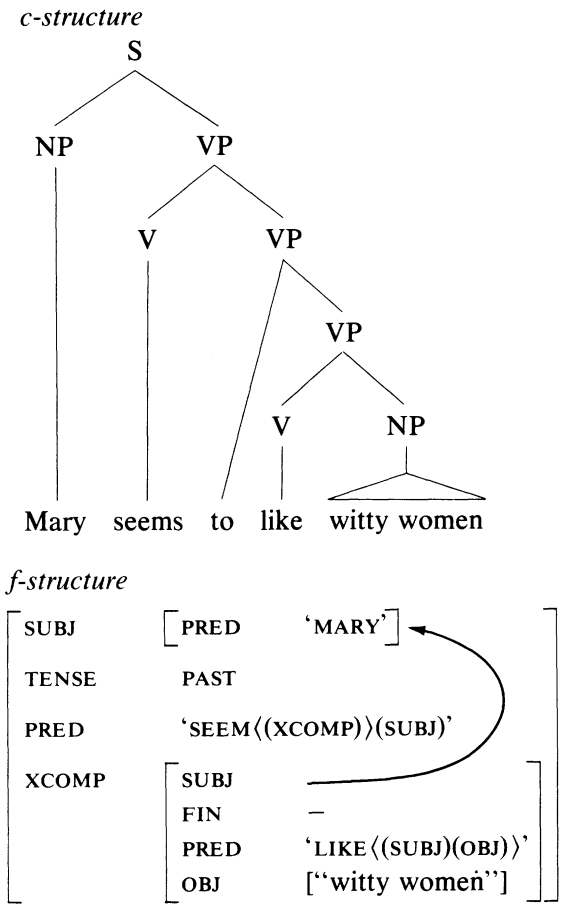


Figure 22
A raising-to-subject construction

with the adverbial material modifying the matrix clauses as shown. They demonstrate that in c-structure, the raised NP must not form a constituent with the verbal phrase that is predicated of it. (Postal (1974) makes a slightly different argument from NP Shift; his (1977) response to objections by Bresnan (1976c) indicates that his argument for a "raised" NP position in the VP is also valid.)

A third source of confirmation for our analysis is the rule of Right Node Raising, which permits a constituent of any category to be omitted from two conjoined clauses and positioned to their right, as shown in (144).

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

In the author's dialect, nonconstituent sequences of categories cannot undergo Right Node Raising.[17] Accordingly, examples (145a,b) contrast with (144a,b).

(145) a. Mary gave Peter the grape, and Catherine did not give Peter the grape.
b. *Mary gave, and Catherine did not give, Peter the grape.

Right Node Raising is a very useful test for constituency because it applies to all categories, including S (Bresnan (1974)):

(146) It is possible that, but it isn't obvious that, your theory is right.

In our analysis of object raising constructions, the raised NP and the predicative phrase that it controls do not form a constituent in c-structure; hence, we correctly predict the following contrasts in the restrictive dialect (cf. footnote 17):

(147) a. Mary believes Peter to be fat, but Catherine doesn't believe Peter to be fat.
b. *Mary believes, but Catherine doesn't believe, Peter to be fat.

A similar argument has been made by Postal (1974).

To account for these facts in the government and binding framework, it has been proposed that a verb can assign case only to an *adjacent* NP; Right Node Raising is held to destroy the connectivity of case assignment by violating the Adjacency Condition (Stowell (1981), Chomsky (1981)). Since, in fact, an object NP can be right-node-raised away from its governing verb (as in (144b)), the claim must be that case is transmitted to such an NP from its trace under Right Node Raising, and that there is no trace left of an NP (such as *Peter* in (147b)) which is properly contained within a right node raised constituent. Thus, we have two competing hypotheses to account for the ungrammaticality of examples like (147b). Our theory claims that *Peter to be fat* is a nonconstituent NP VP sequence; it therefore fails to undergo Right Node Raising (in the restrictive dialect) for exactly the same reason that other nonconstituent sequences (e.g. (144b)) fail to do so. The alternative theory claims that (147b) and (144b) are ill-formed for unrelated reasons. Example (144b) is said to be ill-formed (in the restrictive dialect) because it is a nonconstituent sequence, but *Peter to be fat* in (147b) is a constituent S, according to this theory. However, *Peter* is not case-marked by *believes* because Right Node Raising destroys the connectivity of case assignment.

Can we find evidence to distinguish between the constituency hypothesis and the connectivity hypothesis? Such evidence could take the form of a construction which required exceptional case marking (in terms of the government and binding theory) and yet behaved as a constituent under Right Node Raising. This evidence exists. Consider the "acc–*ing*" construction, illustrated in (148).

[17] Unlike the author, some speakers do weaken the condition on Right Node Raising to allow nonconstituent sequences (cf. Abbott (1976), Peterson (1981)). Thus, the Right Node Raising judgments of these speakers cannot be used as a test for constituency.

(148) a. All his friends worry about Fred losing again.
b. Doctors disapprove of children smoking.

"Acc–*ing*" constructions are similar in their interpretation to "poss–*ing*" (gerund) clauses, but they lack the genitive marker *'s* to assign case to their subject NPs (*Fred* in (148a), *children* in (148b)). In the government and binding theory, these NPs must be assigned case. Given that the acc–*ing* clauses lack a finite INFL, there are only two possible mechanisms for doing this: either the position of the subject NP is inherently case-marked, or it receives case exceptionally from a governing verb or preposition outside the clause. Now if the NP position were inherently case-marked, the acc–*ing* construction would occur in positions where exceptional government by a verb or preposition does not hold, such as the subject position of a finite clause. However, if the NP position were exceptionally case-marked, the acc–*ing* construction would not occur in subject position. Examples like (149) indicate that the subject of the acc–*ing* construction must be exceptionally governed.

(149) a. *Fred losing again is worried about by all his friends.
(cf. Fred's losing again is worried about by all his friends.)
b. *Children smoking is disapproved of by doctors.
(cf. Children's smoking is disapproved of by doctors.)

It is generally true that an acc–*ing* construction cannot function as a subject, even in nonfinite clauses:

(150) *I'd like for Fred losing again not to bother you.
(cf. I'd like for Fred's losing again not to bother you.)

In this respect, then, acc–*ing* constructions resemble object raising constructions. They differ crucially, however, in that the acc–*ing* construction gives every evidence of forming a single constituent:

(151) a. It's Fred losing that I can't stand the thought of.
b. What doctors really disapprove of is children smoking.

Indeed, this assumption (together with some version of the A-over-A Condition (Bresnan (1976a)) is needed by the government and binding theory to explain the failure of NP Movement in examples like **Children are disapproved of smoking by doctors*. We see now, though, that both clefting (151a) and pseudoclefting (151b) preserve the connectivity of exceptional case marking. And in fact, so does Right Node Raising:

(152) a. Mary's not looking forward to, and I can't stand the thought of, Fred losing again.
b. Doctors disapprove of, and parents worry about, children smoking.

Since these examples are perfectly well-formed, we must reject Stowell's and Chomsky's hypothesis, and conclude that Right Node Raising, like other dislocation rules, does preserve the connectivity of exceptional case marking. This conclusion gains plausibility

because Right Node Raising preserves connectivity in general; for example, it preserves the lexically governed control relations in (153a,b).

(153) a. Mary regards Bill, but Sue doesn't regard Bill, as too fond of himself.
b. Mary strikes Bill, but she doesn't strike Peter, as too fond of herself.

Consequently, failure of connectivity cannot be the explanation for the ungrammaticality of (147b) **Mary believes, but Catherine doesn't believe, Peter to be fat*. The nonconstituency hypothesis, which is the simpler account, provides a superior explanation.

In summary, our theory requires an analysis of raising constructions in which raising verbs take nonthematic subjects or objects and predicative phrasal complements. There is strong evidence that the analysis of raising constructions required by our theory is optimal for the grammatical description of English. Recent work by Mohanan (1982a) shows that the essential features of this analysis are supported by raising constructions in Malayalam, a non-Indoeuropean language which differs strikingly from English in its constituent structure properties. Since the optimal grammatical descriptions of raising constructions are required by the representational assumptions of our theory, but inconsistent with those of the government and binding theory, we conclude that our theory is the more explanatory.

9.7. *Pro-Drop*

Pro-drop is a widespread linguistic phenomenon in which, under certain conditions, a structural NP may be unexpressed, giving rise to a pronominal interpretation. Given our constraints on the use of null c-structure categories (section 6) and our postulate that all semantic forms are introduced lexically (section 8.3), it follows that in the unmarked case, pro-drop should arise for the functions SUBJ, OBJ, and OBJ2. The reasoning is exactly analogous to that given for the representation of the controlled ([+U]) 'PRO' in the beginning of section 8.3. Moreover, it follows from our theory of syntactic encoding that when these functions are nonconfigurationally encoded, pro-drop should always be possible. The reason is that nonconfigurational encoding pairs functional schemata with arbitrary X expanded by rules of the form $C \rightarrow X^*$ (section 4). Hence, the categorial expression of the SUBJ, OBJ, or OBJ2 is always optional. From functional completeness (section 5) it follows that when the categorial expressions of SUBJ, OBJ, or OBJ2 are omitted, the optional functional anaphors carried by the governing lexical forms will be required. Thus, in the unmarked case, pro-drop of SUBJ, OBJ, and OBJ2 will always be possible in nonconfigurational languages.

Investigating this consequence will require careful descriptive studies of nonconfigurational encoding, distinguishing between pro-drop and the rather different phenomena of morphologically incorporated oblique pronouns (as in Navajo (Roberts (1981))) and oblique pronominal clitics (as in Warlpiri auxiliaries (Simpson (in preparation))). However, Hale (to appear) has already observed that extensive pro-drop is a hallmark of nonconfigurational languages, and research on nonconfigurational encoding

in Malayalam (Mohanan (1981b; 1982a)) and Warlpiri (Simpson and Bresnan (in preparation), Simpson (in preparation)) bears out this consequence of our theory.

10. Conclusion

All theories of grammar can be regarded as theories of *grammatical relations* in the broad sense, that is, theories of the relations between the surface constituents of sentences and their semantic predicate argument structures. One of the most important issues in current linguistic theory concerns the nature of grammatical relations: are they completely derivative of representations of constituent structure and semantic predicate argument structure, or do they have an independent representation definable in terms of syntactically primitive grammatical functions (e.g. SUBJ, OBJ)? While this issue will undoubtedly continue to be debated, one conception of the problem must be rejected as fundamentally mistaken. That is the view that theories which take only structuralist concepts as syntactic primitives are inherently simpler and more explanatory than theories which take functions as syntactic primitives, because the latter theories involve an extra level of representational concepts. This view is mistaken because in constructing a theory of grammar, we are free to take any set of concepts as primitive, defining in terms of these a larger set of derivative concepts; and the linguistic concepts of function, category, government, argument, case, and constituent structure are so closely interrelated as to be interdefinable. In particular, the structuralist concept of category can be reduced to more primitive functional concepts. Thus, in the theory presented in this study, the grammatical functions (SUBJ, OBJ, etc.) are taken as syntactic primitives and the phrase structure categories (NP, VP, etc.) are derivative notions. Contrary to this common but erroneous view, there is simply no a priori basis for preferring a structuralist theory of grammatical relations to a theory which employs grammatical functions as syntactic primitives. On the contrary, there appears to be good reason to prefer the function-based theory, since only in this type of theory has it been shown explicitly that there is a universal representation of grammatical relations for both configurational and nonconfigurational language types, and that the radical differences in the structural expressions of subcategorization, government, and control relations in these languages are predictable from a single parameter of variation: the surface realization of grammatical functions by c-structure configuration or by morphological case.

It is the axioms, or postulates, of a theory which give meaning to its primitive concepts. For example, our theory postulates a narrowly constrained set of relations between functions and lexical predicate argument structures, on the one hand, and between functions and phrasal and morphological structures, on the other hand. From these we have derived several significant generalizations. For example, it follows from these assumptions of our theory that in configurational structures (a) the head of a phrase governs the maximal projections that are immediately dominated by any of the projections of the head and (b) lexical items will appear to subcategorize for maximal projections of categories. In nonconfigurational structures, it follows (a) that lexical items will

appear to subcategorize for morphological case and not for phrase structure order, and (b) that (in the unmarked case) pro-drop of SUBJ, OBJ, and OBJ2 will always be possible.

Similarly, our theory postulates that the set of grammatical functions is partitioned into classes which differ in their lexical encoding properties; in particular, the semantically restricted functions are lexically encoded in much more limited ways than the semantically unrestricted functions. All lexically represented grammatical relations are thereby constrained, including lexically induced control relations and lexically carried functional anaphors. Among the various consequences we have enumerated, it follows in our theory (a) that only SUBJ, OBJ, and OBJ2 are possible as lexically induced functional controllers; (b) that morphosyntactically unexpressed 'PRO's can bear only these functions; and (c) that lexical transformations of control relations—such as Passivization (Visser's generalization) and Intransitivization (Bach's generalization)—are severely restricted.

Finally, our theory postulates that the SUBJ function is universally the controlled element in functional control relations, and that functional control must be distinguished from anaphoric control. Evidence has been given that in anaphoric control relations, objects as well as subjects can be controlled, as our theory predicts. While other theories of control have been based on rather different assumptions, our theory appears to offer a much more constrained and explanatory account of government, anaphoric control, and raising. For example, we have obtained an explanation of the distribution of governing morphemes, a uniform cross-configurational characterization of government relations, and the predictability of the actual structural configurations in which government relations can appear. This is achieved within a narrowly constrained theory of constituency relations, which precludes the use of "Affix Hopping" and other structure-deforming rules. Similarly, we have shown that anaphoric control relations are subject to a condition definable on f-structure, namely, f-command, but not on c-structure. And last, we have seen that our theory predicts the essential properties of raising constructions: that only subjects undergo raising, that the subjects are raised only into subject or object positions in the matrix clause, and that the propositional complement of the raising predicate is expressed in c-structure discontinuously, in the form of a nonthematic subject or object constituent and a separate phrasal complement constituent. This last result is of particular interest, since it is inconsistent with the assumption that phrase structure relations directly reflect semantic predicate argument structure (as in the Projection Principle and Theta Criterion of Chomsky (1981)).

References

Abbott, B. (1976) "Right Node Raising as a Test for Constituenthood," *Linguistic Inquiry* 7, 639–642.

Amritavalli, R. (1980) "Expressing Cross-Categorial Selectional Correspondences: An Alternative to the $\bar{X}$ Syntax Approach," *Linguistic Analysis* 6, 305–343.

Anderson, S. (1977) "Comments on the Paper by Wasow," in P. Culicover, T. Wasow, and A. Akmajian, eds. (1977).

Andrews, A. (1981) "The Rudiments of Coordinate Structure in LFG," ms., Department of Linguistics, Australian National University, Canberra, Australia.
Andrews, A. (1982) "The Representation of Case in Modern Icelandic," in J. Bresnan, ed. (1982a).
Andrews, A. (to appear) "Long Distance Agreement and Functional Interpretation," in P. Jacobson and G. Pullum, eds. (to appear).
Bach, E. (1979) "Control in Montague Grammar," *Linguistic Inquiry* 10, 515–531.
Bach, E. (1980) "In Defense of Passive," *Linguistics and Philosophy* 3, 297–341.
Baker, C. L. (1979) "Syntactic Theory and the Projection Problem," *Linguistic Inquiry* 10, 533–581.
Bloomfield, L. (1933) *Language,* Holt, Rinehart and Winston, New York.
Borkin, A. (1974) *Raising to Object Position: A Study in the Syntax and Semantics of Clause Merging,* Doctoral dissertation, University of Michigan, Ann Arbor, Michigan; in *University of Michigan Papers in Linguistics* 2.1.
Bresnan, J. (1972) *Theory of Complementation in English Syntax,* Doctoral dissertation, MIT, Cambridge, Massachusetts; published in 1979 by Garland Press, New York.
Bresnan, J. (1974) "The Position of Certain Clause-Particles in Phrase Structure," *Linguistic Inquiry* 5, 614–619.
Bresnan, J. (1975) "Transformations and Categories in Syntax," presented at the fifth International Congress of Logic, Methodology and Philosophy of Science, London, Ontario, Canada, 1975; in R. Butts and J. Hintikka, eds. (1977) *Basic Problems in Methodology and Linguistics,* Reidel, Dordrecht.
Bresnan, J. (1976a) "On the Form and Functioning of Transformations," *Linguistic Inquiry* 7, 3–40.
Bresnan, J. (1976b) "Toward a Realistic Model of Transformational Grammar," presented at the Convocation on Communications, March 9–10, 1976, MIT, Cambridge, Massachusetts.
Bresnan, J. (1976c) "Nonarguments for Raising," *Linguistic Inquiry* 7, 485–501.
Bresnan, J. (1978) "A Realistic Transformational Grammar," in M. Halle, J. Bresnan, and G. Miller, eds., *Linguistic Theory and Psychological Reality,* MIT Press, Cambridge, Massachusetts.
Bresnan, J. (1979) "A Theory of Grammatical Representation," duplicated lecture notes, Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Bresnan, J. (1980) "Polyadicity," in T. Hoekstra, H. van der Hulst, and M. Moortgat, eds. (1980); in revised form in J. Bresnan, ed. (1982a).
Bresnan, J. (1981) "An Approach to Universal Grammar and the Mental Representation of Language," *Cognition* 10, 39–52.
Bresnan, J., ed. (1982a) *The Mental Representation of Grammatical Relations,* MIT Press, Cambridge, Massachusetts.
Bresnan, J. (1982b) "The Passive in Lexical Theory," in J. Bresnan, ed. (1982a).
Bresnan, J. and J. Grimshaw (1978) "The Syntax of Free Relatives in English," *Linguistic Inquiry* 9, 331–391.
Bresnan, J., R. Kaplan, P. S. Peters, and A. Zaenen (in preparation) "Cross-Serial Dependencies in Dutch," forthcoming in *Linguistic Inquiry* 13.4.
Chomsky, N. (1965) *Aspects of the Theory of Syntax,* MIT Press, Cambridge, Massachusetts.
Chomsky, N. (1969) "Quine's Empirical Assumptions," in D. Davidson and J. Hintikka, eds., *Words and Objections: Essays on the Work of W. V. Quine,* Humanities Press, New York.
Chomsky, N. (1970) "Remarks on Nominalization," in R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar,* Ginn, Waltham, Massachusetts.
Chomsky, N. (1977a) "On *Wh*-Movement," in P. Culicover, T. Wasow, and A. Akmajian, eds. (1977).

Chomsky, N. (1977b) *Essays on Form and Interpretation,* North-Holland, New York.
Chomsky, N. (1980a) "On the Representation of Form and Function," presented at the C.N.R.S. Conference at Royaumont, France, June 1980.
Chomsky, N. (1980b) "On Binding," *Linguistic Inquiry* 11, 1–46.
Chomsky, N. (1981) *Lectures on Government and Binding,* Foris, Dordrecht.
Chomsky, N. and M. Halle (1968) *The Sound Pattern of English,* Harper and Row, New York.
Comrie, B. (1977) "In Defense of Spontaneous Demotion: The Impersonal Passive," in P. Cole and J. Sadock, eds., *Syntax and Semantics, Vol. 8: Grammatical Relations,* Academic Press, New York.
Craig, C. (1977) *The Structure of Jacaltec,* The University of Texas Press, Austin, Texas.
Culicover, P., T. Wasow, and A. Akmajian, eds. (1977) *Formal Syntax,* Academic Press, New York.
Dowty, D. (to appear) "Grammatical Relations and Montague Grammar," in P. Jacobson and G. Pullum, eds. (to appear).
Emonds, J. (1970) *Root and Structure-Preserving Transformations,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Emonds, J. (1976) *A Transformational Approach to English Syntax,* Academic Press, New York.
Evers, A. (1975) *The Transformational Cycle in Dutch and German,* Doctoral dissertation, Rijksuniversiteit, Utrecht, Holland.
Falk, Y. (1980a) *Word Order and Rules of Syntax,* undergraduate thesis, Brandeis University, Waltham, Massachusetts.
Falk, Y. (1980b) "The English Auxiliary System: A Lexical-Functional Analysis," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Fassi Fehri, A. (1981) *Complémentation et Anaphore en Arabe Moderne: Une Approche Lexicale Fonctionnelle,* Thèse de Doctorat d'État, Université de Paris III.
Fiengo, R. (1977) "On Trace Theory," *Linguistic Inquiry* 8, 35–61.
Ford, M. (1982) "Sentence Planning Units: Implications for the Speaker's Representation of Meaningful Relations Underlying Sentences," in J. Bresnan, ed. (1982a).
Ford, M., J. Bresnan, and R. Kaplan (1982) "A Competence-Based Theory of Syntactic Closure," in J. Bresnan, ed. (1982a).
Gazdar, G. (in press) "Phrase Structure Grammar," in P. Jacobson and G. Pullum, eds. (to appear).
Grimshaw, J. (1979) "Complement Selection and the Lexicon," *Linguistic Inquiry* 10, 279–326.
Grimshaw, J. (1981) "Subcategorization of Grammatical Relations," presented at the Harvard Mini-Conference on the Representation of Grammatical Relations, Harvard University, December 12, 1981.
Grimshaw, J. (1982) "On the Lexical Representation of Romance Reflexive Clitics," in J. Bresnan, ed. (1982a).
Hale, K. (1978) "Obviation in Modern Irish," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Hale, K. (1979) "On the Position of Walbiri in a Typology of the Base," published in 1981 by the Indiana University Linguistics Club, Bloomington, Indiana.
Hale, K. (to appear) "Preliminary Remarks on Configurationality," *Proceedings of the Twelfth Annual Meeting of NELS,* MIT, Cambridge, Massachusetts.
Halliday, M. (1967) "Notes on Transitivity and Theme," *Journal of Linguistics* 1, 37–81.
Halvorsen, P.-K. (1981) "An Interpretation Procedure for Functional Structures," ms., The Center for Cognitive Science, MIT, Cambridge, Massachusetts.
Hoekstra, T., H. van der Hulst, and M. Moortgat, eds. (1980) *Lexical Grammar,* Foris, Dordrecht.
Hornstein, N. (1977) "S and X′ Convention," *Linguistic Analysis* 3, 137–176.

Hornstein, N. and A. Weinberg (1981) "Case Theory and Preposition Stranding," *Linguistic Inquiry* 12, 55–91.
Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar,* MIT Press, Cambridge, Massachusetts.
Jackendoff, R. (1976) "Toward an Explanatory Semantic Representation," *Linguistic Inquiry* 7, 89–150.
Jackendoff, R. (1977) *$\bar{X}$ Syntax: A Study of Phrase Structure,* Linguistic Inquiry Monograph 2, MIT Press, Cambridge, Massachusetts.
Jacobson, P. and G. Pullum, eds. (to appear) *On the Nature of Syntactic Representation,* Reidel, Dordrecht.
Jenkins, L. (1972) *Modality in English Syntax,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Kaplan, R. and J. Bresnan (1982) "Lexical-Functional Grammar: A Formal System for Grammatical Representation," in J. Bresnan, ed. (1982a).
Keenan, E. (1980) "Passive Is Phrasal Not (Sentential or Lexical)," in T. Hoekstra, H. van der Hulst, and M. Moortgat, eds. (1980).
Kisala, J. (in preparation) "Thematically-Conditioned Control Phenomena in Nominals," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Lakoff, G. (1970) *Irregularity in Syntax,* Holt, Rinehart and Winston, New York.
Lapointe, S. (1980) *A Theory of Grammatical Agreement,* Doctoral dissertation, The University of Massachusetts, Amherst.
Lehiste, I. (1970) *Suprasegmentals,* MIT Press, Cambridge, Massachusetts.
Levin, L. (1982) "Sluicing: A Lexical Interpretation Procedure," in J. Bresnan, ed. (1982a).
Levin, L. and J. Simpson (1981) "Quirky Case and Lexical Representations of Icelandic Verbs," presented at the Chicago Linguistic Society, April 30, 1981.
Lieber, R. (1980) *The Organization of the Lexicon,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
McCloskey, J. (1979) *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish,* Reidel, Dordrecht.
McCloskey, J. (1980) "Is There Raising in Modern Irish?" *Eriu* 31, 59–99.
McCloskey, J. (1981) "Notes on Case and Control in Modern Irish," ms., Department of Irish, University College Dublin, Ireland.
Maling, J. (to appear) "Transitive Adjectives: A Case of Categorial Reanalysis," in F. Heny, ed., *Papers from the Fourth Groningen Round Table, July 1980.*
Maling, J. and A. Zaenen (1978) "The Nonuniversality of a Surface Filter," *Linguistic Inquiry* 9, 475–497.
Manzini, R. (1980) "On Control," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Marantz, A. (1981) *A Theory of Grammatical Relations,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Miller, G. A. and P. Johnson-Laird (1976) *Language and Perception,* Harvard University Press, Cambridge, Massachusetts.
Mohanan, K. P. (1981a) "Move NP or Lexical Rules: Evidence from Malayalam Causativization," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Mohanan, K. P. (1981b) "Grammatical and Anaphoric Control," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Mohanan, K. P. (1982a) "Grammatical Relations and Clause Structure in Malayalam," in J. Bresnan, ed. (1982a).

Mohanan, K. P. (1982b) "Infinitival Subjects, Government, and Abstract Case," *Linguistic Inquiry* 13.2.
Mohanan, K. P. (1982c) *Lexical Phonology,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Montague, R. (1973) "On the Proper Treatment of Quantification in Ordinary English," in J. Hintikka, J. Moravcsik, and P. Suppes, eds., *Approaches to Natural Language,* Reidel, Dordrecht.
Montalbetti, M. (1981) "Consistency and Clitics," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Nash, D. (1980) *Topics in Walbiri Grammar,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Neidle, C. (1982) "Case Agreement in Russian," in J. Bresnan, ed. (1982a).
Neidle, C. (in preparation) *Topics in Russian Syntax,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Partee, B., ed. (1976) *Montague Grammar,* Academic Press, New York.
Perlmutter, D., ed. (to appear) *Studies in Relational Grammar.*
Peterson, P. (1981a) "Conjunction in LFG," ms., The Center for Cognitive Science, MIT, Cambridge, Massachusetts.
Peterson, P. (1981b) *The Progressive Construction in English: An Analysis and Its Implications,* Doctoral dissertation, University of Newcastle, NSW, Australia.
Pinker, S. (1982) "A Theory of the Acquisition of Lexical-Interpretive Grammars," in J. Bresnan, ed. (1982a).
Pinker, S. and D. Lebeaux (1981) *Language Learnability and Language Development,* forthcoming from Harvard University Press, Cambridge, Massachusetts.
Postal, P. (1974) *On Raising: One Rule of English Grammar and Its Theoretical Implications,* MIT Press, Cambridge, Massachusetts.
Postal, P. (1977) "About a 'Nonargument' for Raising," *Linguistic Inquiry* 8, 141–154.
Rappaport, M. (1980) "On the Derivation of Derived Nominals," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Reinhart, T. (1976) *The Syntactic Domain of Anaphora,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Roberts, J. (1981) "Toward a Unified Analysis of the Passive in Navajo," ms., Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.
Robertson, J. (1976) *The Structure of Pronoun Incorporation in the Mayan Verbal Complex,* Doctoral dissertation, Harvard University, Cambridge, Massachusetts; published in 1980, Garland Press, New York.
Ross, J. (1969) "A Proposed Rule of Tree-Pruning," in D. Reibel and S. Schane, eds., *Modern Studies in English,* Prentice-Hall, Englewood Cliffs, New Jersey.
Rothstein, S. (1981) "On Preposition Stranding," ms., Department of Linguistics, MIT, Cambridge, Massachusetts.
Selkirk, E. (1981) "English Word Structure," ms., Department of Linguistics, The University of Massachusetts, Amherst.
Simpson, J. (in preparation) *Control and Predication in Warlpiri,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Simpson, J. and J. Bresnan (in preparation) "Control and Obviation in Warlpiri." [Paper presented at the First Annual West Coast Conference on Formal Linguistics, Stanford University, January 1982.]
Sridhar, S. (1979) "New Evidence for Spontaneous Demotion," *South Asian Languages Analysis* 1.

Stowell, T. (1981) *Origins of Phrase Structure,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Stump, G. (1981) *The Syntax and Semantics of Free Adjuncts and Absolutes*, Doctoral dissertation, Ohio State University, Columbus, Ohio.
Thompson, S. (1973) "On Subjectless Gerunds in English," *Foundations of Language* 9, 374–383.
Visser, F. Th. (1963–1973) *An Historical Syntax of the English Language,* E. J. Brill, Leiden.
Wasow, T. (1977) "Transformations and the Lexicon," in P. Culicover, T. Wasow, and A. Akmajian, eds. (1977).
Williams, E. (1980) "Predication," *Linguistic Inquiry* 11, 203–238.
Zaenen, A. (1981) "Characterizing Syntactic Binding Domains," *Occasional Paper #17,* The Center for Cognitive Science, MIT, Cambridge, Massachusetts.

20D-219
MIT
Cambridge, Massachusetts 02139