to my wife, Brandel, with love

Table of Contents

Introduction

This textbook, like all textbooks, was born of necessity. When I went looking for a suitable textbook for my course on Lexical-Functional Grammar at the Hebrew University of Jerusalem, I discovered that there wasn't one. So I decided to write one, based on my lecture notes. The writing accelerated when, while I was on sabbatical at Stanford University (August 1999–February 2000), Dikran Karagueuzian of CSLI Publications expressed interest in publishing it.

This textbook is not intended as an introduction to syntax. Throughout, it is assumed that the reader is familiar with elementary concepts of syntactic theory and with contemporary derivational syntactic theory (Government/Binding theory and/or the Minimalist Program). I believe that this approach is conducive to opening up a dialog between different "camps" within generative syntactic theory. It is a mistake for any student of contemporary linguistic theory to be taught a single theoretical framework as if it represents an overriding consensus in the field. Being that derivational theories have a recognized centrality within the field, the assumption behind this book is that students are first introduced to a derivational theory, and then at a more advanced level learn alternatives. (Coincidentally, or not so coincidentally, this situation also matches my teaching experience.) This book is aimed at such students, and therefore attempts to motivate the concepts and formalisms of LFG in relation to derivational approaches. It is my hope that this approach will also make this book an appropriate one for professional linguists who wish to acquaint themselves with the basic principles and concepts of LFG.

Unlike most expositions of LFG, this book focuses on English. While much has been done in LFG on other languages, and the typological reach of LFG is one of its strongest points, I believe that there is pedagogical value in focusing on a single language, one that the student knows. Many students are initially turned off by having to wade through data from an unfamiliar language. (I can attest to this from personal experience.) This approach also provides a more cohesive view of the theory than jumping from language to

language would. It allows us to develop a minigrammar for the language, as is standard in textbooks on other formal theories, such as Akmajian and Heny (1975) on the Standard Theory and Sag and Wasow (1999) on Head-driven Phrase Structure Grammar.

This textbook was written by a descriptively oriented generative syntactician for other descriptively oriented generative syntacticians. As a result, there are many issues that are important in LFG that are not raised here in any serious way. For example, there is no discussion of the mathematical properties of the LFG formalisms or of computational applications, even though both of these have always been central concerns in LFG research. Throughout, the formalism is justified on the basis of descriptive linguistic considerations. Similarly, there is no discussion here of "glue-language" semantics or other issues concerning the relation between LFG syntax and other components of the grammar. References are made to the literature on some of these issues, and the interested student can pursue them given the background provided by this book.

Like any living theory, LFG is continually developing, and there are disagreements about certain details among LFG linguists. The writer of a textbook must wrestle with the problem of exactly what perspective to present. Naturally, my own preferences (and research) have influenced the presentation of the material in this book, but I hope that I have been fair to LFG as a whole. Where there is no consensus and I have chosen one particular approach, I have noted this.

I would like to thank people who commented on the manuscript or helped me in other ways: Farrell Ackerman, Paul Bennett, Joan Bresnan, Aaron Broadwell, Mary Dalrymple, Malka Rappaport Hovav, Tsipi Kuper-Blau, Helge Lødrup, Irit Meir, Rachel Nordlinger and Jane Simpson. I would also like to thank my wife Brandel, who looked at parts of the manuscript with an editor's eye and made helpful suggestions on wording. I would like to thank Dikran Karagueuzian, Chris Sosa, and Kim Lewis of CSLI Publications for all their help and support. Most importantly, I would like to thank all my students, past and present, who have taught me how to teach; I hope some of that has found its way into the book. Of course, none of these people is to blame for any remaining problems. My computer accepts full responsibility; it put the mistakes in when I wasn't looking.

Finally, I would like to thank my wife Brandel and my sons Eli, Yoni, Mati, and Gabi for putting up with my obsession to get this textbook finished. Thank you.

To the Student

Welcome!

As stated in the introduction, the purpose of this textbook is to teach the theory of syntax called Lexical-Functional Grammar. The concepts of the theory are built up piece-by-piece throughout the book. As a result, it is important to realize that the individual chapters are not self-contained. Each builds on what came before and the results are subject to revision in subsequent chapters. A number of chapters have less essential appendices at the end; these should be considered optional.

The end-of-chapter exercises are an inherent part of the material in the text. In some cases, they give the student a chance to practice a topic covered in the chapter; in other cases, they point to an addition to the analysis developed in the chapter.

Finally, a few words about bibliography. In general, the important bibliographic references are cited in the end-of-chapter "Additional Readings" section, rather than in the text of the chapter itself. For this reason, the sources of most of the important concepts in LFG will not be mentioned where the concepts themselves are introduced. There are two reasons for this. First, centralizing the bibliography makes it easier to find the references. Second, most of the concepts we will be discussing are widely accepted in one form or another in the LFG community; while it is important to cite the original source, it is also important to recognize that they have become the basis on which all work in LFG is based. Another thing to keep in mind is that the bibliography focuses on LFG material. In general, there are no references to work in other theoretical frameworks on the basic constructions of English, most of which is probably already familiar to you. This is not because they are not important, but simply because the purpose of this book is to focus on LFG analysis.

1

Welcome to Lexical-Functional Grammar

1.1 Introduction

Generative linguistics or **generative grammar**, a field of study that originates in the work of Noam Chomsky, is an attempt to discover the nature of the human language faculty, specifically of Universal Grammar (UG). The immediate goal of this approach to linguistics is to develop mathematical models of various aspects of human language. It is through the development of such models that formal claims about language can be expressed and tested.

Much work in generative linguistics has focused on modeling the syntactic component, the component of language that deals with the combination of words into phrases, clauses, and sentences. This is not coincidental. Syntax, unlike such components as phonetics/phonology, semantics, and pragmatics, is a system that is purely internal to language. It does not interface with nonlinguistic cognitive or motor systems. It thus plays a central role in organizing the entire linguistic system.

Perhaps the best-known model of syntax within the generative tradition is the one known as transformational syntax. This is a model that has been developed by Chomsky and his associates since the 1950s. Various developments of this model are known by names such as the Standard Theory, the Extended Standard Theory, the Revised Extended Standard Theory, Government/Binding theory, and the Minimalist Program. Despite all the changes, reflected by the different names that transformational theory has taken, certain assumptions underlie all transformational theories. Among these assumptions are the following:

- Syntactic representations are immediate-constituent structures, conventionally represented as trees. The configuration of constituent structure trees defines all crucial concepts of syntax (such as c-command).

- Grammatical functions (also called grammatical relations) such as "subject" and "object" are not elements of syntactic representation. These functions/relations are notions derived from the constituent structure, with the subject configurationally higher than the object, and in some sense "external" (outside the VP, outside the V′, etc.).
- A surface syntactic representation is the result of operations that take an existing constituent structure and change it into a similar but not identical constituent structure. These operations are called transformations, and are the source of the name "transformational grammar." While the details of transformations have changed over the years, transformational operations have included movement of constituents from one position in the tree to another, the insertion or merger of new elements into an existing structure, and the deletion or erasure of elements. In such a theory of grammar, the most salient feature is the set of consecutive representations of a grammatical sentence, often called a derivation. For this reason, a transformational approach to syntax can also be called a derivational approach.
- While the role of the lexicon in transformational grammar has changed drastically over the years, it tends to be seen as relatively limited. The lexicon is generally seen as little more than a repository of idiosyncratic information. This is less true of some versions of derivational theories than others.

While transformational theory represents the approach to syntax taken by most generativists, there are other approaches as well. These approaches are based on the rejection of some or all of these underlying assumptions of transformational syntax. This book is about one such alternative approach to syntax: Lexical-Functional Grammar, or LFG.

LFG rejects the assumptions of transformational theory, not its goals. The basic argument for the LFG approach to syntax is simply that certain transformationalist assumptions are incompatible with the search for a theory of Universal Grammar. LFG is therefore a variety of generative grammar, an alternative to transformational theory. In this book, we will occasionally compare the LFG approach with that of transformational theory, generally Government/Binding (GB) theory (Chomsky 1981, Chomsky 1986), and to a lesser extent the Minimalist Program (MP; Chomsky 1995).

LFG was developed in the mid-to-late 1970s, a period in which many different ideas about syntax were being explored. For example, this is the period in which many of the basic concepts of GB were developed. It was in

the late 1970s that Generalized Phrase Structure Grammar (GPSG; Gazdar, Klein, Pullum, and Sag 1984) was developed—a theory that has since evolved into Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994, Sag and Wasow 1999). And although it began in the early 1970s, this was also the formative period of the theory of Relational Grammar (Perlmutter, ed. 1983). Other attempts at modeling the syntactic component of grammar, many since forgotten, were also created then.

LFG developed in this period out of the work of two people. The first was Joan Bresnan, a syntactician and former student of Chomsky's, who had become concerned about psycholinguistic evidence that seemed to show that something was wrong with the concept of transformations. She started developing an alternative approach, which she called a Realistic Transformational Grammar, in which part of the work done by transformations in standard approaches was done in the lexicon instead (Bresnan 1978). The second person was Ronald M. Kaplan, a computational linguist/psycholinguist who was working on a parsing model called the Augmented Transition Network (ATN; Kaplan 1972). They realized that they were pushing in similar directions, and decided to collaborate. It is out of this collaboration that LFG was born, and to this day Bresnan and Kaplan are the key players in the LFG world.

To understand what LFG is and how it differs from transformational syntax, we will begin by examining the name of the theory: what is meant by "lexical," what is meant by "functional," and what is meant by "grammar"? As we discuss the literal meanings of the parts of the theory's names, we will also see related aspects of the theory.

1.2 "Lexical"

A lexical (or lexicalist) theory is one in which words and the lexicon play a major role. To some extent, this is true even in GB: the Projection Principle attributes certain syntactic patters to properties of words. In the Minimalist Program the derivation begins with a "numeration" (set) of lexical items, which are merged into the structure in the course of the derivation. Some versions of GB even recognize the existence of lexical operations, such as alterations to argument structures. These views in GB and MP depart from ideas in earlier transformational theory, and bring them closer to a lexicalist approach.

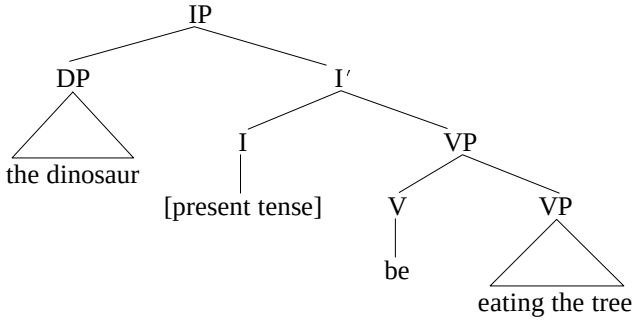
There are, however, some interesting ways in which words are not as important in GB and MP as (perhaps) they ought to be. One crucial way in which words are not important in transformational theory is that it does not,

in any of its incarnations, adopt the Principle of **Lexical Integrity**. We state the principle in preliminary form as (1).

(1) **Lexical Integrity Principle** (preliminary)
Words are the "atoms" out of which syntactic structure is built. Syntactic rules cannot create words or refer to the internal structures of words, and each terminal node (or "leaf" of the tree) is a word.

One example of a violation of the Lexical Integrity Principle in transformational theory can be seen in the standard GB analysis of V-to-I movement constructions. Consider the sentence in (2a). Its underlying (D-structure) representation is shown in (2b).[1]

(2) a. The dinosaur is eating the tree.
b.



Consider the status of the word *is*, one of the "atoms" out of which this sentence is built according to the Lexical Integrity Principle. Under the GB analysis it is not a part of this structure; the syntax builds it through V-to-I movement. It is the syntactic adjunction of the V *be* to the present tense feature in I that creates *is*. And what is present in D-structure under I is not even a word: it is an inflectional feature. This analysis, then, violates the Lexical Integrity Principle, both by virtue of building a word through a syntactic operation and because the syntactic structure is created out of things

[1]Details distinguishing this particular version of the analysis from more elaborated ones (VP internal subject, exploded infl, etc.) are irrelevant.

other than words.[2]

The Lexical Integrity Principle is a proposed principle for a theory of syntax. Like the A-over-A Principle of Chomsky (1973), the Projection Principle of Chomsky (1981), the Greed and Procrastinate of Chomsky (1995), or any other hypothesized principle of grammar, it is a potential step toward the goal of a constrained theory of grammar. All such principles are worthy of exploration; the way to explore such a principle is to examine what kinds of analyses are consistent with it, and to explore its explanatory potential. Inexplicably, while transformationalists have experimented with innumerable principles (and ultimately rejected most of them) they have generally[3] not considered the Lexical Integrity Principle. The ultimate test of any proposed principle of language is its ability to lead to well-motivated analyses of linguistic facts.

The resistance that transformational theory has shown to the Lexical Integrity Principle is all the more surprising because it carries a fair amount of plausibility. The essential claim behind the Lexical Integrity Principle is that syntax cannot see the internal structure of words. It has long been noticed that word structure is different from phrase and sentence structure. This is the reason that while *semantics* and *phonology* refer indifferently to meaning/sound structure both above and below the level of the word, linguists have usually distinguished between structure above the level of the word (*syntax*) and structure below the level of the word (*morphology*). There are many ways to show that word structure is different from phrase and sentence structure. We will mention two here. First, free constituent order in syntax is common cross-linguistically; many languages lack fixed order of the kind that one finds in English. In morphology, on the other hand, order is always fixed. There is no such thing as free morpheme order. Even languages with wildly free word order, such as the Pama-Nyungan (Australian) language Warlpiri (Simpson 1991), have a fixed order of morphemes within the word. Second,

---

[2]In the Minimalist Program the syntax does not build words: *is* is taken from the lexicon. In this respect, it is more consistent with the Lexical Integrity Principle than older versions of transformational syntactic theory. However, as with the GB and earlier accounts, abstract inflectional features still occupy their own structural positions in the Minimalist Program. In addition, feature checking requires the syntax to analyze the internal structure of the inflected verb.

[3]A notable exception is Di Sciullo and Williams (1987).

syntactic and morphological patterns can differ within the same language. For example, note the difference in English in the positioning of head and complement between syntax and morphology.

(3) a. $[_{V'}\ \text{eat}_{\text{head}}\ \text{tomatoes}_{\text{complement}}]$
b. $[_{N}\ \text{tomato}_{\text{complement}}\ \text{eater}_{\text{head}}]$

At the phrasal level, heads precede their complements, while at the level of the word heads follow their complements. If word structure is distinct from phrase and sentence structure, it stands to reason that the component of the grammar responsible for the latter is distinct from the one responsible for the former. This is essentially what the Lexical Integrity Principle says. Consequently, the Lexical Integrity Principle is a plausible component of a theory of syntax.

A theory that respects (some version of) the Lexical Integrity Principle can be said to be a lexicalist theory. This is a theory in which words play a central role in the syntax: syntactic structures are composed of words. It is also a theory in which the lexicon will play a central role, since it is the component in which words are created. LFG is a lexicalist theory in this sense.

Marantz (1997) purports to provide evidence against lexicalism, going so far as to declare lexicalism "dead, deceased, demised, no more, passed on". However, nowhere does he actually address the heart of lexicalism: the Lexical Integrity Principle and the idea that structure above the level of the word differs from structure below the level of the word. Instead, Marantz argues, on the basis of idioms, that words are not unique in (sometimes) having idiosyncratic semantics. Therefore, form-meaning pairs cannot be isolated in the word. Furthermore, Marantz argues that idioms cannot be listed in the lexicon because idiom chunks cannot be Agents. Under Marantz's assumptions, the thematic role Agent is "projected" in the syntax by a functional category rather than being a lexical property of the verb. Therefore, Marantz views the conditions on possible idiomatic meaning as syntactic rather than lexical. However, without the assumption that a lexically unjustified category "projects" the Agent role, the conclusion does not follow. The true generalization about idioms is slightly different in any case; as we will discuss in Chapter 4, it seems to be based on a hierarchy of thematic roles. The issues that Marantz raises are irrelevant to the question of whether syntactic theory should adopt the Lexical Integrity Principle.

However, lexicalism goes beyond the Lexical Integrity Principle.

Consider the passive construction. There have been many analyses of passivization proposed in the long history of transformational theory. Some, such as the incorporation analysis of Baker (1988), see the passive morpheme as a separate syntactic constituent that combines syntactically with the verb. Such analyses clearly violate the Lexical Integrity Principle in the same ways as V-to-I movement: the atoms of syntax are not words, and the syntax builds words.[4] However, there is another transformational analysis, outlined in Chomsky (1981), which treats the passive morpheme as a signal of a lexical change in the verb's argument structure (θ grid in GB terminology). The passive morpheme causes the subject argument to be suppressed. This results in a lexical argument structure with an object argument but no subject argument. As a result of a principle of GB called Burzio's Generalization, the verb also loses its ability to "assign Case." In the syntax, the object argument becomes the subject by undergoing NP movement, a movement triggered by the object not getting Case in situ. The NP movement is thus an indirect result of the lexical change in argument structure. This can be shown informally by the following chart.

(4) One GB analysis of passive

	〈 subject , object 〉
lexical change	↓
	〈 ∅ , object 〉
syntactic change	↓
	〈 ∅ , subject 〉

This is an essentially lexical analysis of the passive, since the syntactic change is triggered by the lexical change. However, the realization of the active object argument as subject is still inexplicably attributed to a derivational syntactic process. From the perspective of lexicalist syntax, there is a clear alternative, in which there is no syntactic derivation. (Again, this is an informal demonstration.)

[4]For arguments against the incorporation analysis of passivization from a GB perspective, and in favor of the lexicalist GB approach, see Falk (1992).

(5) Potential lexicalist analysis of passive

		〈 subject	, object	〉
lexical change		↓	↓	
		〈 ∅	, subject	〉

Such an account is simpler, in that it unifies the two changes associated with the passive construction.

There is also evidence that the lexicalist account is superior to the mixed lexical-syntactic GB approach, as discussed by Bresnan (1982a; 1995b). One such piece of evidence, a very strong one, comes from the fact that passivization feeds other lexical processes. For example, participles with no obligatory nonsubject arguments can be morphologically converted into adjectives through zero-derivation. In the resulting adjectival passive, the subject of the passivized verb is the subject of the adjective.

(6)	a.	The present was given (to the zookeeper).	(Theme)
	b.	the ungiven present	(Theme)
	c.	*The zookeeper was given. (cf. The zookeeper was given a present.)	(Goal)
	d.	*the ungiven zookeeper	(Goal)

(7)	a.	The T-rex was fed.(a Triceratops sandwich)	(Goal)
	b.	an unfed T-rex	(Goal)
	c.	*A sandwich was fed. (cf. A sandwich was fed to the T-rex.)	(Theme)
	d.	*an unfed sandwich	(Theme)

The simplest description of such facts is that the only change is the change of category; there is no change of grammatical functions as a result of the conversion. The appropriate argument is the subject of the adjectival participle because it is the subject of the verbal participle. A transformational account would have to attribute the Theme argument becoming the subject of the adjectival passive to a different process than in the verbal passive, because lexically the Theme is the object of the passive verb.

The preceding discussion shows that a lexicalist theory will have fewer transformations and shorter derivations than a typical transformational theory. The ultimate limit that one can reach is no transformations and no derivation. In fact, lexicalist argumentation generally leads to the conclusion that syntax

is not derivational.[5] For this reason, the term "lexicalist" is often synonymous with "nontransformational" or "nonderivational." LFG is also lexicalist in this sense.

Nonderivational theories are more plausible as psychological and computational models of human language than derivational theories. Transformational theories are, by the nature of what a transformation is, nonlocal theories of syntax. However, it is clear that human language processing is local. Consider the VPs in (8).

(8) a. hears herself
b. *hears myself

Even without the larger context of a full clause, it is clear that (8a) is grammatical and (8b) is not. This is determined from information internal to the VP; the larger IP (or S) is clearly unnecessary. In derivational theories, agreement is a result of feature copying/checking between I (or T or AGR_S or AUX) and its specifier. Thus, although there is no larger structure in these examples, transformational theories must hypothesize one. The grammaticality judgments cannot be determined purely from properties internal to the VP. Theories based on the notion that processing is local are thus more realistic. Further examples of the locality of processing can be found in Bresnan and Kaplan (1982: xlv).

A consequence of taking a nonderivational approach to syntax is that syntactic structures are built **monotonically**; that is to say, information can be added but it cannot be changed. Transformations are, by definition, change of information. Monotonicity is also a computationally plausible constraint on syntax.

Nonderivational theories are also **constraint-based**. Grammaticality cannot be dependent on properties of derivations, since there are no derivations. What determines grammaticality is the satisfaction of static simultaneous constraints. Of course, transformational theories are partially constraint-based as well (GB's Θ Criterion, Case Filter, Binding Principles; MP's Principle of Full Interpretation), but much of the determination of grammaticality is the result of the well- or ill-formedness of the derivation.

So besides being a theory in which the lexicon plays a major role, LFG is a nonderivational theory, one that has no D-structure/S-structure distinc-

[5]An early example of this is Brame (1976).

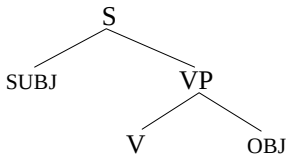
tion. There is just one level of constituent structure. LFG calls this **c-structure**.

1.3 "Functional"

1.3.1 Grammatical Functions

The word *functional* means different things to different people in linguistics. What it means in LFG is **grammatical functions**, notions like subject and object (also called grammatical relations). The role of grammatical functions has long been a matter of dispute in generative syntax. The standard transformationalist view has been that grammatical functions are universally defined on the basis of c-structure configurations, roughly (9).[6]

(9)



Under such a view, grammatical functions are not part of the basic vocabulary of syntax. Syntax deals with c-structural configurations only. Whatever properties grammatical functions are thought to have are derived from the configurations that define them. For example, the fact that only subjects can be controlled is attributed to the unique structural properties of the subject position (in GB, specifically the fact that V does not "govern" the subject position).

However, this view has been challenged. The basic idea behind the alternative is that a major facet of syntax is the fact that each element is there because it has a function (or bears a relation to the clause). Thus, grammatical functions (or grammatical relations) ought to be part of the vocabulary of syntactic theory. It is interesting that while GB claims to reject this view, there are certain relational features to the architecture of the theory. For example, the notion "government" as understood in GB is basically a relational notion: a governee bears some grammatical relation to the

[6]There have been several variants of this, depending the specifics of the theory of structure and categories. The reader should feel free to substitute the appropriate category labels and intermediate nodes.

governor. Similarly, the "complete functional complex" of Chomsky's (1986) Binding Theory is a functionally defined unit. Finally, "Case" as generally used in GB and MP is largely a cover term for grammatical functions.

The first challenge to the c-structural approach to grammatical functions came from Paul Postal and David Perlmutter in a series of lectures at the Summer Institute of the Linguistic Society of America in 1974. These lectures developed into the theory of Relational Grammar (Perlmutter, ed 1983), a theory based on the idea that the syntactic representation of a sentence is a network of grammatical relations, and that syntactic rules are expressed in terms of grammatical relations.

The LFG claim is that grammatical functions are elements of syntactic representation, but of a kind of syntactic representation that exists in parallel to c-structure. This level of representation is not a tree structure, like c-structure. Instead, it is based on the idea that grammatical functions are like features, and the elements that have specific functions are the values of these feature-like functions. The representation of grammatical functions also includes features of a more conventional nature. It is called **f-structure**, where (because of a happy accident of English) one can think of *f* as standing for either *function* or *feature*. (The standard interpretation is that *f-structure* stands for *functional structure*.)

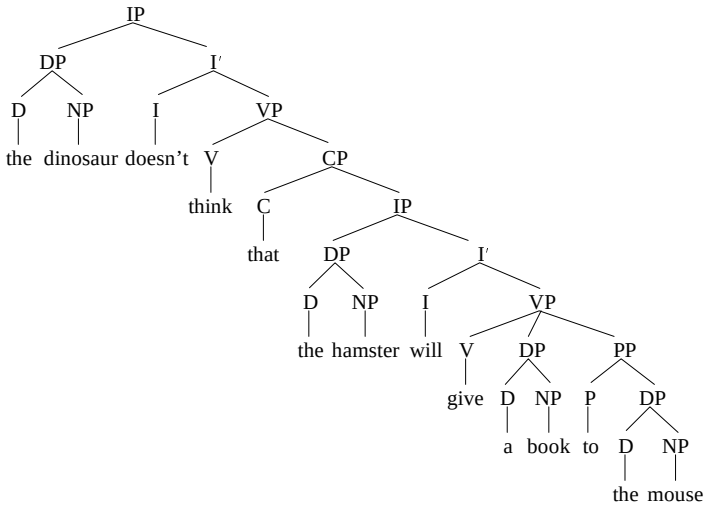
Unlike c-structures, f-structures are not familiar from derivational theories of syntax. We will first examine what an f-structure looks like, and then we will discuss the motivations for hypothesizing f-structure and the consequences for the general architecture of linguistic theory.

1.3.2 F-structure

To make the notion of f-structure concrete, let us consider a sentence and its c-structure and f-structure.

(10) a. The dinosaur doesn't think that the hamster will give a book to the mouse.

b. c-structure



c. f-structure:

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{‘dinosaur’} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{NEG} & + \\ \text{PRED} & \text{‘think } \langle \text{SUBJ, COMP} \rangle\text{’} \\ \text{COMP} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{‘hamster’} \end{bmatrix} \\ \text{TENSE} & \text{FUTURE} \\ \text{PRED} & \text{‘give } \langle \text{SUBJ, OBJ, OBL}_{\text{Goal}} \text{ OBJ} \rangle\text{’} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & - \\ \text{PRED} & \text{‘book’} \end{bmatrix} \\ \text{OBL}_{\text{Goal}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{Goal}} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{‘mouse’} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The f-structure is what is sometimes called an **attribute-value matrix** (or **AVM**). An attribute is a feature or function name; unlike the more familiar notation for features in phonology, the attribute name precedes the value.

Thus, the phonological feature (11a) would appear as (11b) in an AVM.

(11) a. [+voiced]
b. [VOICED +]

Let us take a closer look at the f-structure. It contains five attribute names: SUBJ, TENSE, NEG, PRED, and COMP. To the right of each attribute name is its value. Three of the attributes, TENSE, NEG, and PRED, are features; they have simple values. The other two attributes, SUBJ and COMP, are functions; their values are smaller f-structures (AVMs) within the larger f-structure.[7] Let us consider these one-by-one.

- The feature TENSE is an inflectional feature, like PERS(on), NUM(ber), CASE, GEND(er), etc. Such features are represented in f-structure in LFG, not in c-structure.

- The feature NEG is also an inflectional feature. Note that both [TENSE PRES] and [NEG +] are contributed by the word *doesn't*.

- The feature PRED is very important. The idea behind it is that the existence of meaningful items is relevant to the syntax. Of course, the meaning itself is not part of syntactic representation, but certain aspects of meaning are. First, the syntax needs to be able to distinguish between meaningful elements and dummy (or expletive) elements. The PRED feature serves to represent meaningfulness; its value is represented conventionally as the word itself in single quotation marks. For pronouns, which are meaningful but get their reference elsewhere in the sentence or discourse, the special PRED value 'PRO' is used.[8] In this example, we also see another kind of syntactic relevance of meaning: the verb *think* takes two arguments ("assigns two θ roles" in GB/MP terminology): one bearing the function SUBJ, and the other bearing the function COMP. A PRED value with a specification of arguments is

[7]Note that the term "f-structure" is thus ambiguous: it can refer either to the entire representation of the sentence or to some AVM within the representation.

[8]The f-structure [PRED 'PRO'] should not be confused with the PRO of transformationalist theories.

sometimes called a **lexical form**. It is ultimately derived from the verb's argument structure (**a-structure**). The two functions that appear as attributes in the f-structure are the same ones subcategorized for by the verb.

- The attribute SUBJ is a grammatical function, corresponding roughly to the traditional intuitive notion "subject" (just as N corresponds roughly to the traditional "noun"). Its value is a subsidiary f-structure consisting of the features DEF(initeness) and PRED and their values. The lexical form of *think* specifies that the value of the SUBJ function fills the first argument position of the verb.

- The function COMP(lement) is the grammatical function of clausal complements. It fills the second argument position of *think*, and its value consists of the attributes SUBJ, TENSE, PRED, OBJ, and OBL_{Goal}.

Most of the rest of f-structure (10b) should be straightforward. What does require some explanation is the final argument in the lexical form of *give*, and the representation of the PP that fills this argument position. The PP *to the mouse* consists of a head P *to* and its OBJ *the mouse*. The PP functions as an **oblique** argument: an argument whose "role" is identified morphologically (by a preposition in English). "Role" in this context generally means thematic role, although sometimes the prepositional marking is idiosyncratic. The preposition is similar to semantic Case (in fact, many languages use Cases in this context). For the last argument of *give*, the preposition *to* marks the DP as bearing the thematic role of Goal. In LFG, the oblique functions are treated as a class of grammatical functions OBL_{θ}; in the present case, OBL_{Goal}. Since the preposition *to* is what identifies the argument as an OBL_{Goal}, its prepositional Case (PCASE) feature also has the value OBL_{Goal}. Finally, it is not the PP itself (which has the function OBL_{Goal}) that is the final argument of *give*; instead, it is the OBJ within the PP. For this reason, the lexical form of *give* specifies a path through the f-structure, OBL_{Goal} OBJ, as the syntactic realization of the argument.

One additional clarification is in order concerning f-structures. We have seen that meaningfulness is represented by the feature PRED. Of course, sometimes there are meaningless elements in syntax. Such elements include expletives and idiom chunks, as in:

(12) a. *It* seems that this book will be interesting.
b. The teachers kept *tabs* on the students.

Naturally, these items will not have PRED features. In fact, it is crucial that they not be meaningful elements, i.e. that they lack PRED features. Instead, they have a feature, called FORM, that individuates them and allows them to be selected for. The f-structures associated with *it* and *tabs* are:

(13) a. $\begin{bmatrix} \text{FORM} & \text{it} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix}$

b. $\begin{bmatrix} \text{FORM} & \text{tabs} \\ \text{NUM} & \text{PL} \end{bmatrix}$

The lexical forms of these uses of *seem* and *keep* will indicate that they have nonthematic arguments. Since the argument structure is indicated inside angle brackets, a nonthematic argument can be placed outside the angle brackets:[9] [10]

(14) a. 'seem ⟨COMP⟩ SUBJ'
b. 'keep-tabs-on ⟨SUBJ, OBL_{on} OBJ⟩ OBJ'

In addition, the lexical entries of these verbs will require FORM feature values for their nonthematic arguments. The f-structure of (12b) is:

[9]In some early LFG papers, including many in Bresnan, ed. (1982), nonthematic arguments were omitted from the verb's lexical form. The notation that has been adopted since, and is used here, formalizes the fact that they are selected for by the verb, even though they are not thematic arguments.

[10]Note that "OBL_{on} OBJ" in the lexical form of *keep tabs on* is a single argument, not two arguments.

(15)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`teacher'} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{`keep-tabs-on} \langle \text{SUBJ}, \text{OBL}_{\text{on}}\ \text{OBJ} \rangle\ \text{OBJ'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{tabs} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{OBL}_{\text{on}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{on}} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`student'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

It is important to note that f-structure has a completely different geometry and completely different properties from c-structure. C-structure is built out of NPs, VPs, etc., and represents membership in hierarchically larger and larger groupings of elements. F-structure is composed of attributes (features and functions) and their values. It is not composed of a hierarchical arrangement of categories, and lacks a representation for certain elements of c-structure (such as the VP constituent). That is to say, although some of the information represented in f-structure resembles "underlying" structure information in transformational theory, the levels cannot be related to each other by movement. They are completely different structures.

1.3.3 Motivation

We turn now to the motivation for "functional." That is to say: why hypothesize f-structure in addition to c-structure? We will answer this question from two different perspectives. First, we will discuss the motivation for representing grammatical features at a level distinct from c-structure. We will then address the more central question concerning the role that LFG gives grammatical functions.

We begin with features. The essential observation behind the LFG approach is that features cannot always be associated with the c-structure constituents that they describe. Consider the following sentence:

(16) The deer are in the forest.

The features of the SUBJ of this sentence are:

(17)
$$\begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`deer'} \\ \text{NUM} & \text{PL} \end{bmatrix}$$

However, these features come from two different elements of the c-structure. The DP *the deer* is unspecified for number, as evidenced by the sentence *The deer is in the forest*. The feature structure of the DP is:

(18)
$$\begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`deer'} \end{bmatrix}$$

By virtue of its position in the c-structure tree and English-specific rules relating structure and function, *the deer* will appear in the f-structure of our sample sentence as:

(19)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`deer'} \end{bmatrix} \end{bmatrix}$$

The [NUM PL] feature of the SUBJ comes from lexical entry of *are*. The features of *are* are:[11]

(20)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`be } \langle \ldots \rangle \text{'} \end{bmatrix}$$

That is to say, *are* is a present tense form of *be* with a plural subject.[12]

(19) and (20) are partial f-structures for the sentence. However, since we are building an f-structure for a single sentence, the SUBJ features from the two sources have to come together. The resulting f-structure is (21).

[11]We will ignore the details of the lexical form.

[12]*Are* can also have a singular *you* as subject; we ignore this complication here.

(21)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`deer'} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`be} \langle \ldots \rangle\text{'} \\ \vdots & \vdots \end{bmatrix}$$

This merging of feature structures is called **unification**. Unification is a central concept of feature-based approaches to syntax, including LFG, but also certain other frameworks like HPSG. The point here is that unification is part of the reason to consider f-structure an independent level of syntactic representation. It allows us to represent together features that belong to a single conceptual part of the syntactic structure of the sentence even if the features come from several places in the actual syntactic structure. A theory like LFG, in which grammatical features are represented independently of constituent structure, does not need mechanisms of feature percolation, feature inheritance, etc.

A side-effect of unification is that it accounts automatically for the ungrammaticality of a sentence like:

(22) *The lion are in the forest.

That is to say, agreement is an automatic result of unification. Unlike *deer*, the noun *lion* is inherently singular. It therefore has the lexical feature [NUM SG]. On the other hand, as we have seen, *are* includes the lexical feature [SUBJ [NUM PL]]. Since *the lion* is in the structural position associated with the function SUBJ, the [NUM SG] feature of *the lion* and the [SUBJ [NUM PL]] feature of *are* must unify. However, the result is that the SUBJ NUM feature is inconsistent with itself: it must be simultaneously singular and plural. Since this is impossible, the sentence is ungrammatical. Thus, unlike transformational theories, LFG does not need any special mechanisms like cosuperscripting or feature checking or SPEC-head relations to enforce agreement. Feature checking is part of unification. More generally, much of what is modeled by movement in transformational theory is modeled by unification in LFG. Unlike movement-based theories, a unification-based theory does not need to hypothesize structural arrangements of elements which differ from that which is accessible from the visible, superficial form of a sentence.

The primary justification for f-structure relates to the status of grammati-

cal functions. As discussed earlier, transformational grammar considers grammatical functions to be a derivative concept that can be defined in terms of the c-structure configurations in (9) above. LFG denies this, and claims that grammatical functions are an independent concept. Such a claim, of course, needs to be proven. The way to prove it is to show that there are languages in which concepts like subject and object are relevant for which the c-structure configuration in (9) cannot be supported. It is to this that we now turn.

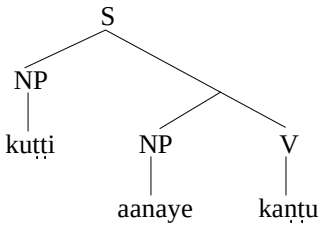
First, however, a caveat. We are interpreting transformationalist statements about constituency as an empirical claim about c-structure. However, the arguments often given for such structures are not based on standard constituency tests such as distribution and order, but on what LFG claims to be function-related phenomena such as anaphora. It is thus possible to view a configuration such as (9) as nothing more than an idiosyncratic way of representing grammatical functions. If putting a constituent in the [SPEC, IP] position is nothing more than a notation for SUBJ, then constituency tests are irrelevant. However, it is a rather strange representation for grammatical functions, and would leave transformational theory with no theory of c-structure.

If, on the other hand, the syntactic structure of transformational theory really is a c-structure, then it must be tested empirically. In fact, there are languages that cast doubt on this kind of approach. In the first place, there are languages that have free constituent order. Japanese is one such language; we will use the example of the Dravidian language Malayalam (Mohanan 1982). Note the possible orders of the words in the sentence 'The child saw the elephant'.

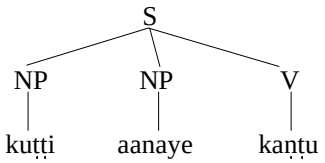
(23)	a.	Kuṭṭi	aanaye	kaṇṭu. (SOV)
		child.NOM	elephant.ACC	saw
	b.	Aanaye kuṭṭi kaṇṭu.		(OSV)
	c.	Aanaye kaṇṭu kuṭṭi.		(OVS)
	d.	Kaṇṭu aanaye kuṭṭi.		(VOS)
	e.	Kaṇṭu kuṭṭi aanaye.		(VSO)
	f.	Kuṭṭi kaṇṭu aanaye.		(SVO)
	'The child saw the elephant.'			

Consider two possible hypotheses as to the structure for such a sentence.

(24) a.



b.



The (a) structure, in some variant, is the transformationalist view, which places the SUBJ in a structurally higher position than the OBJ. Innumerable movement rules would be required to derive all the surface word orders from such a structure. On the other hand, with a flatter structure, as in (b), all that one has to say is that the ordering is free. Since all three constituents are sisters, all of the possible orderings would result.

Of course, the argument in the preceding paragraph can be countered. In a theory with unconstrained movement, any word order can be derived from any D-structure. And if SUBJ-OBJ asymmetries in binding or quantifier scope are taken axiomatically to mean a relation of asymmetric c-command, the (a) structure must be the structure of the sentence. However, the facts of Malayalam present no independent evidence for treating the verb and OBJ as forming a constituent that excludes the SUBJ, and the description of the language is simpler if we assume no such constituent. But if there is no such constituent in Malayalam, SUBJ and OBJ cannot be universally defined in terms of c-structure.

Even more strikingly, there are languages that present positive evidence against a VP constituent. This evidence comes from languages like the Pama-Nyungan Australian language Warlpiri (Simpson 1991) and the non-Pama-Nyungan Australian language Wambaya (Nordlinger 1998). (The examples here come from Wambaya.) In these languages, the auxiliary (infl) occurs in second position. One constituent must precede the infl and the rest follow. With the single exception of the auxiliary, constituent order is completely free. (In these examples, the infl is italicized.)

(25) a. Dawu *gin-* *a* alaji janyi- ni.
bite 3SG.M.ERG- PST boy.ABS dog- ERG
b. Alaji *gin-a* dawu janyi-ni.
c. Alaji *gin-a* janyi-ni dawu.
d. Dawu *gin-a* janyi-ni alaji.
e. Janyi-ni *gin-a* alaji dawu.
f. Janyi-ni *gin-a* dawu alaji.
'The dog bit the boy.'

Multiword constituents can precede infl.

(26) [Naniyawulu nagawulu baraj- bulu]
that.DU.ABS female.DU.ABS old.person- DU.ABS
wurlu- *n* duwa.
3DU.NPST- PROG get.up
'The two old women are getting up.'

This gives us a test for constituenthood in Wambaya: if there is a VP constituent, it should be able to precede infl. Strikingly, it cannot.

(27) a. *[Daguma janji] *ng-* *a* ngawurniji.
hit dog.ABS 1SG.ERG- PST 1SG.ERG
b. *[Janji daguma] *ng- a* ngawurniji.
'I hit the dog.'

This suggests that it is not enough to account for the freedom of constituent ordering in Wambaya by allowing constituents to be moved out of the Wambaya VP; Wambaya does not seem to have a VP! But Wambaya can be shown to have SUBJs and OBJs, just like any other language. SUBJs and OBJs are Case-marked differently and are crossreferenced by different pronominal (agreement) markers on the infl. As in many languages, only SUBJs can serve as the antecedents of reflexives, and only SUBJs can be controlled in nonfinite subordinate clauses. Wambaya also has a switch-reference system in which certain subordinate clauses are marked for whether their SUBJ is the same as or different from the main clause SUBJ. As Nordlinger (1998) shows in detail, attempts that have been made to account for languages like Wambaya within a c-structural/derivational approach have all failed to account for the facts of these languages. In Wambaya , then, we have an example of a language in which SUBJ and OBJ are rather similar to the same concept in more familar

languages, but cannot be distinguished in terms of being part of a VP constituent. This requires some independent representation of grammatical functions.

The conclusion is that while the structure in (9) does characterize SUBJ and OBJ in English, it does not do so universally. This means that grammatical functions cannot be universally dependent on constituent structure position. Languages like English, in which a VP constituent distinguishes SUBJ from OBJ, can be called **configurational** languages, while languages like Japanese, Malayalam, Warlpiri, and Wambaya can be called **nonconfigurational**. The existence of nonconfigurational languages provides crucial evidence for the independence of grammatical functions from c-structure, and thus for f-structure.

1.3.4 Consequences

The conclusion that there is a level of f-structure distinct from c-structure has interesting consequences for an overall theory of the nature of language in general and the nature of syntax in particular. In this section we will explore this.

A sentence is an expression of several different types of linguistic information. We can identify at least the following:

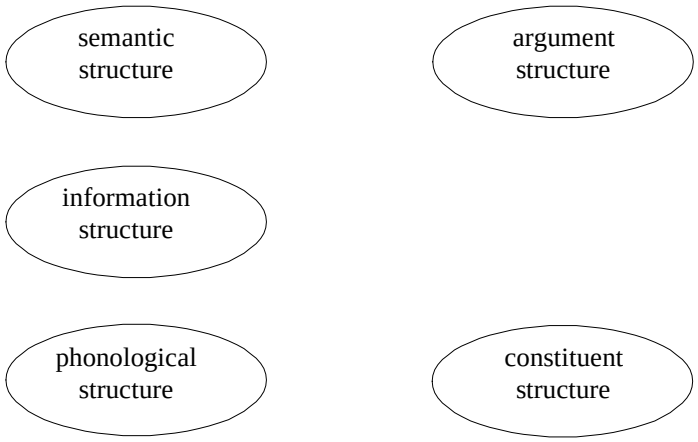
information/discourse/pragmatics
meaning/semantics
argument structure/thematic roles
syntactic constituent structure
sounds (phonology/phonetics)

There are two ways that this can be conceptualized.

The approach taken by transformational theory has generally been that (with the possible exception of phonology) these are different aspects of the same kind of structure. Syntactic constituent structure is taken to be the basic form of structure and the other kinds of information are expressed in terms of it. For example, the thematic role Agent is represented by a chain whose foot is in an "external" position ([SPEC, IP] or [SPEC, VP], depending on the exact version of the theory). Such a theory has a certain conceptual simplicity: all rules of language are stated over the same primitives, and all properties of a single element in the sentence can be determined from a single kind of structure.

There is an alternative approach, which sees each of these kinds of

information as part of a distinct kind of structure. Under this alternative, information structure, semantic structure, argument (or thematic) structure, syntactic constituent structure, and phonological/phonetic structure[13] are distinct subsystems of language, each with its own primitives and its own internal rules of organization. This can be schematized as follows:



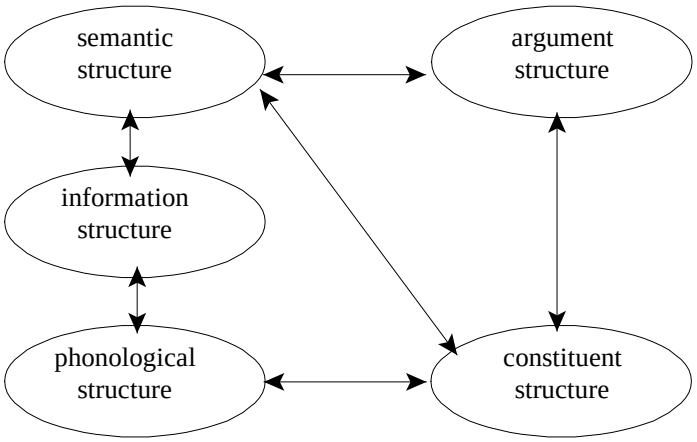
These levels of representation all exist in parallel; no one is prior to any of the others. A theory of language that is based on such a model can be said to have a **parallel** architecture.

However, this is not enough for a theory with parallel architecture. Besides different kinds of primitives and rules for each dimension of linguistic structure, a system of **correspondence** is required to map between the levels.[14]

---

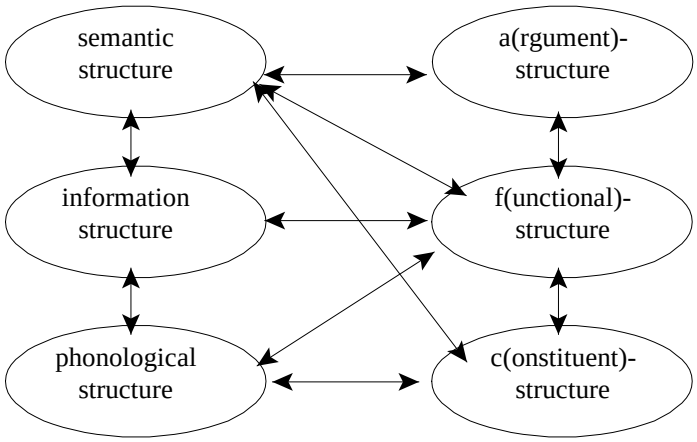
[13]This is an oversimplification. Any specific instantiation of this approach may draw the borders differently, depending on what empirical evidence is found. For example, it is possible that phonological structure and phonetic structure are distinct, or that thematic structure and semantic structure are the same.

[14]This diagram is for purposes of illustration only. I will not argue for any specific aspect of this diagram. In particular, exactly which levels are directly related by correspondence rules needs to be determined independently.



Such a theory therefore needs correspondence functions, or "projection" functions. LFG is said to have a **projection architecture** connecting the different levels of representation. Determining all the properties of a particular element in a modular system requires examining the corresponding item (or items) in each of the projections.

The conclusion reached in the previous section that c-structure and f-structure are formally different representations with their own primitives and their own organization makes sense given the concept of parallel, correspondence-based architecture. It simply adds an additional level: f-structure.



The three structures on the right side of this diagram are the syntactic levels, which are the ones on which LFG focuses.[15] However, by virtue of its adoption of a parallel architecture, other levels can be hypothesized for an LFG grammar.

The LFG assumption of parallel architecture, and its claim that grammatical functions and features are a kind of linguistic information distinct from constituency, provide an elegant solution for a potential problem with the Lexical Integrity Principle. The problem is apparent when we consider the following.

(28) a. The dinosaur *ate* the tree.
b. The dinosaur *did eat* the tree.

(29) a. My dinosaur is *hungrier* than yours.
b. My dinosaur is *more hungry* than yours.

In the (a) examples, a single word is italicized, while in the (b) sentences a two-word sequence is highlighted which seems to serve the same function as the single word in (a). In a nonlexicalist framework such facts can be accounted for by treating *ate* as a combination of *did* and *eat*, and by treating

[15]although there has naturally been work on other aspects of language in LFG, particularly semantics.

hungrier as a combination of *more* and *hungry*. Details aside, this kind of analysis has been standard in transformational syntax since Chomsky (1955). The challenge for a lexicalist theory is how to express such relationships within the confines of Lexical Integrity, which does not allow words to be built in the syntax.

A closer look shows that cases like these pose no problem for LFG. The Lexical Integrity Principle designates words as the atoms out of which "syntactic structure" is built. However, as we have seen, there are two levels of "syntactic structure" in LFG: c-structure and f-structure. The one that is built out of words is c-structure; f-structure consists of abstract attributes (features and functions) and their values. We can state the Lexical Integrity Principle as follows.

(30) **Lexical Integrity Principle**
Morphologically complete words are leaves of the c-structure tree and each leaf corresponds to one and only one c-structure node.

However, the equivalence of the (a) and (b) sentences above is in grammatical features. The verb form *ate* includes within it both the lexical properties of *eat* (the PRED feature, in LFG terms) and the past tense feature. With *did eat*, these two features are separated. Since features are involved, the level of representation at which *eat* and *did eat* are equivalent is f-structure. The f-structure representation of the two sentences in (28) is:

(31)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{`eat} \langle \text{SUBJ, OBJ} \rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`tree'} \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}$$

Lexical integrity as understood by LFG is thus limited to c-structure. It is a limited sort of lexical integrity, which is better able to deal with featural equivalence of words and word sequences than an approach in which all aspects of the internal structure of a word is invisible to the syntax. LFG's version of the Lexical Integrity Principle balances the similarities and the differences between words and phrases.

1.4 "Grammar"

Like transformational theory, LFG is a variety of generative grammar, an approach to the study of language that has its origins in the work of Noam Chomsky. Generative grammar has several central aims.

- The discovery of linguistic universals in an attempt to determine the nature of Universal Grammar (UG):

  > The main task of linguistic theory must be to develop an account of linguistic universals that, on the one hand, will not be falsified by the actual diversity of languages and, on the other, will be sufficiently rich and explicit to account for the rapidity and uniformity of language learning, and the remarkable complexity and range of the generative grammars that are the product of language learning. (Chomsky 1965: 27–28)

- the discovery of a psychologically real model of linguistic competence that can be incorporated into a performance model, and the study of the mathematical properties of the competence model

  > No doubt, a reasonable model of language use will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer's knowledge of the language; but this generative grammar does not, in itself, prescribe the character or functioning of a perceptual model or a model of speech production.… To my knowledge, the only concrete results that have been achieved and the only clear suggestions that have been put forth concerning the theory of performance … have come from studies of performance models that incorporate generative grammars of specific kinds… (Chomsky 1965: 9,10)
  >
  > In brief, mathematical study of formal properties of grammars is, very likely, an area of linguistics of great potential. It has already [1965] provided some insights into questions of empirical interest and will perhaps some day provide much deeper insights. (Chomsky 1965: 62)

- the formal explicit statement of the machinery of the theory of language and rules of specific languages

  > We can determine the adequacy of a linguistic theory by developing rigorously and precisely the form of the grammar corresponding to the set of levels contained within this theory, and then investigating the possibility of constructing simple and revealing grammars of this form

> for natural languages. (Chomsky 1957: 11)
>
> If the grammar is, furthermore, perfectly explicit—in other words, if it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution—we may (somewhat redundantly) call it a *generative grammar*. [italics original]. (Chomsky 1965: 4)

As the above citations show, these are all aims that one finds expressed very explicitly in Chomsky's early writings laying out the generative approach. Oddly, one can very seriously question the degree to which Chomsky's work over the past two decades still has these as its goals. For example, recent transformational theory has tended to ignore counterexamples to some of its basic claims, often taking refuge behind an artificial distinction between "core grammar" and "periphery", as in the following quote from Chomsky (1981: 8).

> [E]ach actual "language" will incorporate a periphery of borrowings, historical residues, inventions, and so on, which we can hardly expect to—and indeed would not want to—incorporate within a principled theory of UG.…What a particular person has inside his head is an artifact resulting from the interplay of many idiosyncratic factors, as contrasted with the more significant reality [sic] of UG (an element of shared biological endowment) and core grammar (one of the systems derived by fixing the parameters of UG in one of the permitted ways).

The continued inability to come to grips with the challenge posed by nonconfigurational languages illustrates this as well. Recent derivational approaches have also not taken facts about linguistic performance and mathematical properties of grammars to be linguistic evidence. They also have eschewed formal statements of rules and principles of the kind that was typical of earlier derivational theories, in which phrase structure rules and transformations were stated in painstaking detail, and is still typical of LFG and other nonderivational theories. In this sense, LFG may be truer to the goals of generative grammar than Government/Binding theory and the Minimalist Program.

The search for linguistic universals must be based on research into typologically different languages. This is implicit in Chomsky's statement that the theory of universals should be one that "will not be falsified by the actual diversity of languages." Consequently, generative linguistics can only be properly carried out in conjunction with typological work. As we have

already seen, there are typological problems with GB/MP, such as the assumption that grammatical functions are uniformly represented in constituent structure. LFG, on the other hand, has always involved the description of typologically disparate languages, without preconceptions about how languages might differ. In the words of Austin and Bresnan (1996: 263), "theoretical economy and explanatory elegance are unreliable guides to truth." As a result, LFG is a typologically more plausible model of language, in which the constraints on syntax are derived from a broader understanding of linguistic diversity. Ultimately, this approach is more likely to provide true explanations for linguistic phenomena.

The development of LFG has involved the collaboration of people working on linguistic description, computation, and psycholinguistics. As mentioned at the outset, LFG began with the collaboration of a theoretical/descriptive linguist and a computational linguist/psycholinguist. Bresnan and Kaplan (1982) discuss the relation between linguistic competence and linguistic performance. They show that transformational theories of linguistic competence do not meet the goal expressed by Chomsky (1965) that a theory of linguistic performance will incorporate a theory of competence as one of its components. As discussed above, LFG is designed to conform to what is known about the computation of language, and thus is more likely to be incorporable into a theory of performance.

LFG also has, as we will see in the next three chapters, a well-developed formalism. As in early transformational grammar, and unlike GB/MP, linguistic descriptions must be expressed in a rigorous formalism and not in informal prose. It is thus possible to examine whether an analysis conforms to the data.

In this textbook, we will develop an explicit grammar for much of the syntax of English as we develop the formalism of the theory. We will thus see how LFG can be used to produce an actual grammar.

Additional Readings

The conceptual basis for LFG is laid out in Bresnan and Kaplan (1982). Early psycholinguistic studies can be found in Bresnan, ed (1982) as well; for a more recent discussion, see Pinker (19xx). The properties of unification-based grammars are discussed in Shieber (1986). Parallel, correspondence-based architecture is discussed and argued for (from a non-LFG perspective) by Jackendoff (1997), who calls it "representational modularity."

The argument for a lexical analysis of the passive construction dates back to Bresnan's pre-LFG work (Bresnan 1978), and was further developed in Bresnan (1982; 1995; 2000 Chapter 3). The Lexical Integrity Principle and the c-structure/f-structure distinction is discussed in Bresnan

and Mchombo (1995).

Nonconfigurational languages have featured prominently in work on LFG, with continual arguments against transformational and configurational analyses. Early discussions can be found in Mohanan (1982) and Simpson (1983). Chapter 1 of Nordlinger (1998) provides extensive critical discussion of recent analyses of nonconfigurational languages in the GB/MP tradition.

Mathematical properties of LFG grammars have been discussed in Kaplan and Bresnan (1982) and many of the papers in Dalrymple, Kaplan, Maxwell, and Zaenen, eds. (1995).

Exercises

1. What English sentence does each of the following f-structures represent?

a.
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{PERS} & 1 \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`believe } \langle \text{SUBJ, COMP} \rangle\text{'} \\ \text{COMP} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`constituent'} \\ \text{NUM} & \text{PL} \\ \text{ADJ} & \{[\text{PRED} \quad \text{`syntactic'}]\} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`move } \langle \text{SUBJ} \rangle\text{'} \\ \text{ADJ} & \{[\text{PRED} \quad \text{`quickly'}]\} \end{bmatrix} \end{bmatrix}$$

. b.
$$\begin{bmatrix} \text{OBJ} & [\text{PRED} \quad \text{`Hammerstein'}] \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{`send } \langle \text{SUBJ, OBJ, OBJ2} \rangle\text{'} \\ \text{SUBJ} & [\text{PRED} \quad \text{`Rodgers'}] \\ \text{OBJ2} & \begin{bmatrix} \text{PRED} & \text{`song'} \\ \text{DEF} & - \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}$$

c.
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`idea'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`colorless'} \end{bmatrix} \begin{bmatrix} \text{PRED} & \text{`green'} \end{bmatrix} \right\} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`sleep } \langle \text{SUBJ} \rangle \text{'} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`furiously'} \end{bmatrix} \right\} \end{bmatrix}$$

2. Note the values of the function ADJ (adjunct) in 1. Why do you suppose the value of ADJ is a set of smaller f-structures instead of a single smaller f-structure?

3. Give the f-structures of the following sentences.

   a. Mary had a little lamb.
   b. A funny thing happened on the way to the forum.

2

Constituent Structure

2.1 Constituent structure in LFG

The basic concepts of c-structure are familiar from transformational theories. Constituent structure is an organization of the words that make up a sentence into successively larger and larger units, where each unit (**constituent**) belongs to a **category**. By separating constituent structure from grammatical functions, LFG is more able than structurally-based theories to reflect the actual constituency properties that one finds in different languages. We will discuss both the similarities and the differences between the LFG theory of c-structure and the transformationalist theory.

C-structure is the overt expression of the features and functions that make up a syntactic expression. Unlike GB, MP, and other structurally-based theories, LFG does not require c-structure to contain all the syntactic properties of a constituent. Thus, for example, GB/MP and related theories require a set of empty categories to account for elements that receive no overt realization. Such empty categories are the result of the mixture of c-structural and f-structural information, and do not reflect purely c-structural properties.

As the overt aspect of syntax, c-structure is subject to the principle of Economy of Expression.

(1) **Economy of Expression**
All syntactic phrase structure nodes are optional and are not used unless required to create a well-formed f-structure or to add semantic content.

This principle severely limits the use of empty elements in c-structure, since (as we will see) most of the empty categories of transformational theories are redundant in LFG.

2.2 $\bar{X}$ theory

2.2.1 Lexical categories and their projections

The theory of c-structure assumed in most generative work, including LFG, is generally known as **$\bar{X}$ (X-bar) theory**. $\bar{X}$ theory is the result of research in the 1970s which aimed to constrain c-structure rules on the one hand and express cross-category generalizations on the other.

The basic **lexical categories** are N (noun), V (verb), A (adjective), P (preposition), and perhaps ADV (adverb).[1] These are the categories of words that carry meaning. The conventional view is that they are analyzed into binary distinctive features. Although we will not pursue this here, the proposed feature systems of Chomsky (1981) and Bresnan (2000) are presented in (2).

(2) a. Chomsky's features:
[±N] (nominal)
[±V] (verbal)

	[+N]	[−N]
[+V]	A	V
[−V]	N	P

b. Bresnan's features:
[±trans] (transitive)
[±pred] (predicative)

	[+trans]	[−trans]
[+pred]	V	A
[−pred]	P	N

A more worked out distinctive feature analysis of categories is needed to properly express generalizations across categories. Although there have been other proposals for feature systems (see Wunderlich 1996 and references

[1]The relation between adjectives and adverbs is unclear. In this section, we will ignore adverbs.

cited there), no study has been undertaken as thorough as the classic but now outdated Jackendoff (1977). For the purposes of this textbook, we will improvise with disjunctive expressions in curly brackets.

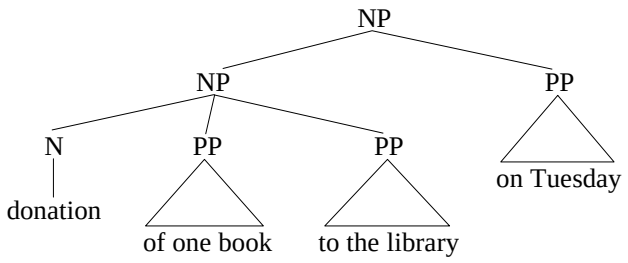
The basic insight behind the notion of constituent structure is that groups of words form constituents, or **phrases**, which can be identified by their ability to occur in different places in the sentence. The lexical categories are the **heads** of these phrases. The phrases headed by the lexical categories N, V, A, and P are called NP, VP, AP, and PP respectively.

(3) a. NP [donation of a book to the library on Tuesday]
b. VP [donate a book to the library on Tuesday]
c. AP [proud of the library]
d. PP [on Tuesday]

The head of a phrase defines the properties of the whole phrase. The phrase *donation of a book to the library on Tuesday* belongs to the category NP because the head belongs to the category N. The category NP is said to be the phrasal **projection** of N. Other properties of the NP are also determined by the head N. For example, while the NP *donation of a book to the library* is singular, the NP *donations of a book to the library* is plural. In structurally-based theories, this is achieved through some mechanism of feature percolation or feature passing. In LFG, a lexical head and its phrasal projection correspond to the same piece of f-structure, so their features unify. There is no need in LFG to stipulate a mechanism for passing features from head to projection; it is a consequence of the functional identification of head and projection.

In configurational languages like English, in which there is a close correlation between c-structure configurations and grammatical functions, arguments and adjuncts occupy different positions. Constituents that function as arguments are sisters of the head, i.e. in **complement** position, while adjuncts are **adjoined** to the phrasal node:

(4)



When examining the various phrases in a language, we discover that there is usually consistency in the relative ordering of heads and complements. In some languages, such as English and Hebrew, the head uniformly precedes its complement(s). In others, such as Japanese and Hindi, the head follows. This observation, originally due to Greenberg (1963), is captured in $\bar{X}$ theory by treating all categories as fundamentally identical in structure.[2]

On the other hand, different categories do have different properties. We can distinguish categories by their distributional properties. For example, in English the only category that freely ellipts is VP. Note the following contrast, where an ellipted element is shown by being crossed out.

(5) a. The hamster started running when the dinosaur stopped ~~running~~.
b. *The hamster started the race when the dino stopped ~~the race~~.[3]

Another category-based difference relates to the category of modifying elements: NPs are modified by adjectives while other categories are modified by adverbs.

(6) NP
a. urgent nomination of a candidate
b. *urgently nomination of a candidate

[2]For a detailed study of cross-category generalizations of this kind in English, see Jackendoff (1977).

[3]Note that this sentence is grammatical if the ellipted element is some action, determinable from the larger discourse, that the dinosaur stopped. But then what is ellipted is a VP.

(7) VP
a. urgently nominated a candidate
b. *urgent nominated a candidate

(8) AP
a. absolutely complete
b. *absolute complete

(9) ADVP
a. absolutely completely
b. *absolute completely

(10) PP
a. completely under the table
b. *complete under the table

Properties like these can be used as tests for category identity. Furthermore, while syntactic theory has to express the similarities between categories, differences like these also must be expressible.

2.2.2 Functional Categories

An interesting result of continuing research into $\bar{X}$ theory is the hypothesis that in addition to lexical categories, there are also **functional categories**. Consider the following nominal phrases.

(11) a. [the donation to the library]
b. [a donation to the library]

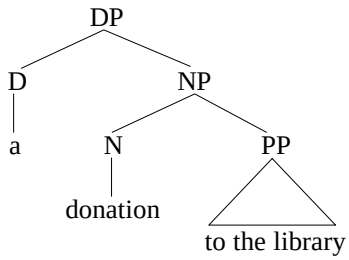
Their f-structures are as follows.

(12) a.
$$\begin{bmatrix} \text{DEF} & + \\ \text{NUM} & \text{SG} \\ \text{PRED} & \text{`donation } \langle \text{OBL}_{\text{Goal}}\ \text{OBJ} \rangle\text{'} \\ \text{OBL}_{\text{Goal}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{Goal}} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`library'} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{DEF} & - \\ \text{NUM} & \text{SG} \\ \text{PRED} & \text{`donation } \langle \text{OBL}_{\text{Goal}}\ \text{OBJ} \rangle\text{'} \\ \text{OBL}_{\text{Goal}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{Goal}} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`library'} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Definiteness, like number, is a grammatical feature of the nominal phrase, and therefore represented at f-structure. Consequently, (12a) (the f-structure of the definite nominal) has the feature [DEF +] while (12b) (the f-structure of the indefinite nominal) has the feature [DEF –]. Unlike [NUM SG], however, these properties are due not to the nominal's lexical head *donation*; instead, they come from the initial *a* or *the*. These particles belong to a category often called D(eterminer). What these examples show is that the determiner has the headlike property of setting properties of the phrase. The determiner also seems to occupy a head position relative to the NP: it precedes it, just as all heads in English precede their complements. Because of these headlike properties, it has become widely accepted that the determiner is the head of a phrase, a DP, within which the NP is a complement (Brame 1982, Abney 1987).

(13)



The determiner is called a functional category[4] because its purpose is to provide features for its phrase. The actual lexical semantic content (PRED feature) is provided by the head of the NP complement to the determiner. From the perspective of f-structure, the determiner and noun are **co-heads** of the DP.

[4]This terminology originated among GB theorists. The word *functional* here is not related to the technical LFG sense of grammatical function.

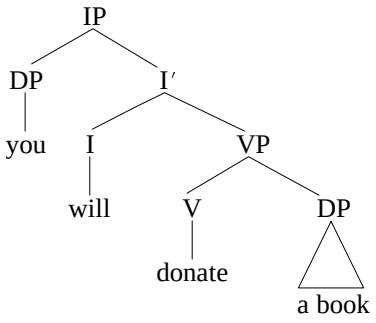
Similarly, consider the following.

(14) a. [You will donate a book]
b. [You did donate a book]

In (14a), the clause has the feature [TENSE FUT] while in (14b) it has the feature [TENSE PAST]. The tense feature comes not from the verb, but from the words *will* and *did*. These words are called auxiliaries in traditional terminology; in current generative syntactic terminology the category they belong to is usually called Infl (or I). Like the determiner in the nominal phrase, the infl in the clause acts like a head. As in our analysis of the nominal phrase, we can treat the clause as the phrasal projection of the functional category Infl (i.e. IP), with the VP in complement position (Falk 1984, Chomsky 1986, Kroeger 1993).

We also see another aspect of phrase structure in (14). The DP which is assigned the function SUBJ occupies (in English) a special structural position within the IP: it is a daughter of IP and a sister to a node (an intermediate projection of infl) which contains the infl and its VP complement. This special position is called the **specifier** position, and the intermediate projection is called I′ (I-bar).

(15)



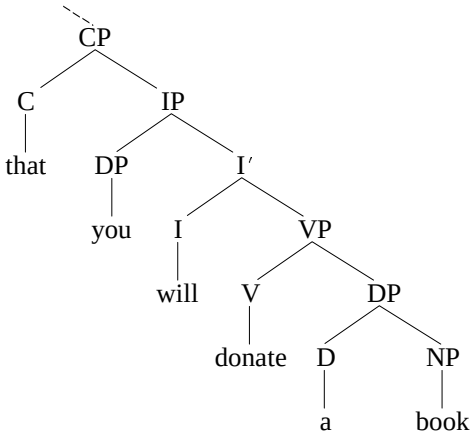
In the GB/MP tradition, functional categories other than D and I have been proposed. The question of other functional categories emphasizes a central difference between the LFG conception of c-structure and the transformationalist notion. In GB/MP, functional categories can be sublexical features that merge with the lexical head in the course of the derivation. As a result, various sublexical features have been proposed as functional heads, such as agreement, aspect, transitivity, etc. While these surface in some

languages as verbal affixes, they are not independent words. The Lexical Integrity Principle does not allow such affixal features to be analyzed as distinct c-structural nodes. The functional categories we have discussed, D and I, are categories of *words*, not sublexical affixes. In a language with, for example, no free form defining tense, I could not be analyzed as tense features that must merge with the verb. Such a language might simply lack the category I, or finite verbs might be lexically specified as belonging to the category I instead of V. (The latter analysis corresponds to GB's "V-to-I movement.")

Even within the more limited role assigned to functional categories under the Lexical Integrity Principle, there are additional candidates for functional categories. The clearest case is the complementizer (C) (Fassi-Fehri 1981, 1982, Chomsky 1986), which (usually) takes an IP as its complement. The complementizer determines certain properties of the clause, such as whether it is declarative or interrogative. Like determiners and infls, it precedes its sister (IP) in headlike fashion.

(16) a. … that you will donate a book

b.



Another potential candidate is Case (K) (Fillmore 1968, Bittner and Hale 1996, Falk 1997, Butt and King 1999), which also displays (in languages where it is a free form and not a nominal affix) head-like properties. If K is a functional category, it (usually) takes a DP as its complement and heads a KP, which is the full nominal category. We will not be concerned with K in this book.

The specifier position also represents a difference between LFG and GB/MP. In MP (and some versions of GB), specifier position is hypothesized to be the locus of agreement or feature checking. Many elements that are said to assume specifier positions do not do so overtly, and are hypothesized to move there at the covert level of "logical form." In LFG, specifier position is an overt position. The three functional categories that we are hypothesizing all have specifier positions. As we have seen, the [SPEC, IP] position is the structural position of the subject in English (although this is not necessarily true cross-linguistically: in some languages subjects occupy other positions, and in some languages [SPEC, IP] is a focused or topicalized position). [SPEC, CP] is the position occupied by fronted *wh* elements. This is harder to show since, unless *to* is a complementizer,[5] there is no lexical complementizer when [SPEC, CP] is filled.

(17) a. I don't know [what I should read].
b. *I don't know [what that I should read].
c. I don't know [what to read].

We will return to this below. Similarly, [SPEC, DP] is a subject-like position (possessor) within the nominal phrase; here again, when the specifier position is filled the functional head is absent.

(18) a. I admire [Dave's hamster].
b. *I admire [Dave's the hamster].
c. I object to [Dave's reading that book].

It is not clear whether lexical categories have specifier positions. Bresnan (2000) hypothesizes that only functional categories have specifiers. On the other hand, it has been hypothesized in GB/MP that the true subject position is [SPEC, VP], and that the [SPEC, IP] position is the result of movement.[6] For concreteness, we will assume Bresnan's hypothesis in what follows. As a result, we depart from the approach of "traditional" $\bar{X}$ theory, which analyzes all categories as heaving a uniform X-X′-XP (or X″) projection. What we are calling VP, NP, AP, and PP correspond most closely to what are

[5]We will argue in Chapter 5 that *to* is a complementizer.

[6]Kroeger (1993) argues that the alleged VP-internal subject in Tagalog is not inside the VP.

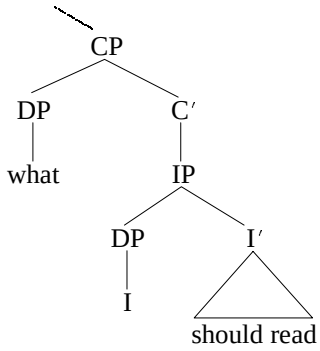
usually called V′, N′, A′, and P′.

2.2.3 Endocentricity

One of the claims of $\bar{X}$ theory is that all phrases have heads of the same category: an NP is headed by an N, an IP is headed by an I, etc. This property is called **endocentricity**. However, it turns out that this claim must be qualified somewhat. One counterexample is discussed here; a second is included in the exercises for this chapter.

As discussed above, heads of functional phrasal categories are sometimes absent. Consider the subordinate clause in (17a) above. Its structure is:

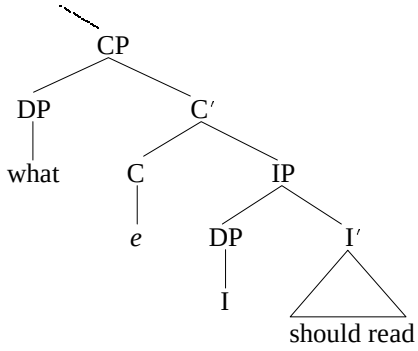
(19)



In this structure, the CP has no head C. This is a violation of endocentricity.

One possible solution to this problem would be to postulate an "empty" head complementizer:

(20)



This hypothetical empty C violates the principle of Economy of Expression. It contributes nothing to either the f-structure or the meaning, and is thus superfluous. It is nothing more than an artificial device for forcing a headless CP to appear to have a C head.

Bresnan (2000) suggests that the headless CP is grammatical because the specifier itself provides the features that would be provided by the head. A complementizer such as *if* or *whether* would mark the clause as a question; that is the function of these complementizers.[7] Formally, it adds a question feature to the f-structure corresponding to the clause. Here, however, the fronted *wh* element provides this feature. The specifier's contribution to the f-structure renders the head superfluous. Under Economy of Expression, superfluous c-structure nodes do not exist. Therefore, we expect the head to be absent. Under this approach, c-structure headedness is dependent on f-structure information.

It is instructive to compare this approach with the empty complementizer alternative. There are two variants of the empty complementizer approach. One variant would treat the complementizer as truly empty. It is unclear that there is any advantage to this approach. As noted above, it is nothing more than the use of an artificial device to save endocentricity after a counterexample has been discovered. A more sophisticated approach is to treat the empty complementizer itself as containing a question feature, and entering into an agreement relation with the specifier. Here again, it is not clear what the purpose of the complementizer is. This is not agreement in the usual sense, where some overt marking on a head corresponds to some feature of a dependent. Since the complementizer is not pronounced, there is no overt marking. Furthermore, such an account does not explain why the complementizer can be empty precisely in a case where the information that would come from the complementizer comes from some other source.

Finally, a word about another kind of construction that is sometimes thought to involve a null complementizer. Compare the complement clauses in the following sentences.

(21) a. I know [that the world is flat].
b. I know [the world is flat].

In (21a), the complement of the verb *know* is a CP, headed by the

[7]Recall that C is a functional category.

complementizer *that*. In (21b), on the other hand, there is no complementizer. Such cases are sometimes analyzed as involving an empty complementizer which is an allomorph of *that*. Unlike the case discussed earlier, here there is no other element in the clause to render the complementizer superfluous. However, there is no reason to posit an empty complementizer because there is no reason to posit a CP. An alternative analysis would be to treat the complement clause here as a bare IP. Importantly, *that*-less clauses have a different distribution from CPs. For example, CPs (with or without an overt complementizer) can be subjects; *that*-less clauses cannot.

(22) a. [What I should read] is a question that disturbs me.
b. [That the world is flat] disturbs me.
c. *[The world is flat] disturbs me.

The different distribution suggests very strongly that the *that*-less clause belongs to a different category. This is thus not a case of a missing complementizer.

2.3 Phrase structure rules

It is clear from examining many languages that different c-structures are grammatical in different languages. This, in fact, was one of the points made in Chapter 1 when we demonstrated the need to separate grammatical functions from c-structure configurations. This means that the grammar of each language will have rules defining (or licensing) well-formed c-structures. This contradicts the approach taken in contemporary transformational theories, in which phrase structure configurations are taken to be the result of universal or near-universal principles (such as the Projection Principle) relating the lexicon to the syntax.

The traditional formal device for licensing c-structures is the **phrase structure rule**. A first approximation at the phrase structure rules for English is (23).

(23) a. Functional maximal projections

$$\text{CP} \rightarrow \text{XP}\ \text{C}'$$

$$\text{IP} \rightarrow \begin{Bmatrix}\text{DP}\\ \text{CP}\\ \text{PP}\end{Bmatrix}\ \text{I}'$$

$$\text{DP} \rightarrow \text{DP}\ \text{D}'$$

b. Functional single-bar projection
$C' \rightarrow C\ IP$
$I' \rightarrow I\ VP$
$D' \rightarrow D\ NP$

c. Lexical phrases

$$VP \rightarrow V\ \ DP\ \ DP\ \ PP^*\ \left\{\begin{matrix} IP \\ CP \end{matrix}\right\}$$

$$PP \rightarrow P\ \ DP\ \ PP\ \ IP$$

$$NP \rightarrow N\ \ PP^*\ \ CP$$

$$AP \rightarrow A\ \ PP\ \left\{\begin{matrix} IP \\ CP \end{matrix}\right\}$$

d. Adjuncts[8]

$$XP \rightarrow \left\{\begin{matrix} AP \\ ADVP \end{matrix}\right\}\ XP\ ,\ XP \text{ a lexical category}$$

$$XP \rightarrow XP\ \ PP\ ,\ XP \text{ a lexical category}$$

In these rules, the categories on the right are licensed as c-structure daughters of the categories on the left, in the order shown. Note that the different distribution of CP and IP, mentioned in the previous section, is expressed in these rules.[9]

A few comments are in order about the formalism of phrase structure rules. First, obligatory and optional constituents are often distinguished by putting parentheses around optional constituents. We will not do this here because the Economy of Expression principle states that all c-structure nodes are optional. Second, the asterisk after PP in (23f) is an operator, called the **Kleene star**, which means 'any number of'. Thus, according to this rule, a VP can contain a maximum of one V, two DPs, any number of PPs, and a

[8]We assume that an appropriate feature system will make it possible to specify that AP goes with NP and ADVP with other categories. Also note that ADVP in VP need not precede the VP; we will return to this.

[9]It has been argued (Koster 1978, Bresnan 1994) that "subject" CPs are not in the same structural position as subject DPs. We will be assuming that they are in the same position.

constituent that can be either IP or CP. A related notation is the **Kleene plus**, which requires at least one. For example, coordinated phrases consist of a sequence of phrases, followed by a conjunction and one more phrase, as in *lions, tigers, and bears*. There must be at least one phrase before the conjunction. We can express this with the following phrase structure rule.

(24) $XP \rightarrow XP^{+}$ CONJ XP

The Kleene plus defines an exception to the statement in the Economy of Expression principle that c-structure nodes are optional. It is not just the pre-CONJ part that must appear, the conjunction and the final conjunct must also be present. This seems to be a special property of coordinate structures.

Phrase structure rules have been much maligned in post-1980 generative syntax. In the GB framework (Chomsky 1981, Stowell 1981) it has been claimed that they are redundant because they follow from more general principles of the theory (primarily endocentricity and the Projection Principle). Therefore, it has been argued, they serve no purpose in the grammar. The idea that c-structure configurations are completely predictable from independent principles is shown to be false by the wide range of c-structural properties available in different languages. However, even if the premise were true, the conclusion would not follow. A formal theory of c-structure must have a formal device to define what is a well-formed c-structure. Positing phrase structure rules does not necessarily mean that these rules are primitive; they could very well be derived from more general principles of grammar. It is instructive to recall that while $\bar{X}$ theory is generally taken in GB to constrain possible structures, earlier work in $\bar{X}$ theory (e.g. Jackendoff 1977) understood it to constrain phrase structure rules.

Another, more cogent, line of attack on phrase structure rules is that they conflate dominance and ordering relations. This argument has been made at various times in the history of generative syntax (Stahl 1967, Gazdar, Klein, Pullum, and Sag 1985, Falk 1983a, inter alia), and it has become standard in the theoretical framework of HPSG (and its predecessor GPSG) where the job of phrase structure rules is divided between **ID (immediate dominance) rules** and **LP (linear precedence) rules**. Such an approach allows the capturing of ordering generalizations which traditional phrase structure rules are unable to express. ID rules are like phrase structure rules, except that they don't specify relative ordering. A comma is placed between the daughter nodes. Our existing phrase structure rules are unchanged, except for the

adjunct rules, which can be simplified as a single rule if ordering is factored out.

(25) Adjunct ID rule

$$XP \rightarrow XP, \left\{ \left\{ \begin{array}{c} PP \\ AP \end{array} \right\} \right\}, \text{ XP a lexical category}$$
$$\left\{ ADVP \right\}$$

The generalizations about ordering are expressed by a separate set of LP rules.[10]

(26) a. X^0 initial
b. DP $\prec$ PP
c. PP ($\prec \left\{ \begin{array}{c} CP \\ IP \end{array} \right\}$) final
d. SPEC initial
e. $\left\{ \begin{array}{c} AP \\ ADVP \end{array} \right\} \prec \left\{ \begin{array}{c} NP \\ AP \\ PP \end{array} \right\}$

We will not restate the phrase structure rule for coordination (24) in terms of ID and LP rules. As we saw when we introduced the rule, it differs from other phrase structure rules in that its daughter nodes are obligatory. The ordering is also a crucial part of the coordination construction. We hypothesize that these are special constructional properties of coordination. This is expressed by the special status of the coordination phrase structure rule.

By factoring out ID and LP rules, more elegant statements of the generalizations about phrase structure can be achieved. We can express the fact that both PP and A(DV)P can be adjoined to any lexical category, with PP following because PPs always occur near the end of a phrase, and A(DV)P preceding in categories other than VP. Since no ordering is expressed for ADVP in VP, either ordering is possible.

---

[10](26c) is a makeshift notation. It indicates that PP either is itself final or is followed by a CP/IP which is final. This could be expressed as follows:

(i) PP final $\vee$ (PP $\prec \left\{ \begin{array}{c} CP \\ IP \end{array} \right\} \wedge \left\{ \begin{array}{c} CP \\ IP \end{array} \right\}$ final)

The LFG literature does not generally adopt the factorization of phrase structure rules (although Bresnan 2000 does, as does King 1995). We will adopt it in this text.

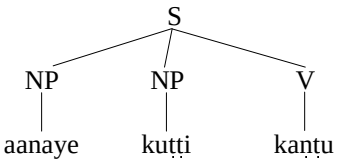
2.4 Exocentricity

In a theory of c-structure based solely on $\bar{X}$ theory (as in GB theory), all phrasal categories are projected from heads. That is to say, all categories are endocentric. It is clear why such a theory is desirable: it is restrictive and it is strongly lexical (since phrases are projected from lexical items).

Unfortunately, there is evidence that this is incorrect. For example, as we saw in Chapter 1, there are languages that lack the category VP. Consider again the following sentence from Malayalam.

(27) a. Aanaye kuṭṭi kaṇṭu.
elephant.ACC child.NOM saw
'The child saw the elephant.'

b.



In a language with flat clausal structures of this type, the usual concepts of $\bar{X}$ theory do not apply. There is no specifier-head-complement structure, no functional category - lexical category progression. Formally, the clause has no c-structure head. A phrasal category with no c-structure head is said to be **exocentric**.

The claim within generative grammar that exocentric structures are possible is not original with LFG. The first generativist to propose two different types of syntactic organization was Kenneth Hale in the 1970s (see, for example, Hale, Jeanne, and Platero 1977), who distinguished between configurational and nonconfigurational languages, or "X-bar" (languages that obey the principles of $\bar{X}$ theory) and "W-star" (languages that have the phrase structure rule $S \rightarrow W^*$, where W is a word). Bresnan (2000) follows up on this by referring to languages involving **endocentric organization** and **lexocentric organization**. Under endocentric organization, grammatical functions are encoded in c-structure configurations. Under lexocentric organization, grammatical functions are encoded by lexical means, such as

Case and agreement morphology. Languages may have a mixture of endocentric and lexocentric organization.

LFG enriches the theory of c-structure with a single nonprojective exocentric category, S, distinct from IP. S is not part of the $\bar{X}$ system. It does not have a c-structure head, and therefore cannot be identified with any lexical category. It can consist of a string of words, or a string of phrases with a single lexical item to serve as the functional head, or a NP (or DP) subject and predicate of any category.

(28) a. $S \rightarrow X^*$
b. $S \rightarrow XP^*, X^0$
c. $S \rightarrow NP, XP$

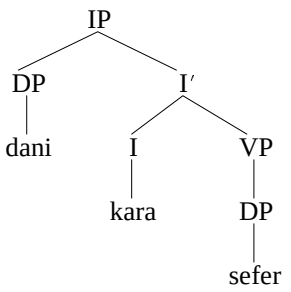
Languages can combine endocentric and exocentric structures. For example, consider Hebrew.

(29) a. dani kara sefer.
Danny read a.book
'Danny read a book.'

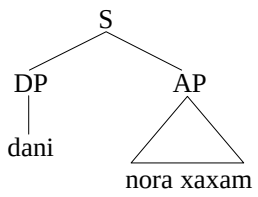
b. dani nora xaxam.
Danny awful smart
'Danny is awfully smart.'

(29a) is an ordinary sentence, with an IP-over-VP structure. (As in many languages, the finite verb in Hebrew is in I rather than V.) In (29b), on the other hand, there is no c-structure head to the sentence. It consists of a DP subject and an AP predicate. The c-structures of the sentences are therefore as follows.

(30) a.



b.



Both IP and S can be complements to CP in Hebrew. So the syntax of Hebrew will include the following phrase structure rules.

(31) a. $CP \rightarrow XP \ C'$

b. $C' \rightarrow C \left\{ \begin{matrix} IP \\ S \end{matrix} \right\}$

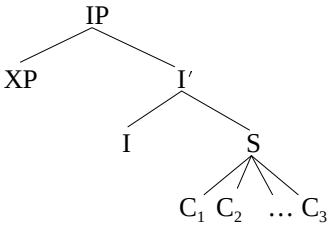
c. $IP \rightarrow DP \ I'$

d. $I' \rightarrow I \ VP$

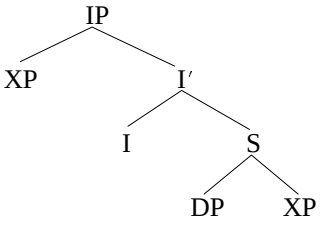
e. $S \rightarrow DP \ XP$

Even radically nonconfigurational languages like Warlpiri and Wambaya can have endocentric structures. Recall from Chapter 1 that in these languages the "auxiliary" must be in second position. The "auxiliary" is infl, which in these languages includes tense and agreement features and is always separate from the verb. It has been proposed by Austin and Bresnan (1996) for Warlpiri and Nordlinger (1998) for Wambaya that the single constituent before the verb is a focused or topicalized element in [SPEC, IP]. However, following the infl there is no evidence for internal structure, and constituents (including the head verb) can occur in any order. The proposed structure, then, has an S as complement to I, as in (32a). Similar analyses have been proposed for Irish (32b) and Tagalog (32a,b) (Kroeger 1993).

(32) a.



b.



Additional readings

The literature on $\bar{X}$ theory is vast and goes far beyond LFG. Here, we will outline some highlights.

$\bar{X}$ theory has its origins in Chomsky (1970), and was explored by many researchers in the 1970s. The best known (and most thorough) of these studies is Jackendoff (1977). The notion of functional categories, although popularized in GB in the late 1980s, has its origin in work done by lexically-oriented syntacticians in the late 1970s and early 1980s, as noted in the text. Thus, lexically justifiable functional categories such as D, I, and C (and maybe K) are adopted in LFG not in imitation of GB but because lexical properties justify them. More recent LFG studies on $\bar{X}$ theory include King (1995), Kroeger (1993), and Bresnan (2000). All of these take an approach to functional categories similar to the one taken here. For a more skeptical view, see Börjars, Chisarik, and Payne (1999).

The Economy of Expression principle is inherent in most work in LFG, and is formulated in Bresnan (2000).

The discussion of endocentricity is based largely on Bresnan (2000). For an analysis of "mixed categories" like the English gerund (a DP headed by a VP), see Bresnan (1997; 2000 Chapter 13).

The exocentric category S was introduced in Bresnan (1982b), and has figured prominently in LFG studies of nonconfigurational languages, but was first clearly distinguished from IP by Kroeger (1993). Simpson (1991), Austin and Bresnan (1996), and Nordlinger (1998) discuss nonconfigurationality and how to analyze it.

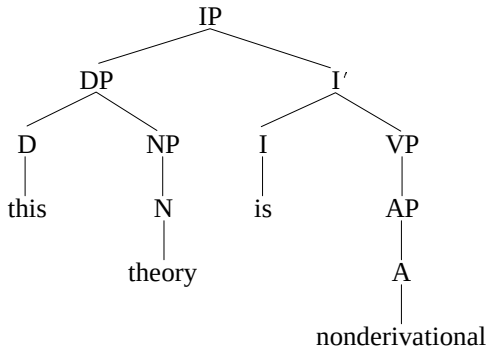
Exercises

1. Over the 40+ year history of generative syntax, the auxiliary system of English has received a fair amount of attention from a variety of theoretical frameworks. We will assume here that the following c-structural analysis is correct.
   ① the finite forms of auxiliary *have* and all uses of *be* occupy the structural position of I, as do finite forms of the verb *do* in "supportive" use
   ② all other finite verbs occupy the V position
   ③ all nonfinite verb forms (including nonfinite forms of *have* and *be*) are in V position
   ④ modal auxiliaries, which have no verb-like properties, are in I
   ⑤ elements of I (modals; finite *be*, auxiliary *have*, supportive *do*) are in C position in direct questions ("Subject-Auxiliary Inversion")

These assumptions are widely accepted, and are supported by a large body of evidence. (For a (somewhat outdated) LFG perspective, see Falk 1984.) Now consider the sentence:

This theory is nonderivational.

On the assumption that the complement of I must be VP, its c-structure would have to be the following.



In a derivational theory like Government/Binding theory, this structure is the consequence of V-to-I movement, and *is* leaves a trace in the position of the head of the VP. In LFG, there is no movement and no evidence for a trace in V position. As discussed in the text, *is* is an I because it is lexically specified as belonging to the category I. That is to say, instead of a movement analysis of the I-hood of *is*, LFG offers a lexical analysis. Without motivation (either theoretical or empirical) for an empty V in the VP, the VP is unheaded.

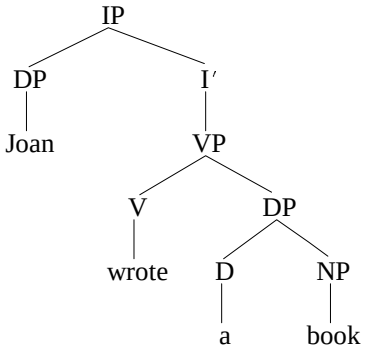
a. What kind of evidence might there be to choose between a movement analysis and a lexical analysis? Is there any evidence in English favoring a lexical analysis?

b. Is there empirical evidence for the VP node? (Does the constituent *nonderivational* act like a VP in any way?)

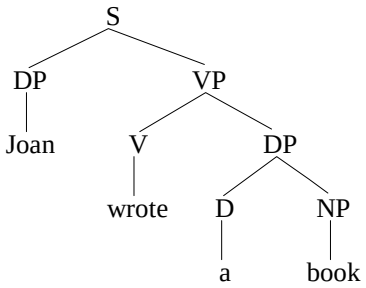
c. In the text, we discussed the problem raised for endocentricity by complementizer-less CPs, and determined that the problem can be solved by referring to the f-structure correspondent of the CP. Can you think of a similar approach that would make the headless VP not problematic?

2. Unlike GB, LFG has a nonprojective (non-$\bar{X}$) category S. Consider a sentence with no auxiliary, such as *Joan wrote a book*. Should it be analyzed as an IP (as in (i)) or an S (as in (ii))?

(i)



(ii)



Explain your answer.

3. We have argued that determiners head a phrase (DP) of which the NP is a complement. Consider the bracketed phrases in the sentences in (i).

   (i) a. [Students] don't like to do homework.
   b. They threw a party for [practitioners of generative basket-weaving].
   c. The book received [great reviews].

   Are these phrases NPs or DPs? Explain your answer. What implications does this have for the ID rules for English?

3

Functional Structure

In this chapter, we will discuss the properties of f-structure. F-structures will be more mysterious to you than c-structures, since they do not have a formal analog in transformational theories. As discussed in Chapter 1, f-structures are attribute-value matrices (AVMs), in which the attributes are features and grammatical functions, and the values can be either atomic entities or smaller f-structures.

3.1 Grammatical functions

The crucial concept behind f-structure is, of course, grammatical functions. Grammatical functions are postulated in LFG because they display internal coherence, but do not correspond to a uniform semantics or a uniform c-structural realization. Just as LFG's theory of c-structure builds on ideas from other theoretical frameworks, like GB, the theory of grammatical functions also draws on ideas first raised in other theories. The first generative theory based on grammatical functions is Relational Grammar (RG; Perlmutter, ed. 1983), so not surprisingly ideas from RG have influenced LFG.[1] LFG has also drawn on ideas from typological and functionalist approaches.

The most basic function that syntactic elements serve is to express arguments of predicates. Consequently, the most basic grammatical functions are the **argument functions**. The best understood of these are SUBJ (subject), OBJ (object), OBJ2 (secondary object),[2] and the OBL_θ (oblique) family of

[1]We will not discuss differences between LFG and RG.

[2]We will have more to say about OBJ2 in Chapter 4.

functions. Additional functions of this type are POSS (possessor), which is used for arguments of nouns, and COMP (complement), which is used in languages like English instead of OBJ for certain arguments that are realized as CP or IP. There are also **nonargument functions**, such as ADJ (adjunct), FOCUS, and TOPIC. This distinction between argument functions and nonargument functions is similar to the GB distinction between A positions and Ā positions, but in LFG the distinction is between grammatical functions rather than c-structure positions. The distinction is equally valid for languages that distinguish arguments from nonarguments structurally and those that do not.

Research on grammatical functions has shown that further distinctions can be made among the argument functions. There is a fundamental distinction between grammatical functions like SUBJ, OBJ, and OBJ2 on the one hand, and the OBL_θ family on the other.[3] The former are called the **core** (or **term**) functions, while the latter are **noncore** (or **nonterm**). The core functions are the ones that are typically realized as DPs in languages like English and nominative or accusative[4] Case in languages with morphological Case. In addition, function-based phenomena most frequently involve the core functions. Conversely, noncore functions are rarely implicated in function-based constructions, and are typically marked with prepositions or Cases expressing their thematic roles. Core functions are more strictly *grammatical* functions, while noncore functions are more closely tied to semantics.

A finer-grained distinction can be made among the core functions as well. As first noted by Keenan and Comrie (1977), the argument functions are arranged in a **relational hierarchy**, indicating their relative accessibility to grammatical processes such as relativization (the process discussed by Keenan and Comrie), antecedence of anaphors, etc.

(1) $\text{SUBJ} > \text{OBJ} > \text{OBJ2} > \text{OBL}_\theta$

Note that the core functions outrank the noncore functions on the relational hierarchy. More interestingly, the core functions are themselves hierarchi-

[3]We will not discuss the status of the POSS and COMP functions in terms of the core/non-core distinction; it is not clear how this distinction applies to them.

[4]or ergative, in ergative and active languages.

cally arranged. The hierarchy corresponds very closely to the structural positions associated with the argument functions in GB and MP, with hierarchy-related phenomena attributed to the structural relation "c-command." The difference between the relational hierarchy and c-command is that the structural configurations on which c-command is based are not independently justified in all languages.

Argument functions (and the function ADJ) represent the clause-internal aspect of syntactic elements. However, clauses do not exist in isolation: they are embedded in each other and form parts of discourses. So, as a secondary function, a syntactic element can relate to its clause's place in larger syntactic or discourse structures. These secondary functions can be called **overlay functions** (as in Johnson and Postal 1980).

It is generally assumed in the LFG literature that there are at least three such overlay functions: TOPIC (or TOP), which expresses the topic of the discourse (and thus old information); FOCUS (or FOC), which expresses new information; and SUBJ (subject), which is the default discourse topic, and is a topic-like element connecting clauses in the same sentence (it is often shared between coordinated clauses, it is involved in "raising" constructions, etc.).[5] Note that, while the SUBJ function serves (at least partially) to connect clauses within a sentence, the other overlay functions relate a sentence to the larger discourse. For this reason, functions like TOPIC and FOCUS can be called **(grammaticized) discourse functions** (Bresnan 2000). Discourse functions are not part of discourse representation, any more than argument functions are part of lexical semantics. They are grammatical (i.e. syntactic) functions that express relations that are relevant for discourse grammar.

To conclude, the following is the set of grammatical functions that we will be assuming.

[5]The characterization of SUBJ as default topic is due to Andrews (1985) and Bresnan (2000). The cross-clausal sentence-internal properties of SUBJ are discussed in Falk (1999; 2000).

(2)

<table>
<tr><td>overlay</td><td colspan="3">nonoverlay</td><td>overlay</td></tr>
<tr><td colspan="3">argument</td><td colspan="2">nonargument</td></tr>
<tr><td colspan="2">core</td><td>non-core</td><td></td><td></td></tr>
<tr><td rowspan="2">SUBJ</td><td>OBJ
OBJ2</td><td>OBL$_\theta$</td><td rowspan="2">ADJ</td><td rowspan="2">FOCUS
TOPIC
etc.</td></tr>
<tr><td colspan="2">COMP
POSS</td></tr>
</table>

Nonoverlay argument functions (in other words, all argument functions other than SUBJ) are called **complement functions** and, as discussed above, nonargument overlay functions (i.e. all but SUBJ) are discourse functions. We will use the following abbreviations for classes of grammatical functions:

GF any grammatical function
AF argument function
$\bar{A}$F nonargument function
CF core function
CPF complement function
OF overlay function
DF discourse function

3.2 Well-formedness conditions

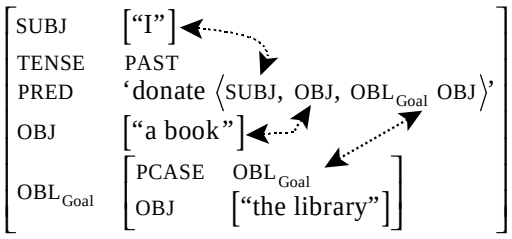
There are several well-formedness conditions on f-structures. These constraints play a role in ruling out certain sentences as ungrammatical.

The most basic kind of constraint on sentences in any theory is related to arguments: the argument-filling elements must be matched up with selected argument types. In classical generative grammar, this is captured by the notion of subcategorization, in GB by the Θ Criterion, and in MP by part of the Principle of Full Interpretation. This is why (3a) is grammatical while (3b) and (3c) are ungrammatical: in (3b) an argument is missing and in (3c) there is an extra argument.

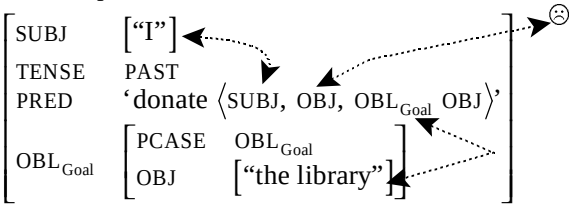
(3) a. I donated a book to the library.
b. *I donated to the library.
c. *I donated the university a book to the library.

In LFG, this is formalized in terms of a relationship between the argument functions specified in the lexical form of the head and the argument functions appearing as attributes in the f-structure. An f-structure in which all argument functions selected by the head actually appear is a **complete** f-structure; conversely, one that is missing (at least) one argument is **incomplete**. Similarly, an f-structure in which all the argument functions appearing as attributes are selected by the head, and thus fit into the argument structure, is said to be **coherent** (i.e. interpretable), while one in which there is (at least) one argument that is not selected is **incoherent**.

(5) a. a grammatical f-structure[6]

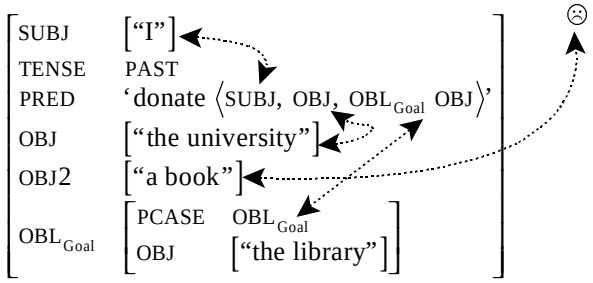


b. an incomplete f-structure



[6]A quick comment about notation: Note the use of double quotes as a shorthand. This is done when we do not want to specify the details of the subsidiary f-structure, just as a triangle is used in c-structure as an alternative to displaying the internal structure of a constituent.

c. an incoherent f-structure



We can state the Coherence and Completeness conditions as follows. We emphasize that completeness and coherence must be local; that is to say, the same AVM within the f-structure must contain both the governing PRED and the governed argument functions. (Recall that the term *f-structure* is ambiguous between the representation of the entire sentence and the various local f-structures that comprise it.)

(4) a. **Completeness Condition** (first approximation)
All argument functions specified in the value of the PRED feature must be present in the local f-structure.

b. **Coherence Condition** (first approximation)
All argument functions in an f-structure must be selected by their local PRED.

There is something missing from these statements. We have seen that idiom chunks and expletive elements are arguments, but differ from ordinary arguments in two ways. First, they are represented in the syntax as meaningless; they lack the PRED feature. Second, they are non-thematic arguments; they do not receive a thematic role from the verb. These two properties are related to each other, and to the Completeness and Coherence conditions. If a semantically empty element were placed in a thematic argument slot, the result would be the same as having nothing there.

(5) a. *I donated there to the library.
b. *Tabs donated a book to the library.

Conversely, a meaningful item placed in a non-thematic argument slot would be incoherent.

(6) a. *The sentence seems that the subject is an expletive.
b. *We kept surveillance on the suspects.

We can therefore add this to the statements of the conditions.

(7) a. **Completeness Condition**
All argument functions specified in the value of the PRED feature must be present in the local f-structure. All functions that receive a thematic role must have a PRED feature.

b. **Coherence Condition**
All argument functions in an f-structure must be selected by their local PRED. Any argument function that has its own PRED feature must be assigned a thematic role.

Although the Completeness and Coherence conditions have roughly the same effect as GB's Θ Criterion, there are some interesting differences. In the first place, the Θ Criterion relates to the mapping between argument structure and c-structure positions. The Completeness and Coherence refer to f-structure, not c-structure, and thus cannot be used to motivate the existence or nonexistence of a node in c-structure. This is important because in the GB literature the Θ Criterion (and its cousin, the Projection Principle) are often used to motivate the presence of empty categories. Arguments of this kind are impossible in principle in LFG. Second, Completeness and Coherence are not a "θ" criterion; they do not impose a one-to-one mapping between thematic roles and arguments. Finally, the Θ Criterion as generally interpreted includes the stipulation that the same element cannot receive thematic roles from two different heads. We have not incorporated this stipulation into the Completeness and Coherence conditions. It is at best unmotivated, and, as we will see when we discuss control constructions, there is reason to believe that it is incorrect.

The notion of coherence can be extended to nonargument functions. Recall that these come in two varieties: overlay functions and ADJ. For an overlay function to be coherent, it must be identified with some clause-internal position; FOCUS and TOPIC are simply highlighted elements that would be in the clause anyway.[7] Adjuncts are grammatical if they modify

[7]Recall that they are "overlay" functions: laid over the more basic a-functions.

meaningful elements; an adjunct modifying an expletive would not be incorporable into the meaning of the sentence.

(8) **Extended Coherence Condition**
All functions in an f-structure must be incorporated into the semantics. Argument functions are subject to the Coherence Condition. Overlay functions must be identified with argument positions. Adjuncts must be in f-structures containing PREDs.

There is one additional well-formedness constraint which we have been assuming implicitly. The concept of unification is based on the idea that an f-structure attribute can only have a single, or unique, value, or alternatively that its value must be consistent. This condition is called the Uniqueness Condition or the Consistency Condition.

(9) **Uniqueness Condition** (also called **Consistency Condition**)
Every attribute has a unique value.

That is to say, for example, that a clause cannot have two different SUBJs, or two different TENSEs.

The four conditions (Completeness, Coherence, Extended Coherence, and Uniqueness/Consistency) fulfill the same role in LFG as the Principle of Full Interpretation and the Θ Criterion do in transformationalist theories. They insure that the various parts of the sentence fit together. However, the LFG version is more specific that Full Interpretation in the requirements that it sets, and, as we have seen, differs from the Θ Criterion in its effects.

3.3 Some formalism: the c-structure–f-structure mapping

3.3.1 Overview

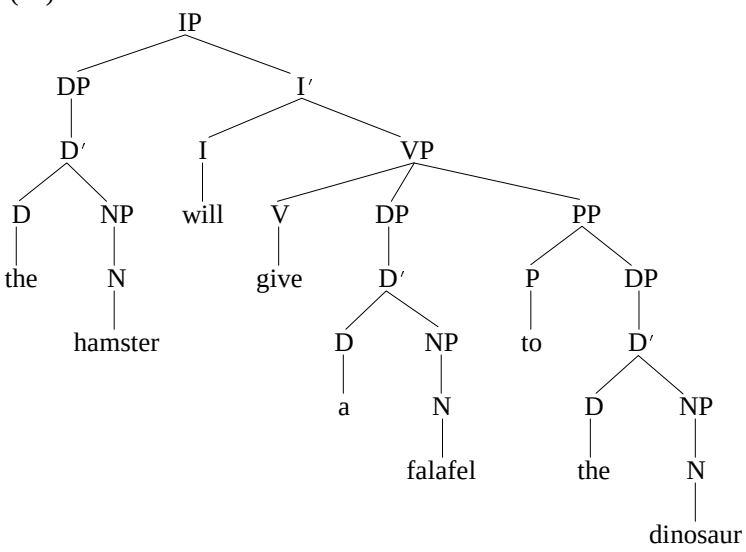
Now that we have discussed the properties of c-structure and f-structure, we will discuss the mapping between them. This mapping is the heart of the descriptive power of LFG, since it deals with the relationship between overt syntactic elements and the features they represent. The formal system is a mathematical algorithm; to students familiar with only semi-formalized theories like GB/MP it may look forbidding at first glance. The approach we will take here is to build up the concepts intuitively by starting with a sample c-structure and f-structure and working backwards.

The sample sentence we will use is:

(10) The hamster will give a falafel to the dinosaur.

We will assume the following structures.

(11)



$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`hamster'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{FUT} \\ \text{PRED} & \text{`give} \langle \text{SUBJ, OBJ, OBL}_{\text{Goal}}\ \text{OBJ} \rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & - \\ \text{PRED} & \text{`falafel'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{OBL}_{\text{Goal}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{Goal}} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

3.3.2 Correspondence

Recall that LFG is based on the idea that there is a relation of correspondence between nodes in the c-structure and parts of the f-structure. To make this

concrete, consider the outermost f-structure in the f-structure. The value of the TENSE feature comes from I, and the value of the PRED feature comes from V. The presence of the SUBJ attribute is due to a property of the IP, namely that it has a DP daughter. Finally, the presence of the OBJ and OBL_{Goal} attributes are due to properties of the VP. The outermost f-structure thus corresponds to a region of the c-structure composed of the IP-I projection and the VP-V projection. Similarly, the leftmost DP node and everything it dominates correspond to the value of SUBJ.

As we observed in Chapter 1, mathematical relations of correspondence are central to theories like LFG. The mapping relation from c-structure to f-structure is called **ϕ (phi)**; the mapping from f-structure to c-structure is therefore ϕ^{-1} (the inverse of ϕ). These mapping relations can be incorporated into grammatical rules. This is because all levels of structure exist simultaneously at different dimensions; there is no derivational relationship. For example, in English the constituent corresponding to the grammatical function OBJ precedes the one corresponding to OBJ2. This can provisionally be expressed in the following LP rule.

(12) $\phi^{-1}(\text{OBJ}) \prec \phi^{-1}(\text{OBJ2})$

This is not quite right, however. The problem is that several c-structure nodes can correspond to one f-structure; the c-structure correspondent of an f-structure is therefore not a node but a set of nodes. The solution is to define a precedence-like relation at f-structure derivative from c-structure precedence. This notion has come to be called **f-precedence**, symbolized with a subscripted *f* following the precedence symbol. Two slightly different definitions have been proposed in the literature. We will state them both in prose and with symbols.

(13) a. (from Kaplan 1987; Zaenen and Kaplan 1995)
An f-structure α f-precedes an f-structure β iff all nodes in the set $\phi^{-1}(\alpha)$ precede all the nodes in the set $\phi^{-1}(\beta)$.

$\alpha \prec_f \beta$ iff for all $n_1 \in \phi^{-1}(\alpha)$ and for all $n_2 \in \phi^{-1}(\beta)$, $n_1 \prec n_2$.

b. (from Bresnan 1995a; 2000)
An f-structure α f-precedes an f-structure β iff the rightmost nodes in the nonempty set $\phi^{-1}(\alpha)$ precedes the rightmost node in the nonempty set $\phi^{-1}(\beta)$.

$\alpha \prec_f \beta$ iff $\phi^{-1}(\alpha)$ and $\phi^{-1}(\beta)$ are nonempty and for n_1 the rightmost element of $\phi^{-1}(\alpha)$ and n_2 the rightmost element of $\phi^{-1}(\beta)$, $n_1 \prec n_2$.

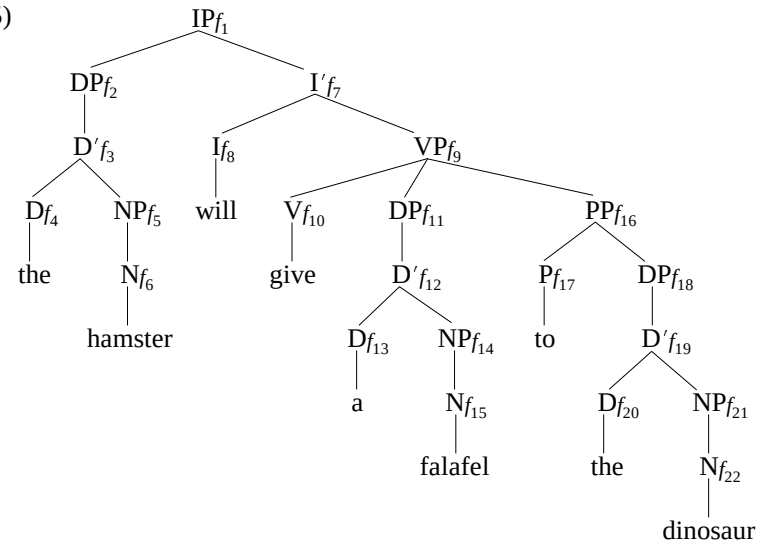
There are some interesting differences between these two definitions. For example, under the first definition an f-structure element with no corresponding c-structure realization vacuously f-precedes and follows everything, while under the second empty elements do not enter into f-precedence relations. For our purposes, either definition will suffice. We can restate our LP rule.

(14) $\text{OBJ} \prec_f \text{OBJ}2$

Since f-precedence, under either definition, is defined in terms of c-structure–f-structure correspondences, this LP rule still involves correspondence.

These correspondences can be formalized by assigning a **variable** to each corresponding pair, a symbol that can be used to represent the pair. By convention, the variables are indicated as f_1, f_2, etc. They can be marked on both the c-structure and the f-structure.

(15)



$$
\begin{bmatrix}
f_1, f_7, f_8, f_9, f_{10} & \text{SUBJ} & \begin{bmatrix} f_2, f_3, f_4, f_5, f_6 & \text{DEF} & + \\ & \text{PRED} & \text{‘hamster’} \\ & \text{NUM} & \text{SG} \end{bmatrix} \\
& \text{TENSE} & \text{FUT} \\
& \text{PRED} & \text{‘give} \left\langle \text{SUBJ, OBJ, OBL}_{\text{Goal}} \text{ OBJ} \right\rangle\text{’} \\
& \text{OBJ} & \begin{bmatrix} f_{11}, f_{12}, f_{13}, f_{14}, f_{15} & \text{DEF} & - \\ & \text{PRED} & \text{‘falafel’} \\ & \text{NUM} & \text{SG} \end{bmatrix} \\
& \text{OBL}_{\text{Goal}} & \begin{bmatrix} f_{16}, f_{17} & \text{PCASE} & \text{OBL}_{\text{Goal}} \\ & \text{OBJ} & \begin{bmatrix} f_{18}, f_{19}, f_{20}, f_{21}, f_{22} & \text{DEF} & + \\ & \text{PRED} & \text{‘dinosaur’} \\ & \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

3.3.3 F-descriptions

Given the formal expression of correspondence in terms of variables, we can define the mapping between the c-structure and the f-structure in our sample sentence. We do this with a series of equations called an **f-description** (functional description). For example, we want to say that f_1 and f_7 are the

same f-structure; i.e. the f-structure correspondents of constituents 1 and 7 are identical. This can be done with the **functional equation**:

(16) $f_1 = f_7$

Similarly, we want to say that if you find the f_1 and check the value of its SUBJ attribute, it will be f_2. So what we want to be equal to f_2 is a path through the f-structure starting at f_1 and continuing through the attribute SUBJ. This can be expressed as:

(17) $(f_1\ \text{SUBJ}) = f_2$

(Note that a path through the f-structure is enclosed in parentheses.)

The full f-description is:

(18) $(f_1\ \text{SUBJ}) = f_2$
$f_2 = f_3$
$f_3 = f_4$
$(f_4\ \text{DEF}) = +$
$f_3 = f_5$
$f_5 = f_6$
$(f_6\ \text{PRED})$ = 'hamster'
$(f_6\ \text{NUM}) = \text{SG}$
$f_1 = f_7$
$f_7 = f_8$
$(f_8\ \text{TENSE}) = \text{FUT}$
$f_7 = f_9$
$f_9 = f_{10}$
$(f_{10}\ \text{PRED})$ = 'give $\langle (f_{10}\ \text{SUBJ})\ (f_{10}\ \text{OBJ})\ (f_{10}\ \text{OBL}_{\text{Goal}}\ \text{OBJ})\rangle$'
$(f_9\ \text{OBJ}) = f_{11}$
$f_{11} = f_{12}$
$f_{12} = f_{13}$
$(f_{13}\ \text{DEF}) = -$
$(f_{13}\ \text{NUM}) = \text{SG}$
$f_{12} = f_{14}$
$f_{14} = f_{15}$
$(f_{15}\ \text{PRED})$ = 'falafel'
$(f_{15}\ \text{NUM}) = \text{SG}$

$(f_9\ \text{OBL}_{\text{Goal}}) = f_{16}$
$f_{16} = f_{17}$
$(f_{17}\ \text{PCASE}) = \text{OBL}_{\text{Goal}}$
$(f_{16}\ \text{OBJ}) = f_{18}$
$f_{18} = f_{19}$
$f_{19} = f_{20}$
$(f_{20}\ \text{DEF}) = +$
$f_{19} = f_{21}$
$f_{21} = f_{22}$
$(f_{22}\ \text{PRED}) =$ 'dinosaur'
$(f_{22}\ \text{NUM}) = \text{SG}$

Compare the f-description with the c-structure and f-structure above. You should see how the c-structure–f-structure relation is expressed by the f-description.

A couple of observations are in order about the f-description:

- In the DP *a falafel* the singularity is a property of both the D *a* and the N *falafel*. This is why the f-description includes both $(f_{13}\ \text{NUM}) = \text{SG}$ and $(f_{15}\ \text{NUM}) = \text{SG}$ (D is f_{13} and N is f_{15}). Since $f_{12} = f_{13}$ and $f_{12} = f_{14} = f_{15}$, and therefore $f_{13} = f_{15}$, these values will unify in f-structure. This explains the ungrammaticality of **a falafels*, in which the same f-structure would have the values [NUM SG] and [NUM PL]. This kind of specification is inconsistent (i.e. it violates the Consistency/Uniqueness Condition). As already observed, LFG does not require any extra machinery to enforce this sort of agreement.

- Note the PRED value of *give*. The arguments, which we have been representing until now as SUBJ, etc., are more properly the values of the SUBJ, etc., functions in *give*'s f-structure. The equation in the f-description has been so notated; the f-structure should be updated accordingly.

Mathematically, the f-description is a set of simultaneous equations, and the f-structure is the solution (technically, the minimal solution, since any f-structure with additional material would also be a solution to the f-description).

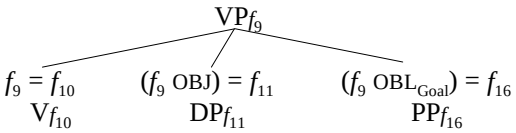
3.3.4 Functional annotations

The equations of the f-description are associated with parts of the c-structure.

For example, the equation (f_9 OBJ) = f_{11} is due to the fact that f_{11} is a DP which is a daughter of the VP associated with the variable f_9. In English, this structural position is associated with the grammatical function OBJ. This is a fact that the grammar of English must express.

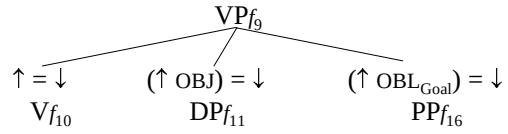
The equations can be annotated to the c-structure nodes with which they are associated to make the connection clearer. For example, the piece of tree containing the daughters of VP would be:

(19)



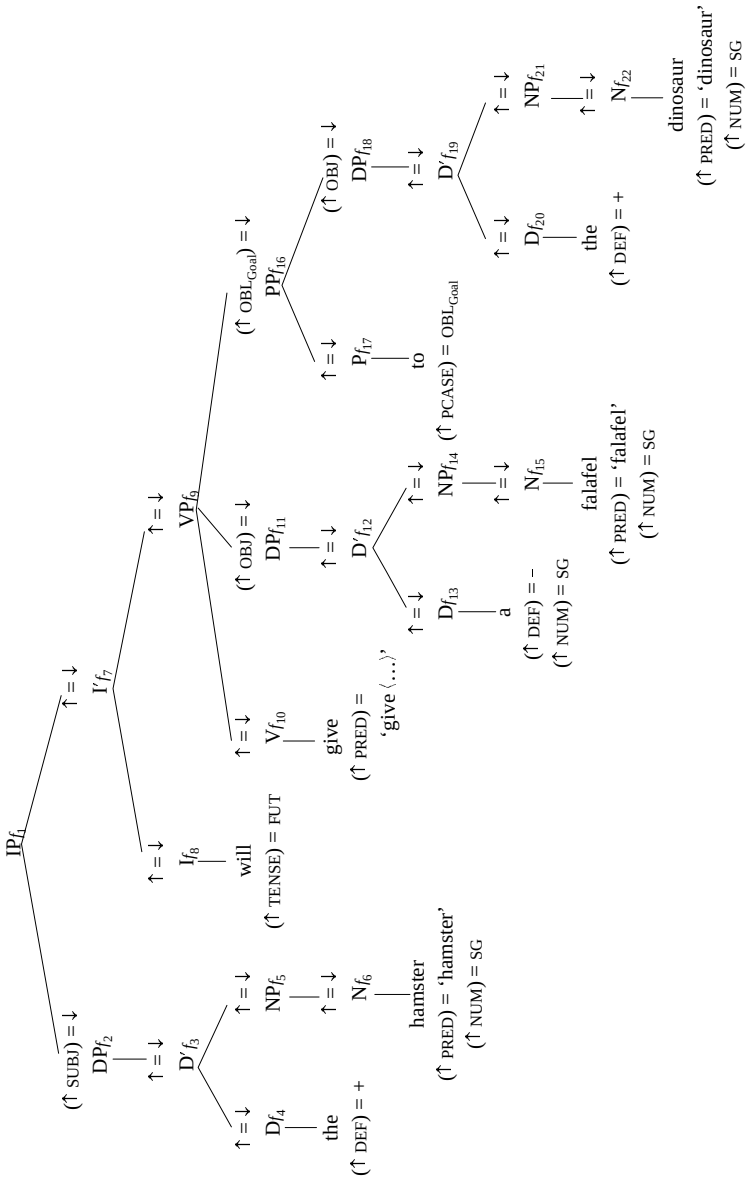
An examination of this tree fragment reveals a very interesting fact: the equations express local relations between mother and daughter nodes. We already know that locality is a desirable feature of a theory of syntax because syntactic processing is local. We can make this explicit by replacing the actual variables for this specific structure with variables for the variables, or **metavariables**, which will stand for constituents in a particular configuration in any sentence. To make this mnemonic, we use ↑ for the mother node and ↓ for the daughter node.

(20)



A c-structure with added functional information is called an **annotated c-structure**. The full annotated c-structure for the sample sentence is shown on the next page.

The annotations show clearly where the functional equations come from, and how they can be incorporated into the grammar of English. Some are introduced with ID (or phrase structure) rules, while others are lexical properties.


IPf_1
(↑ SUBJ) = ↓
↑ = ↓
DPf_2
↑ = ↓
D′f_3
↑ = ↓
Df_4
the
(↑ DEF) = +
↑ = ↓
NPf_5
↑ = ↓
Nf_6
hamster
(↑ PRED) = ‘hamster’
(↑ NUM) = SG
I′f_7
↑ = ↓
If_8
will
(↑ TENSE) = FUT
↑ = ↓
VPf_9
(↑ OBJ) = ↓
Vf_{10}
give
(↑ PRED) = ‘give ⟨…⟩’
(↑ OBL$_{Goal}$) = ↓
PPf_{16}
↑ = ↓
Pf_{17}
to
(↑ PCASE) = OBL$_{Goal}$
↑ = ↓
DPf_{18}
(↑ OBJ) = ↓
↑ = ↓
D′f_{19}
↑ = ↓
Df_{20}
the
(↑ DEF) = +
↑ = ↓
NPf_{21}
↑ = ↓
Nf_{22}
dinosaur
(↑ PRED) = ‘dinosaur’
(↑ NUM) = SG
↑ = ↓
D′f_{12}
↑ = ↓
Df_{13}
a
(↑ DEF) = –
(↑ NUM) = SG
↑ = ↓
NPf_{14}
↑ = ↓
Nf_{15}
falafel
(↑ PRED) = ‘falafel’
(↑ NUM) = SG

The equations that represent lexical properties are parts of lexical entries. The lexical items in the sample sentence are:

(21)

a.	*the*	D	$(\uparrow \text{DEF}) = +$
b.	*a*	D	$(\uparrow \text{DEF}) = -$
			$(\uparrow \text{NUM}) = \text{SG}$
c.	*hamster*	N	$(\uparrow \text{PRED}) =$ 'hamster'
			$(\uparrow \text{NUM}) = \text{SG}$
d.	*falafel*	N	$(\uparrow \text{PRED}) =$ 'falafel'
			$(\uparrow \text{NUM}) = \text{SG}$
e.	*dinosaur*	N	$(\uparrow \text{PRED}) =$ 'dinosaur'
			$(\uparrow \text{NUM}) = \text{SG}$
f.	*will*	I	$(\uparrow \text{TENSE}) = \text{FUT}$
g	*give*	V	$(\uparrow \text{PRED}) =$
			give $\langle(\uparrow \text{SUBJ})\ (\uparrow \text{OBJ})\ (\uparrow \text{OBL}_{\text{Goal}}\ \text{OBJ})\rangle$'
h.	*to*	P	$(\uparrow \text{PCASE}) = \text{OBL}_{\text{Goal}}$

Most of the ID rules are straightforward as well. In (22), we annotate most of the rules from Chapter 2.

(22) a. Functional maximal projections

$$\text{CP} \rightarrow \underset{(\uparrow \text{FOC}) = \downarrow}{\text{XP}}, \underset{\uparrow = \downarrow}{\text{C}'}$$

$$\text{IP} \rightarrow \left\{\begin{matrix}\text{DP}\\ \text{NP}\\ \text{CP}\\ \text{PP}\end{matrix}\right\}, \underset{\uparrow = \downarrow}{\text{I}'} \quad (\uparrow \text{SUBJ}) = \downarrow$$

$$\text{DP} \rightarrow \underset{(\uparrow \text{POSS}) = \downarrow}{\text{DP}}, \underset{\uparrow = \downarrow}{\text{D}'} \quad (\uparrow \text{DEF}) = +$$

b. Functional single bar projection

$$C' \rightarrow \underset{\uparrow = \downarrow}{C} \quad , \quad \underset{\uparrow = \downarrow}{\left\{ \begin{matrix} S \\ IP \end{matrix} \right\}}$$

$$I' \rightarrow \underset{\uparrow = \downarrow}{I} \quad , \quad \underset{\uparrow = \downarrow}{VP}$$

$$D' \rightarrow \underset{\uparrow = \downarrow}{D} \quad , \quad \underset{\uparrow = \downarrow}{NP}$$

A complication arises when we consider adjuncts and coordinate structures. Recall from Chapter 1 that adjuncts are treated as elements of a set. Similarly, coordinated constituents can be treated as members of a set. We can use the mathematical symbol $\in$ 'is an element of'.

(23) a. $$XP \rightarrow \underset{\uparrow = \downarrow}{XP} \quad , \quad \underset{\downarrow \in (\uparrow ADJ)}{\left\{ \begin{matrix} PP \\ \left\{ \begin{matrix} AP \\ ADVP \end{matrix} \right\} \end{matrix} \right\}}$$, XP a lexical category

c. $$XP \rightarrow \underset{\downarrow \in \uparrow}{XP^{+}} \quad CONJ \quad \underset{\downarrow \in \uparrow}{XP}$$

There is another interesting complication in the ID rules introducing complements of lexical categories. Consider VP. As a first approximation, we can formulate the ID rule as follows.

(24)

$$VP \rightarrow \underset{\uparrow = \downarrow}{V} \quad , \quad \underset{(\uparrow OBJ) = \downarrow}{\left\{ \begin{matrix} DP \\ NP \end{matrix} \right\}} \quad , \quad \underset{(\uparrow OBJ2) = \downarrow}{\left\{ \begin{matrix} DP \\ NP \end{matrix} \right\}} \quad , \quad \underset{(\uparrow OBL_{Goal}) = \downarrow}{PP^{*}} \quad , \quad \underset{(\uparrow COMP) = \downarrow}{\left\{ \begin{matrix} CP \\ IP \\ S \end{matrix} \right\}}$$

We can simplify this somewhat. This rule has two nominal positions, for the functions OBJ and OBJ2. We can use the Kleene star to show that more than one nominal is possible in the VP, and that each one can be annotated with

any "object" function. The reason no more than two nominals are possible is that there are exactly two object functions. Temporarily lacking a feature system to express classes of grammatical functions (a gap we will fill in Chapter 4), we express the annotation as a disjunction.

(25) $$\text{VP} \rightarrow \text{V} \quad , \quad \left\{\begin{matrix}\text{DP}\\ \text{NP}\end{matrix}\right\}^* \quad , \quad \text{PP}^* \quad , \quad \left\{\begin{matrix}\text{CP}\\ \text{IP}\\ \text{S}\end{matrix}\right\}$$
$$\uparrow = \downarrow \quad \left\{\begin{matrix}(\uparrow \text{OBJ}) = \downarrow\\ (\uparrow \text{OBJ}2) = \downarrow\end{matrix}\right\} \quad (\uparrow \text{OBL}_{\text{Goal}}) = \downarrow \quad (\uparrow \text{COMP}) = \downarrow$$

But there is a problem with the PP complement: the grammatical function specified by the functional equation. While all PP complements (other than predicative complements, which we will discuss later) are some sort of thematically restricted OBL, they are not all restricted to being Goals. (In fact, multiple OBL_{Goal}s would be ruled out by the Uniqueness Condition.) One might think that we could leave the exact thematic role unspecified, as follows:

(26) $$\text{VP} \rightarrow \text{V} \quad , \quad \left\{\begin{matrix}\text{DP}\\ \text{NP}\end{matrix}\right\}^* \quad , \quad \text{PP}^* \quad , \quad \left\{\begin{matrix}\text{CP}\\ \text{IP}\\ \text{S}\end{matrix}\right\}$$
$$\uparrow = \downarrow \quad \left\{\begin{matrix}(\uparrow \text{OBJ}) = \downarrow\\ (\uparrow \text{OBJ}2) = \downarrow\end{matrix}\right\} \quad (\uparrow \text{OBL}_{\theta}) = \downarrow \quad (\uparrow \text{COMP}) = \downarrow$$

But that is not quite correct either: the thematic role to which the complement is restricted comes from somewhere! Where?

The thematic restriction of the oblique PP complement comes from the head P. The P carries a Case-like feature which has been dubbed PCASE. This feature was introduced in Chapter 1 and appears in our sample f-structure. It is the PCASE value of the oblique that determines exactly which oblique function it has; it does this by having, as its value, an oblique function name. So what we want is, in place of OBL_{Goal} or OBL_{θ}, a notation that means "the PP's PCASE value." The notation for that is:

(27) $(\downarrow \text{PCASE})$

If we insert this designator into the ID rule for VP, we get the following

result.

(28) $$\begin{array}{lcccccc} \text{VP} \rightarrow & \text{V} & , & \left\{\begin{array}{c}\text{DP}\\ \text{NP}\end{array}\right\}^* & , & \text{PP}^* & , & \left\{\begin{array}{c}\text{CP}\\ \text{IP}\\ \text{S}\end{array}\right\} \\ & \uparrow = \downarrow & & \left\{\begin{array}{c}(\uparrow \text{OBJ}) = \downarrow \\ (\uparrow \text{OBJ}2) = \downarrow\end{array}\right\} & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

This looks rather forbidding, but it expresses exactly what we want to express. The expansions of NP, AP, and PP will be similar.

(29) $$\begin{array}{lcccc} \text{NP} \rightarrow & \text{N} & , & \text{PP}^* & , & \text{CP} \\ & \uparrow = \downarrow & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

$$\begin{array}{lcccc} \text{AP} \rightarrow & \text{A} & , & \text{PP}^* & , & \left\{\begin{array}{c}\text{CP}\\ \text{IP}\\ \text{S}\end{array}\right\} \\ & \uparrow = \downarrow & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

$$\begin{array}{lcccccc} \text{PP} \rightarrow & \text{P} & , & \left\{\begin{array}{c}\text{DP}\\ \text{NP}\end{array}\right\} & , & \text{PP} & , & \left\{\begin{array}{c}\text{IP}\\ \text{S}\end{array}\right\} \\ & \uparrow = \downarrow & & (\uparrow \text{OBJ}) = \downarrow & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

3.3.5 Tying it all together

Functional equations, then, govern the mapping between c-structure and f-structure, and thus form an essential part of the LFG formalism. Much of LFG's descriptive power comes from functional equations.

Since LFG is a declarative theory rather than a derivational theory, the proper way to interpret the equations is as defining a well-formed mapping between an existing c-structure and an existing f-structure. However, the system can also be used to build an f-structure out of an existing c-structure, phrase structure rules, and lexical entries. Essentially, one follows our discussion backwards to do this:

① Annotate the appropriate phrase structure or lexical equations to each node in the c-structure tree.
② Assign each node in the tree an *f* variable.

③ Substitute *f* variables for ↑ and ↓ in each equation: that node's *f* variable for ↓ and the mother node's *f* variable for ↑. The resulting equations are the f-description.
④ Go through the f-description equation-by-equation, building the minimal f-structure consistent with the equations.

As we have seen, the *f* variables of the mapping formalism provide a way to refer to f-structure elements. The process of building an f-structure is easier with the variables annotated to the f-structure. However, f-structures are normally drawn without the variables. This poses a problem for lexical forms with arguments. Since these arguments are specified relative to their local f-structure, the lexical form has lexically specified ↑'s. The usual notation is to leave the ↑'s in the f-structure.

(30)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{'hamster'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{FUT} \\ \text{PRED} & \text{'give} \left\langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ})\ (\uparrow \text{OBL}_{\text{Goal}}\ \text{OBJ}) \right\rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & - \\ \text{PRED} & \text{'falafel'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{OBL}_{\text{Goal}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{Goal}} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{'dinosaur'} \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

3.3.6 Constraining equations, etc.

The functional equations we have discussed are **defining equations**. They define an f-structure attribute as existing and having a particular value. However, sometimes we want to require a particular feature to exist and have a particular value by virtue of a different part of the c-structure. For example, consider the following sentences.

(31) a. I helped the dinosaur [sing].
b. I kept the dinosaur [singing].

Each of the main verbs takes a verbal complement, the function of which we

will provisionally call VCOMP. However, the VCOMPs differ in their inflectional features: *help* takes the verb in its uninflected (bare infinitive) form, while *keep* takes a present participle.[8] Like all grammatical features, inflectional features are represented at f-structure. The lexical entry of *keep* will require its VCOMP to have the feature [PART PRES], while *help* will disallow inflectional features. As a first approximation, we could hypothesize that the lexical entry of *keep* includes the equation below.

(32) *keep*: (↑ VCOMP PART) = PRES

However, this equation does not do what we want. In particular, it does not rule out the following ungrammatical case.

(33) *I kept the dinosaur [sing].

This is because the equation in (32) defines a feature value for the VCOMP, whether it is supplied by the morphology of the subordinate verb or not. The ungrammatical sentence will have the feature [VCOMP [PART PRES]] by virtue of the equation.

Instead of a defining equation, what we want here is a **constraining equation**, an equation that requires a particular feature value to be present. Constraining equations are distinguished notationally by subscripting the letter *c* to the equal sign.

(34) *keep*: (↑ VCOMP PART) $=_c$ PRES

This specification will achieve the result we want.

A similar use of constraining equations involves the possessor in a DP. The possessor must be marked with the genitive Case marker. The following phrase structure rule insures this.

(35) $\text{DP} \rightarrow \underset{\substack{(\uparrow \text{POSS}) = \downarrow \\ (\uparrow \text{DEF}) = + \\ (\downarrow \text{CASE}) =_c \text{GEN}}}{\text{DP}}$, $\underset{\uparrow = \downarrow}{\text{D}'}$

[8]On the inflectional features used here, see the Appendix to this chapter.

Another place where constraining equations are needed is idioms. Consider the idiom *keep tabs on*. As discussed in Chapter 1, the noun *tabs* has a FORM feature:

(36) *tabs*: (↑ FORM) = 'tabs'
(↑ NUM) = PL

The verb *keep* (in this usage), has the following lexical entry:

(37) *keep*: (↑ PRED) = 'keep-tabs-on ⟨(↑ SUBJ) (↑ OBL_{on} OBJ)⟩ (↑ OBJ)'
(↑ OBJ FORM) $=_c$ 'tabs'

Unlike equality, inequality can only be understood as constraining. For example, consider the infl *was* in English: its SUBJ is singular and nonsecond person. This can be specified in the lexical entry as follows.

(38) (↑ SUBJ NUM) = SG
(↑ SUBJ PERS) ≠ 2

There is no way to interpret the inequality as defining a value; it simply constrains the feature not to have a particular value. Because there is no ambiguity, the *c* subscript is not used for inequality.

It is also possible for a lexical item to require a feature to be present without requiring that it have a particular value. For example, the complementizer *that* requires its clause to have a finite verb, i.e. a verb with the feature TENSE. The value of TENSE is irrelevant. This can be indicated as:

(39) *that* (↑ TENSE)

Conversely, the complementizer *to* cannot cooccur with the feature TENSE. It can include the following specification:

(40) *to* ¬(↑ TENSE)

Similarly, the verb *help*, discussed above, will have the following equations in its lexical entry to ensure that its VCOMP is uninflected.

(41) *help* ¬(↑ VCOMP TENSE), ¬(↑ VCOMP PART)

3.3.7 Outside-in vs. inside-out designators

In the previous section, we extended the formal system by adding constraining equations and designators stipulating existence. In this section, we will discuss one final addition to the formal system.

Consider the lexical item *what*. It is a *wh* word, but it differs from other *wh* words in an interesting respect. While most other *wh* words can introduce both questions and relative clauses, *what* can only introduce questions in standard English.

(42) a. [Who did you see]?
b. I asked [who you saw].
c. the syntactician [who you saw]

(43) a. [Where did you read that]?
b. I asked [where you read that].
c. the newspaper [where you read that]

(44) a. [What did you eat]?
b. I asked [what you ate].
c. *the falafel [what you ate]

We will assume that there is a feature for clause type, which we will call TYPE, with values such as Q, REL, and the like. The lexical item *what* must lexically specify that the clause in which it occurs has the feature [TYPE Q]. That is to say, the f-structure in which *what* appears must look like this (where DF is some discourse function):[9]

(45)
$$\begin{bmatrix} \text{TYPE} & \text{Q} \\ \text{DF} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{PRON} & \text{WH} \end{bmatrix} \end{bmatrix}$$

This differs from cases we have seen until now, because here the word *what* imposes a constraint on the larger structure in which it is found. From the perspective of c-structure, this constraint is bottom-up instead of top-down.

[9]We follow Butt, King, Niño, and Segond (1999) in hypothesizing that *wh* elements have a "pronoun type" (PRON) feature with the value WH.

Because of the orientation of f-structures, we cannot talk about top-down and bottom-up. Instead, we can refer to functional designators as **outside-in** (the f-structure equivalent of top-down) or **inside-out** (the f-structure equivalent of bottom-up). The designators with which we have dealt up to this point, designators like "(↑ SUBJ)", are outside-in. They define a path into the f-structure from a point specified by ↑.

In the present case, we need an inside-out designator to specify the f-structure containing *what*. The designator would say that starting at ↑ we move in a path outward past the function DF. The f-structure thus reached is the one to which we want to refer. The notation for inside-out designators is to put the ↑ at the end:

(46) (DF ↑)

We need to combine this with an outside-in designation of the TYPE feature. The full specification in the lexical entry of *what* is therefore as follows:

(47) ((DF ↑) TYPE) = Q

This specification will be part of the lexical entry, along with the equations specifying the value of PRED and PRON. Similarly, *which* will specify in its lexical entry that if it is in a relative clause it will have the [PRED 'PRO'] feature.

(48) ((DF ↑) TYPE) = REL ⇒ (↑ PRED) = 'PRO'

The use of inside-out designators is newer in the development of LFG than outside-in designators. It has transpired, however, that these designators are necessary to express certain syntactic phenomena. In this textbook we will be using them primarily in the analyses of long-distance dependencies ("*wh* movement") and anaphora. Another use for inside-out designators involves what is called **constructive morphology**. In constructive morphology, an inflectional morpheme imposes an existential constraint that the f-structure of which it is a part must bear a certain function in the larger f-structure. Metaphorically, the morpheme "constructs" its own context, whence the name. For example, in English the genitive suffix *'s* can only be used on a noun that heads a constituent bearing the grammatical function POSS. In fact, if one hears an utterance such as "*Spock's*", one knows, even without context, that Spock must be the possessor of something. We can

analyze the genitive suffix as adding the following equation to any noun to which it is affixed.

(49) (POSS ↑)

It also should be noted that inside-out designators are a natural concept in LFG. Recall that the grammar consists of constraints on structures, not instructions on how to build structures. An instruction view of grammar would be uneasy with combining outside-in and inside-out instructions. However, the static constraint view is perfectly consistent with the idea that some constraints go one way and other constraints go the other way.

3.4 More on c-structure and f-structure

In this chapter we have discussed the formal properties of c-structure and f-structure and the mapping between them. The formal system for c-structure–f-structure mapping imposes certain mathematical constraints on the nature of the mapping, and thus on the hypothesized processing involved. It also puts certain limits on the linguistic capabilities of the system of mapping. However, from the linguistic perspective it is not sufficiently constrained. It does not restrict the relations between specific c-structure configurations and specific f-structure elements.

Purely structural theories, such as those in the transformational tradition, essentially hold that there is a single universal mapping between constituent structure (LFG's c-structure information) and grammatical functions (LFG's f-structure information). The independent existence of f-structure was motivated in large part by the observation that this is false. However, we are now left with a system that allows any relation between c-structure and f-structure.

There are many potential relations between c-structure and f-structure that could be formalized in LFG but seem impossible. For example, there is no reason that there is no language in which the OBJ would occupy a higher structural position than the SUBJ. However, the available evidence strongly suggests that such a language is impossible. It is thus necessary, in addition to the formal system, to posit linguistic conditions on the mapping. To do this, we need to consider what kinds of c-structure–f-structure mappings might be plausible. We restrict our attention to configurational (endocentric) encoding of grammatical functions. Since, by definition, nonconfigurational (lexocentric) encoding does not mark grammatical functions by structural position, we would expect S to have greater freedom than the $\overline{\text{X}}$ categories.

Let us begin by considering the complement positions of lexical categories and functional categories. A lexical category like N or V is a meaningful item that has the potential to take arguments. Elements in structural complement positions fill these argument slots, or at least those assigned to complement functions. On the other hand, a functional category like D or I contributes features, but is not an argument-taking predicate. The NP which is a complement of D is not a complement in the functional sense. In f-structure, the D and the NP unify, with the D contributing one or more features to the joint f-structure correspondent of the D (and DP) and NP. In general, we can say that the structural complements of functional categories are co-heads, not functional complements. We can state these as conditions on annotations to phrase structure rules.

(50) a. In a phrase structure rule introducing a complement of a lexical head, annotate the phrase in complement position (↑ CPF) = ↓.
b. In a phrase structure rule introducing a complement of a functional head, annotate the phrase in complement position ↑ = ↓.

Next, let us consider overlay functions. Unlike argument functions, overlay functions relate not to the head predicate but rather connect the clause to other clauses in the sentence (SUBJ) and other sentences in the discourse (FOCUS, TOPIC). It is thus to be expected that, in a configurational structure, they will be structurally higher than complements. In fact, they occupy specifier position. We can thus add the following condition:

(51) In a phrase structure rule introducing a specifier (of a functional head[10]), annotate the phrase in specifier position (↑ OF) = ↓.

While this does not exhaust all the possibilities (we have not discussed adjuncts, for example), it is sufficient for our purposes. Note that under this proposal, it is impossible for an OBJ to be mapped to [SPEC, IP] or for a SUBJ to be mapped to a complement position in the VP.

We have assumed the validity of these conditions in our analysis of a sample sentence above. So, for example, we treated the modal *will* not as an

[10]Recall that we have adopted the working hypothesis that only functional categories have specifiers.

argument-taking predicate, but simply as contributing the feature [TENSE FUT] to the f-structure. This is consistent with the condition above, but for a contrary view see Falk (1984). Note, however, that the alternative involves the theory of control, which we have not dealt with yet. Another example is the oblique PP complement. In our analysis it bears the function OBL_{Goal} and the DP within it is an OBJ. This is the approach apparent, for example, in Bresnan (1982a) and in Kaplan and Bresnan (1982), although OBL_{Goal} is called TO there, and in Bresnan (2000). The other approach, as in Bresnan (1982b). For example, is that the DP is not an OBJ, but a co-head. This results in a flatter f-structure, and in an argument structure for *give* in which the final argument is simply ($\uparrow$ OBL_{Goal}) rather than ($\uparrow$ OBL_{Goal} OBJ). Again, the f-structure assumed here is more consistent with the proposed constraints.

The crucial point is that the hypothesis that grammatical configurations and grammatical functions are elements at distinct levels of representation does not entail that the mapping between them is completely free. A substantive linguistic theory will include linguistically based constraints.

3.5 Appendix: Verbal inflectional features

In our discussion, we have made certain implicit assumptions about features representing the inflectional properties of verbs. First, we have placed all inflectional features in f-structure, following the standard LFG analysis. Second, we have assumed the following set of features.[11]

(52) [TENSE PRES/PAST/FUT/POSSIBILITY/…]
[PART PRES/PAST]

For standard verb forms, this results in the following feature structures:

(53)

present tense:	[TENSE PRES]
past tense:	[TENSE PAST]
ing form	[PART PRES]
ed/*en* form:	[PART PAST]
infinitive	nothing

[11]Our treatment of modals is undoubtedly oversimplified. A better approach would have separate TENSE and MOOD features, rather than having features like [TENSE POSSIBILITY]. We will stick with the simplified version here, rather than explore the relation between TENSE and MOOD.

In this appendix, we will first compare our set of features with two others that have been proposed in LFG work. We will then briefly consider some problems that arise if all inflectional features are part of f-structure.

In Butt, King, Niño, and Segond (1999), a single set of features is proposed for grammars of English, French, and German. Simplifying slightly, they propose a semantically relevant feature TENSE[12] and the following morphologically relevant features:

(54) [FIN +/−]
[VFORM BASE, PERFP, PRESP]

The verb forms have the following feature realizations:

(55) present tense: $\begin{bmatrix} \text{TENSE} & \text{PRES} \\ \text{FIN} & + \end{bmatrix}$

past tense: $\begin{bmatrix} \text{TENSE} & \text{PAST} \\ \text{FIN} & + \end{bmatrix}$

ing form: $\begin{bmatrix} \text{VFORM} & \text{PRESP} \\ \text{FIN} & - \end{bmatrix}$

ed/*en* form: $\begin{bmatrix} \text{VFORM} & \text{PERFP} \\ \text{FIN} & - \end{bmatrix}$

infinitive: $\begin{bmatrix} \text{VFORM} & \text{BASE} \\ \text{FIN} & - \end{bmatrix}$

In this feature system, nonfinite verb forms have a single feature VFORM which specifies the inflectional properties of the verb, whereas the feature system used here distinguishes participles (which have the feature PART) from infinitives, which have no feature. In the system used here, the infinitive form represents something of a default form, with no inflectional properties per se. This reflects the fact that the infinitive form is used as a citation form for verbs, and also the fact that (in English, at least) the infinitive bears no inflectional affix. Our analysis also allows us to express the fact that participles behave as a class, both morphologically and syntactically. Morphologically, participles can be converted into adjectives while other verb forms cannot. We will return shortly to the syntactic evidence.

[12]This feature is actually part of a feature complex which also expresses mood and aspect.

A feature system slightly more similar to the one proposed here is that of Andrews (1994). Andrews uses the following features:

(56) [TENSE PRES/PAST]
[MOOD FIN/INF/PART]

This realizes the following six options.

(57) present tense: $\begin{bmatrix} \text{TENSE} & \text{PRES} \\ \text{MOOD} & \text{FIN} \end{bmatrix}$

past tense: $\begin{bmatrix} \text{TENSE} & \text{PAST} \\ \text{MOOD} & \text{FIN} \end{bmatrix}$

ing form: $\begin{bmatrix} \text{TENSE} & \text{PRES} \\ \text{MOOD} & \text{PART} \end{bmatrix}$

ed/*en* form 1: $\begin{bmatrix} \text{TENSE} & \text{PAST} \\ \text{MOOD} & \text{PART} \end{bmatrix}$

ed/*en* form 2: $\begin{bmatrix} \text{TENSE} & \text{PAST} \\ \text{MOOD} & \text{INF} \end{bmatrix}$

infinitive: $\begin{bmatrix} \text{TENSE} & \text{PRES} \\ \text{MOOD} & \text{INF} \end{bmatrix}$

Andrews identifies the two *ed*/*en* forms as the passive and the perfect respectively. That is to say, while he considers the passive form to be a kind of participle, he treats the perfect form as a kind of infinitive. His evidence for this comes from the syntactic properties of participles alluded to earlier. He points out that VPs in certain syntactic positions (such as certain adjuncts) must be participles.

(58) a. Thinking about his book, John walked into a puddle.
b. Insulted by a taxi-driver, John got flustered.

(59) a. People thinking about their books are often absent-minded.
b. People insulted by taxi-drivers often get flustered.

Finite VPs and infinitival VPs cannot appear in these positions. Strikingly, the *ed*/*en* forms must be passives, perfects are not permitted. On the other hand, a fronted perfect VP can take on the infinitival form.

(60) Buy/bought a house though they have, they still haven't solved their book-storage problems.

Andrews takes this to be evidence for his analysis. While the facts that Andrews cites are interesting, they may not be enough to argue for this split of what is always in English a single morphological form. Other factors, such as the semantics of the perfect forms, may conspire to require the auxiliary, thus blocking the participle from appearing alone. We will leave the ability of a fronted perfect VP to be an infinitive as an unsolved problem. We note in passing that it is a greater problem for a derivational theory in which the VP actually moves from one position to the other. We prefer the analysis which treats the infinitive as lacking in inflectional features.

The other assumption, that inflectional features are represented at f-structure, has turned out to be problematic when auxiliaries are analyzed in more detail. The problem is that, under the analysis we have adopted here, auxiliaries and main verbs are co-heads, so their f-structures merge. Consider the following sentence.

(61) Spock might have analyzed the data.

Under our feature system, this sentence would have the following f-structure.[13]

(62)
$$\begin{bmatrix} \text{SUBJ} & [\text{“Spock”}] \\ \text{TENSE} & \text{POSSIBILITY} \\ \text{ASP} & \text{PERF} \\ \text{PART} & \text{PAST} \\ \text{PRED} & \text{‘analyze } \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle \text{’} \\ \text{OBJ} & [\text{“the data”}] \end{bmatrix}$$

Note that there is no way to ensure that the verb form following *might* is an infinitive and the one following *have* is a past participle. If we were to include the specification ¬(↑ PART) in the lexical entry of *might* as a way of ensuring that the following verb be an infinitive, this sentence would be ruled out. Things get worse when more than one participle appears in the clause,

[13]We are assuming for concreteness that the auxiliary *have* introduces an aspectual feature [ASP PERF].

as in the following case.

(63) Spock might have been analyzing the data.

Here, the same f-structure would have the features [PART PAST] and [PART PRES]. It would therefore be inconsistent.

There are several possible approaches one could take. The one which currently commands something of a consensus is that there is another level of structure, called morphosyntactic structure, or **m-structure**, at which these inflectional features are represented. If this is correct, the kinds of lexical specifications we discussed for complement-taking verbs would be specifications of m-structure rather than f-structure. Another possibility is that our analysis of auxiliaries is incorrect, and each auxiliary is an argument-taking predicate.

Additional readings

One of the most interesting (and controversial) issues in the theory of grammatical functions is the status of SUBJ in languages that are different from English in certain ways (such as syntactically ergative languages). These languages suggest that SUBJ is not the unified function presented in this text or generally assumed in work in LFG and other frameworks. For two very different approaches, see Manning (1996) and Falk (1999, 2000).

The formal system for c-structure–f-structure mapping is detailed in Kaplan and Bresnan (1982), along with a discussion of its mathematical properties. These are further discussed in many of the articles in Dalrymple, Kaplan, Maxwell, and Zaenen, eds. (1995).

The well-formedness conditions on f-structure also originate in Kaplan and Bresnan (1982). They are further discussed in Bresnan (2000), where there is extensive discussion of the linguistic constraints on the c-structure–f-structure mapping.

Constructive morphology was introduced in Nordlinger (1998; 2000), based on an idea from Andrews (1996). An interesting application to English is the analysis of the "contraction" of *have* given in Barron (1998).

On m-structure, see Butt, King, Niño, and Segond (1999) and Frank (2000). The question of the representation of finiteness is discussed by Barron (2000). The analysis of auxiliaries which we have adopted, under which auxiliaries do not have a PRED feature, is motivated by Bresnan (2000), Schwartze (1996), Butt, King, Niño, and Segond (1999), and others, and is the generally accepted analysis. An analysis under which auxiliaries are argument-taking predicates with VP complements, which is similar to the HPSG analysis, was originally proposed by Falk (1984) and has been argued to be the correct analysis in Norwegian by Dyvik (1999).

We briefly considered coordination in this chapter, but there is much more to be said. Among the studies of coordination in LFG we can mention Kaplan and Maxwell (1988b),

Dalrymple and Kaplan (1997), and Sadler (1999).

Exercises

1. Each of the following is ungrammatical because it violates one or more well-formedness condition on f-structure. Indicate which well-formedness condition(s) each one violates. Explain.

   a. *We put the book.
   b. *The mice am in the cage.
   c *My pen, I like my writing implement.
   d. *Who do you think I saw you?
   e. *Some linguists think transformations.
   f. *The librarian read me the book to the kids.
   g. *I wrote a papers about auxiliaries.
   h. *The dinosaur spoke to the trilobite to the mastodon.
   i. **Seinfeld* was watched *Friends*.
   j. *This theory seems that LFG is constraint-based.

2. Assume the ID and LP rules developed in the last two chapters and the following lexicon fragment.

a:	D	(↑ DEF) = –
		(↑ NUM) = SG
dinosaur:	N	(↑ PRED) = ‘dino’
		(↑ NUM) = SG
gorilla’s:	N	(↑ PRED) = ‘gorilla’
		(↑ NUM) = SG
		(↑ CASE) = GEN
		(POSS ↑)
green:	A	(↑ PRED) = ‘green’
hamsters:	N	(↑ PRED) = ‘hamster’
		(↑ NUM) = PL

house:	N	(↑ PRED) = 'house' (↑ NUM) = SG
in:	P	(↑ PCASE) = OBL_{Loc}
might:	I	(↑ TENSE) = MIGHT
sell:	V	(↑ PRED) = 'sell ⟨(↑ SUBJ) (↑ OBJ2) (↑ OBJ)⟩'
sell:	V	(↑ PRED) = 'sell ⟨(↑ SUBJ) (↑ OBJ) (↑ OBL_{Goal} OBJ)⟩'
summer:	N	(↑ PRED) = 'summer' (↑ NUM) = SG
the:	D	(↑ DEF) = +
to:	P	(↑ PCASE) = OBL_{Goal}

Consider the following sentence:

> The hamsters might sell the gorilla's house to a green dinosaur in the summer.

Show how the formal system of LFG creates the structures for this sentence. Show:

- ☞ the annotated c-structure with f-variables on the nodes
- ☞ the f-description
- ☞ the creation of the f-structure: show what it looks like after every five equations in the f-description

4

Argument Structure

In this chapter we will consider the nature of argument structure and the mapping between the lexicon and the syntax.

4.1 Function-changing processes

Let us begin by considering the active-passive alternation.

(1) a. The hamster placed the cage in the garbage.
b. The cage was placed in the garbage by the hamster.

It is uncontroversial that this alternation is regular, and therefore must be governed by rules or principles of grammar. As we mentioned in Chapter 1, in LFG passivization is taken to be a lexical process.

The lexical status of passivization was discussed in Chapter 1. We saw that passivization serves as the input to derivational processes, specifically, conversion into an adjective. More generally, it can be shown that passivization cannot involve the movement of a DP from object position to subject position. (This argument is originally due to Grimshaw 1982a.) Consider the verb *capture*: it takes an OBJ, but not a COMP. The rules of English allow OBJ to be a DP or NP, but not CP. This DP OBJ can have propositional meaning, but it must be a DP. Since SUBJs can be either DP/NP or CP, the passive allows either as the argument corresponding to the OBJ.

(2) a. The theory captures the fact that passives are derived.
b. *The theory captures that passives are derived.
c. The fact that passives are derived is captured by the theory.
d. That passives are derived is captured by the theory.

Sentence (2d) cannot be derived in a movement theory, since its D-structure is ungrammatical.

If passivization is a lexical process, what that means is that active and passive verb forms are both listed in the lexicon. Consider an active-passive pair *place* and *placed*. Each is listed in the lexicon with its own properties. The property we are concerned with here is the verb's lexical form, the value of its PRED feature. They will be listed as follows.

(3) a. *place* 'place ⟨(↑ SUBJ) (↑ OBJ) (↑ OBL_{Loc})⟩'
b. *placed* 'place ⟨∅ (↑ SUBJ) (↑ OBL_{Loc})⟩'

That is to say, both express the same three-place predicate 'place', but with different mapping of the arguments. In the active, the first argument (the Agent) is mapped to the grammatical function SUBJ and the second (the Patient) is mapped to OBJ. In the passive, it is the second argument that is mapped to SUBJ. The first argument is unexpressed in the syntax.[1] In both, the third argument (Location) is mapped to OBL_{Loc}.

While both lexical forms are listed in the lexicon, they must be related to each other. This relation must be based on the mapping of arguments, and in LFG this means that it must be based on grammatical functions. Since the active mapping is the more basic one, we can see the passive lexical form as the result of a remapping operation. Using the mathematical symbol ↦ 'maps into', we can characterize passivization in English as follows.

(4) (↑ SUBJ) ↦ ∅
(↑ OBJ) ↦ (↑ SUBJ)
Morphology: participle

Apart from the morphology, this rule is a universal characterization of the passive. While languages differ in word order effects, Case and other morphological effects, etc., a rule of passive in terms of grammatical functions expresses what unifies passivization cross-linguistically.[2]

---

[1]We are following the usual assumption that the *by* phrase is an adjunct. If it is an argument, the lexical form of the passive verb will have the designator (↑ OBL_{by} OBJ) as an alternative to ∅.

[2]This argument was originally made in the framework of Relational Grammar by Perlmutter and Postal (1977).

The *capture* example discussed earlier can be explained under the lexical analysis. The lexical forms of the active and passive forms of *capture* would be as follows.

(5) a. active: 'capture ⟨(↑ SUBJ) (↑ OBJ)⟩'
b. passive: 'capture ⟨∅ (↑ SUBJ)⟩'

The second argument is lexically specified as OBJ in the active and SUBJ in the passive. As we observed above, this correctly prohibits the second argument from being a CP in the active but allows it to be one in the passive.

Other passivization facts in English also follow from this account. Consider the ditransitive construction and its passive.[3]

(6) a. Quark handed Morn a drink.
b. Morn was handed a drink (by Quark).
c. *A drink was handed Morn (by Quark).

This array of facts follows from the analysis that we have been assuming for the ditransitive: the immediately postverbal nominal bears the function OBJ and the second nominal bears the function OBJ2. Since it is the OBJ which is remapped to SUBJ by the rule of passivization, no more needs to be said about this construction.

Passivization is one of a class of constructions that can be characterized as function-changing. We will briefly consider another one, the (morphological) causative. The causative construction does not exist in English but does in many other languages. Consider the following examples from Turkish (taken from Aissen 1979: 8). In particular, note the disposition of the subject argument of the basic verb. (The morpheme glossed CAUS is the causative morpheme.)

(7) a. Hasan öl- dü.
Hasan die- PST
'Hasan died.'

[3]This is the situation for most contemporary varieties of English. There are some varieties of British English in which (6c) is grammatical.

b. Mehmet Hasan-ı öl- dür- dü.
Mehmet Hasan- ACC die- CAUS- PST
'Mehmet caused Hasan to die.'

(8) a. Kasap et- i kes- ti.
butcher meat- ACC cut- PST
'The butcher cut the meat.'

b. Hasan kasab- a et- i kes- tir- di.
Hasan butcher- DAT meat- ACC cut- CAUS- PST
'Hasan had the butcher cut the meat.'

As we can see, if the base verb is intransitive its subject is realized as OBJ in the causative, while the subject of a transitive verb is realized as a dative OBJ2.[4] While this is not the only pattern that one sees with causatives cross-linguistically, it is a very common one. As with the passive, it can be formalized as a remapping rule.[5]

(9) a. $(\uparrow \text{SUBJ}) \mapsto (\uparrow \text{OBJ})$
b. $(\uparrow \text{SUBJ}) \mapsto (\uparrow \text{OBJ2})$

This remapping rule will be associated with the appropriate morphology and with the addition of a new Causer argument which gets mapped to SUBJ. While this formulation is incomplete, it illustrates the basic point that causativization can be treated formally in the same way as passivization.

4.2 Problems with the remapping analysis

The grammatical-function based remapping analysis outlined in the previous section is the earliest LFG analysis of passive, as in Bresnan (1982a). It is

[4]Actually, what we can see is that the Case marking is accusative and dative, respectively. These Cases are typical of the functions OBJ and OBJ2 in Turkish, and following much of the literature on the subject, we assume that this is the correct functional analysis of Turkish causatives.

[5]Following Grimshaw (1982b), we state it as a pair of disjunctively ordered rules. If there already is an OBJ in the lexical form of the verb, the a. version will be unable to apply and the b. version will apply instead. If the a. version can apply, the b. version is skipped.

important to note that a lexical rule of this kind is not monotonic: it takes existing information and changes it. This is ruled out in principle in the syntax on grounds of processing: syntactic information cannot be changed. But a lexical rule is not a syntactic rule. Lexical rules do not represent on-line processing, but rather regularities relating stored lexical items. When a lexical rule is applied productively, the result is stored as a new lexical item. For this reason, the usual LFG constraint against changing information is inapplicable here.

Although the remapping analysis does not violate any processing-based constraints, it transpires that there are problems with it. The problems arise primarily from the descriptive linguistic perspective. We will discuss some of these here, and show how they point to a better analysis of the passive and other function-changing constructions.

One of the earliest sources of evidence that something is amiss came from an examination of passivization in ditransitive constructions in Bantu languages. Bantu languages have a more extensive ditransitive construction than English, involving a wider class of elements than the English Goal and Beneficiary . As in English, the Theme argument is realized as OBJ2 and the other argument (often called the applied argument) as the full OBJ. In some of the Bantu languages, such as Chicheŵa, the passive works the same as in English, and as predicted by the remapping analysis: the non-Theme argument becomes SUBJ, as shown by the following example from Alsina and Mchombo (1993).[6]

(10) a. Chitsîru chi- na- gúl- ír- a atsíkána
7.fool 7SUBJ- PST- buy- APPLIC- VWL 2.girls
mphâtso.
9.gift
'The fool bought the girls$_{OBJ}$ a gift$_{OBJ2}$.'

b. Atsíkāna a- na- gúl- ír- idw- á mphâtso.
2.girls 2SUBJ- PST- buy- APPLIC- PASS- VWL 9.gift
'The girls were bought a gift.'

[6]In the Bantu glosses, the number before each noun represents the noun class to which it belongs. The agreement marker on the verb agrees with this noun class.

c. *Mphâtso i- na- gúl- ír- idw- á átsíkāna.
9.gift 9SUBJ- PST- buy- APPLIC- PASS- VWL 2.girls
'A gift was were bought the girls.'

Languages of this kind have been dubbed **asymmetrical**. However, in other languages, **symmetrical** languages such as Kichaga and Kinyarwanda, either the Theme (OBJ2) or non-Theme (OBJ) argument can become SUBJ. Note the following example from Kichaga (from Bresnan and Moshi 1990).

(11) a. N- ā- ī- lyì- í- à m̀- kà k- élyà.
FOC- 1SUBJ- PRES- eat- APPLIC- VWL 1- wife 7- food
'He is eating food for/on his wife.' (Lit. 'He is eating his wife_{OBJ} $\text{food}_{\text{OBJ2}}$.')

b. M̀- kà n- ā- ī- lyì- í- ò k- élyà.
1- wife FOC- 1SUBJ- PRES- eat- APPLIC- PASS 7- food
'The wife is being benefitted/adversely affected by someone eating the food.' (Lit. 'The wife is being eaten the food.')

c. K- élyà k- ī- lyì- í- ò m̀- kà.
7- food 7SUBJ- PRES- eat- APPLIC- PASS 1- wife
'The food is being eaten on/for the wife.' (Lit. 'The food is being eaten the wife.')

This is contrary to what is predicted by the classical LFG analysis. Furthermore, this difference seems to be something deeper than just a difference in passivization: the omissibility of an object argument is related to its ability to become SUBJ under passivization. The conclusion is that something is different about the nature of the object arguments in the two types of languages, and that this difference is reflected both in the differing passivization properties and in omissibility.

Another kind of empirical problem with the remapping theory of function-changing constructions comes from the Romance languages. These languages have a causative construction that seems to operate like the causative in Turkish, that is to say the SUBJ of the base verb becomes either an OBJ or an OBJ2 depending on whether the base verb is transitive. This is illustrated in the following examples from French (from Grimshaw 1982b).

(12) a. On fera parler Jean de son dernier voyage.
one CAUS.FUT talk John of his last journey
'One will make John talk of his last journey.'

b. Elle a fait visiter la ferme à ses parents.
she has CAUS.PART visit the farm DAT her relatives
'She made her relatives visit the farm.'

However, as shown by the French orthography, the causative "morpheme" is really a separate word; in fact, it is a main verb with an infinitival clause in complement position. Therefore, unlike the Turkish case, Romance causatives cannot be analyzed lexically.[7] Instead, in French there is a combination of two distinct lexical predicates into a single predicate in the f-structure (i.e. in the syntax). Under LFG assumptions about processing, a remapping analysis is impossible in the syntax. Constructions of this kind have come to be called **complex predicates**. Complex predicates thus form a second empirical problem with the remapping analysis of function-changing constructions.

It can also be objected that the remapping analysis is not explanatory. An explanatory theory is one that predicts what is a possible rule and what is not. Now consider the following two remappings:

(13) a. $(\uparrow \text{OBJ}) \mapsto (\uparrow \text{SUBJ})$
b. $(\uparrow \text{SUBJ}) \mapsto (\uparrow \text{OBJ})$

The remapping (13a) is part of passivization, and that in (13b) is part of causativization. Passivization and causativization are different kinds of processes: passivization is simply a remapping of arguments, while causativization involves embedding the original verb under an additional predicate. In the languages of the world, there are no simple remapping rules like passive in which the remapping in (13b) takes place. There is nothing in

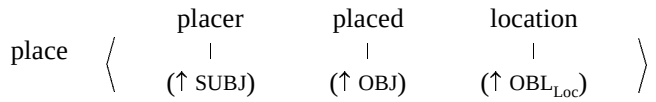
[7]This is perhaps too strong a statement. The original LFG analysis, in Grimshaw (1982b), was lexical, and a lexical approach has been argued to be correct by Frank (1996). Under a lexical analysis, the VP in complement position would be annotated $\uparrow=\downarrow$ and the causativized lexical form of the verb would require the 'cause' verb as co-head. However, the position taken in the text is the usual one in LFG, and is a priori preferable because it reflects the c-structure reality of two separate verbs. Ultimately, though, this is an empirical question which has not yet been resolved.

the remapping analysis that would exclude (13b). A theory of rules like passivization should predict this.

Furthermore, and somewhat surprisingly, while early LFG had a theory of *re*mapping, it had no theory of the initial mapping. A lexical form such as (14a) represents the syntactic expression of the thematic (semantic/conceptual) arguments of the predicate. This can be expressed informally as (14b).

(14) a. 'place ⟨(↑ SUBJ) (↑ OBJ) (↑ OBL_{Loc})⟩'

b.



The nature of this mapping was not considered in early LFG. This lack of a theory of the mapping of arguments was problematic for three reasons: first, because this is a central lexical process of importance to the syntax; second, because the same mapping is found in all languages, and is therefore not arbitrary; and third, because resolving the nature of the mapping is a potential direction for reconsidering *re*mapping operations such as passivization.

In the late 1980s and early 1990s, **Lexical Mapping Theory (LMT)** was developed to address these issues. The basic idea behind LMT is a(rgument)-structure, a representation of the syntactic arguments of a predicate. A-structure is the locus of the mapping between thematic roles and grammatical functions. The mapping, as we will see, is monotonic. This allows some of the mapping to be done on-line in the syntax without violating computational constraints. Thus, complex predicates can be incorporated within LMT.

Unlike the mapping between c-structure and f-structure, LMT is not purely syntactic in nature. It maps from a semantic/conceptual representation of thematic roles, henceforth "θ-structure", to a syntactic representation of grammatical functions, f-structure, via an intermediate lexical representation called a-structure. LMT must thus interact with a theory of θ-structure and come to grips with issues of modularity: the degree to which representations of different dimensions of linguistic structure interact with each other and the degree to which they are isolated from each other. The nature of θ-structure is orthogonal to the primary questions addressed by LFG, different approaches have been taken by different researchers within the LMT literature.

We will outline one such approach here. The question of the connections between the various dimensions of structure relates to the nature of the LFG projection architecture, and there are still many open questions in that area.

4.3 Lexical Mapping Theory

4.3.1 Thematic structure

LMT is about the syntactic realization of the arguments of a predicate. Argumenthood is a semantic/conceptual concept: arguments fit empty positions in the meaning of a predicate. Arguments can be identified by their role in the predicate's meaning; in recent generative work such roles have come to be called **thematic roles** or **θ roles**. As generally used by syntacticians, thematic roles are vague labels (Agent, Patient, Theme, Goal, Source, Experiencer, etc.) used for an imprecise characterization of conceptual roles. In this section, we will outline in an informal way a more precise theory of thematic roles, based on Jackendoff (1990), but without Jackendoff's formalism. Many (though by no means all) researchers in LFG have adopted some version of Jackendoff's approach.

According to Jackendoff, there are two aspects to the way we conceptualize the meanings of verbs. One such conceptualization is that of an action, involving an entity that acts and/or one that is acted on or affected. These two entities are called Actor and Patient (or Undergoer) respectively. An Undergoer which is positively affected is often called a Beneficiary. The other conceptualization is a spatially-based one. We conceptualize elements in terms of location or movement, either in physical space or in some abstract space such as possession, physical properties, or time. The thing moving or being located is called the Theme, and it is either placed at a Location or moved along a Path. A Location or Path (the latter sometimes specifying Source and/or Goal) is usually specified in terms of some sort of Reference Object. Sometimes the movement is initiated by an element other than the Theme; this external initiator can be called the Instigator.Although the Instigator is not strictly speaking a spatial participant in the action, it can be considered to be part of the spatial conceptualization because it sets the moevement in motion.[8]

We will make this concrete with a few examples. Consider the transitive

[8]This is a simplification of the very complex subject of the nature of causation. See Jackendoff and references cited there for more detailed discussion.

verb *roll* as in:

(15) a. Sisko rolled the ball to second base.
b. Sisko rolled the ball from his office.
c. Sisko rolled the ball along the baseline.

As an action, in each of these cases Sisko is acting on the ball. We can show this by making the notion of one entity acting on the other explicit:

(16) What Sisko did to the ball was roll it to second base.

Sisko is therefore the Actor and *the ball* is the Patient. In spatial terms, Sisko is the cause of the spatial event of the ball moving along a path; *Sisko* is the Instigator, *the ball* is the Theme, and the objects of the prepositions are the Reference Objects of the Paths expressed by the prepositions. The Reference Object in (15a) is a Goal and the one in (15b) is a Source. The conceptualization of the verb thus involves three arguments:

(17) roll: [Actor/Instigator] … [Patient/Theme] … [Path]

On the other hand, note the following sentence:

(18) Sisko rolled down the field intentionally.

Here, there is no Instigator: Sisko is Theme and Actor.

Next, consider the transitive *melt*:

(19) The phaser melted the metal.

As an action, the phaser is acting on the metal.

(20) What the phaser did to the metal is melt it.

The phaser and metal are thus Actor and Patient respectively. In (pseudo-) spatial terms, the metal is "moving" along a path from solidness to liquidness. The "motion" is caused by the phaser. Therefore, *the phaser* is the Instigator and *the metal* is the Theme. The Path is not grammatically expressed; rather, it is semantically part of the meaning of the verb.

(21) melt: [Actor/Instigator] … [Theme] … [Path=from solid to liquid]

Finally, consider the verb *feed*.

(22) Quark fed his customers cheap food.

Here, Quark is affecting his customers.

(23) What Quark did to his customers was feed them cheap food.

Thus, *Quark* is the Actor and *his customers* is the Patient. Spatially, there is amovement of the food to the customers: literally spatially in the sense that the food is actually placed in front of the customers, and also in the pseudo-spatial field of possession, since the food is transferred to the customers' ownership. This spatial event of the food moving to the customers is initiated by Quark. Thus, *Quark* is the Instigator, *the food* is the Theme, and *the customers* are the Goal (or the Reference Object of a Path).

Under this view, thematic relations are not primitives of linguistic theory; they are descriptions of certain aspects of cognitive conceptualization. In the full version of the theory, there is a level of conceptual structure at which these notions are represented. A "θ-structure" with role labels is a simplification. It is crucial for a theory of argument structure that thematic roles not be opaque primitives to be manipulated by the syntactician; when this happens, the theory becomes vacuous.

The examples that we have given are interesting in terms of the relationship between the action conceptualization and the spatial conceptualization. In all three cases, where both the Actor and Instigator roles are present, the same element serves as both. This is not always the case: it is possible for a Theme to be an Actor, as in *I rolled down the hill on purpose*. However, whenever there is an Instigator, it is also the Actor. For this reason, we will not treat Actor and Instigator as distinct thematic roles; instead, following common usage, we will refer to an Actor/Instigator as an Agent. On the other hand, while Patient and Theme often correspond, *feed* shows that they need not. In much of the literature, the terms Patient and Theme are used interchangeably; from the perspective of the thematic theory outlined here, this is a mistake.

In fact, Jackendoff suggests that the same spatial conceptualization can be given different action conceptualizations. Consider the concept of transferring possession (i.e. instigating movement in the possessional field)

to someone. Is the affected entity (Patient or Beneficiary) the thing being transferred (Theme) or the person receiving it (Goal)? Both conceptualizations are plausible. Jackendoff argues that these differing conceptualizations are what lie behind the ditransitive (or "dative") alternation.

(24) a. $\text{Quark}_{\text{Agent}}$ handed a $\text{drink}_{\text{Patient/Theme}}$ to $\text{Morn}_{\text{Goal}}$.
b. $\text{Quark}_{\text{Agent}}$ handed $\text{Morn}_{\text{Beneficiary/Goal}}$ a $\text{drink}_{\text{Theme}}$.

Finally, we note that this approach to thematic roles predicts a hierarchical relation between them. In the action conceptualization, the Actor has primacy over the Patient/Beneficiary; where both are present, the Patient or Beneficiary is affected as a result of something the Actor does. In the spatial conceptualization, the Instigator has primacy over the Theme, which in turn is more prominent than the Path, Location, or Reference Object. If we furthermore stipulate that the action conceptualization takes priority over the spatial, we derive the following **thematic hierarchy**.

(25) Agent > Patient/Beneficiary > Theme > Path/Location/Ref Obj

Such hierarchies have played a prominent role in theories of thematic roles and argument mapping. One advantage of a theory such as Jackendoff's is that the hierarchy is not an independent stipulation of the theory, but a consequence of the conceptualization. We have not yet discussed Instruments; we will assume that they go between Patient/Beneficiary and Theme. (For discussion of the conceptual status of Instruments, see Jackendoff 1990.) Our thematic hierarchy is therefore:

(26) Agent > Patient/Ben > Instrument > Theme > Path/Location/Ref Obj

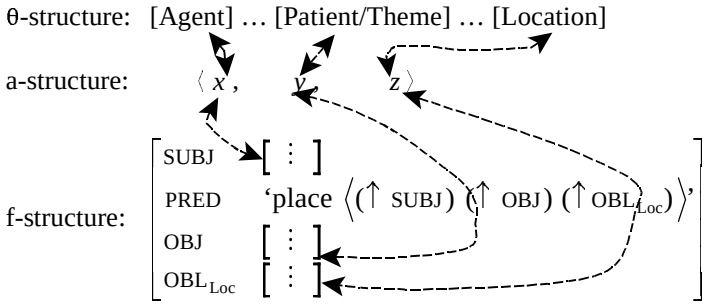
The thematic hierarchy has been implicated in other grammatical phenomena as well. For example, as discussed by Bresnan (2000), idiom chunks can be Locations or Themes, but are generally not higher up on the thematic hierarchy.

4.3.2 Mapping

Now that we have a rudimentary theory of θ-structure, we can discuss the place of a-structure in LFG and LMT. Like "argument structure" or "θ grid" in GB theory, a-structure is a representation of the syntactic argument-taking properties of a lexical item. Just as the formal system introduced in the

previous chapter maps between the syntactic levels of c-structure and f-structure, LMT maps between θ-structure and a-structure[9], and between a-structure and f-structure.

(27) *place*:



A-structure differs from θ structure in several ways. In the first place, it abstracts away from the conceptual structure details of thematic roles, representing simply the thematic hierarchy and, as we will see shortly, providing a special representation for Patients and Themes. It is a more strictly linguistic representation than the lexical conceptual structure of which θ-structure is a simplification. Second, it defines what grammatical functions each argument can be potentially mapped to. Third, as a syntactic representation, it only deals with syntactically relevant aspects of θ structure and is the locus of constraints. For example, as we will see, the distinction between symmetrical and asymmetrical languages is expressed in terms of an a-structure condition. Nonthematic arguments, such as expletives and idiom chunks, are also represented at a-structure, although they naturally have no role to play at conceptual structure.

The basic idea of LMT is that there is a classificatory system for argument grammatical functions, and that this classification is the basis for the syntactic mapping of thematic roles. The essential observations are the following:

- Themes and Patients are mapped to either SUBJ (in passives and

[9]The term "a-structure" is sometimes used in a sense closer to what we are calling θ-structure. The term seems to be used ambiguously by Bresnan (2000).

unaccusatives) or OBJ (in transitives)

- Non-Theme/Patient arguments cannot be mapped to OBJ. They may be mapped to SUBJ (Agents, Locations, possibly Instruments) or OBL_θ (Locations, Instruments, Sources, Goals, etc.)
- In verbs with more than one Theme/Patient argument, the "secondary" one (defined to some extent on a language-particular basis) is mapped either to OBJ or to "OBJ2." Since "OBJ2" is characteristically restricted to a single thematic role (non-Patient Theme in English) or a small set of thematic roles (such as below a particular position on the thematic hierarchy), a better designation for the "secondary object" function is OBJ_θ[10], and we will henceforth use this name for the function. The mapping of "secondary" Themes and Patients to OBJ_θ is apparently a marked option: not all languages have OBJ_θ, and even in those that do non-Patient Themes are only mapped to OBJ_θ if there is also an OBJ.

These three mappings are summarized in the following table; there is no generalization for the fourth grouping.

(28)

	non-Theme/Patient	secondary Theme/Patient
Theme/Patient	SUBJ	OBJ
	OBL_θ	OBJ_θ

This classification also makes sense from the perspective of characterizing the argument functions themselves. The functions SUBJ and OBJ are distinguished by not being inherently linked to thematic roles; each can represent a variety of thematic roles, or even non-thematic (expletive) arguments. The following are examples of expletive *it* in SUBJ and OBJ

[10]This notation is preferable for another reason as well. It predicts that a language might have multiple "secondary objects" with different thematic roles in the same sentence, since they would be individuated by the thematic role subscript. If there is a single "secondary object" function, Uniqueness/Consistency would rule out multiple secondary objects. In some Bantu languages this prediction seems to be correct. Much of the early work on LMT was carried out with special reference to Bantu languages, particularly Chicheŵa and Kichaga. See the references cited in the Additional Readings for this chapter.

position.

(29) a. *It* is obvious that the world is flat.
b. I take *it* that the world is flat.

On the other hand, each member of the OBL_θ family is explicitly marked (by Case or preposition) for its thematic role and the OBJ_θ family is restricted by language-specific constraints to a small set of thematic roles. This is the property of **restrictedness**: SUBJ and OBJ are unrestricted, while OBJ_θ and OBL_θ are restricted. This can be formalized as a binary distinctive feature [±*r*]. On the other hand, OBJ and OBJ_θ are both object functions, while SUBJ and OBL_θ are not. These two pairs of functions can be distinguished by a feature [±*o*].

(30)

	[−*o*]	[+*o*]
[−*r*]	SUBJ	OBJ
[+*r*]	OBL_θ	OBJ_θ

A-structure is a list of the syntactically expressed arguments of a predicate with partial specification of the grammatical realization. The a-structure is organized in terms of the thematic hierarchy, from most prominent argument to least prominent argument. The thematically most prominent argument is called $\hat{\theta}$ (**theta-hat**). The θ-structure to a-structure mapping principles (sometimes called the "intrinsic classification" of arguments) are listed in (31).

(31) **θ-structure to a-structure mapping**
Patients and Themes map to [−*r*]
"secondary" Patients and Themes map to [+*o*] as a marked option
non-Theme/Patient arguments map to [−*o*]

The characterization of "secondary" Patients and Themes differs cross-linguistically. In English, it is Themes which are not also Patients; in Bantu languages (Alsina and Mchombo 1993) it is any applied argument at the lower end of the thematic hierarchy.

The a-structure of our sample verb *place* is (32).

(32) $\langle [-o], [-r], [-o] \rangle$

This is the lexical representation of the arguments. These arguments have to map into f-structure functions. The Agent argument, represented in the a-structure as $[-o]$ and (by virtue of its position adjacent to the left bracket) $\hat{\theta}$, must map to SUBJ. This is achieved by adding the additional feature specification $[-r]$. The Patient/Theme argument, represented as $[-r]$, maps to OBJ by adding the feature $[+o]$. And the Location argument, $[-o]$, maps to OBL_{Loc} by adding the feature $[+r]$.

(33) **a-structure to f-structure mapping**

SUBJ Mapping 1:	A $[-o]$ argument which is $\hat{\theta}$ maps to SUBJ
SUBJ Mapping 2:	$[-r]$ may map to SUBJ.
NonSUBJ Mapping:	Add positive values of features where possible.

In addition, the following conditions apply.

(34) a. **Function-argument biuniqueness**
Each a-structure role corresponds to a unique f-structure function, and each f-structure function corresponds to a unique a-structure role.

b. **The Subject Condition**
Every verb must have a SUBJ.

With the verb *place*:

- the $\hat{\theta}/[-o]$ argument (the Agent) maps to SUBJ by SUBJ Mapping 1.
- the $[-r]$ argument (Theme/Patient) maps to OBJ by NonSUBJ Mapping, which adds the feature $[+o]$; although $[-r]$ is allowed to map to SUBJ (by SUBJ Mapping 2) doing so here would lead to a violation of Function-argument biuniqueness, since it would result in two a-structure roles mapping to the same function. The feature $[+r]$ cannot be added because it would contradict the existing value for the feature $[r]$.
- the other $[-o]$ argument (Location) is subject to nonSUBJ mapping, which adds the feature $[+r]$. The resulting grammatical function is OBL_{θ}, more specifically, OBL_{Loc}.

The a-structure to f-structure mapping is reflected in the PRED value of the

verb as it appears in f-structure.[11]

The mapping from θ-structure to a-structure to f-structure functions can be summarized as follows.

(35)

place ⟨ Agent, Patient/Theme, Location ⟩

a-structure:	[− *o*]	[− *r*]	[− *o*]
	|	|	|
f-structure:	S	O	OBL_{θ}

It should also be noted that some verbs involve idiosyncratic mapping of arguments. This is why some verbs, such as *enter* and *visit*, have OBJs which are neither Themes nor Patients, but rather Locations. Such verbs would have the Location argument idiosyncratically mapped to [− *r*]. Once this idiosyncratic mapping is established, these verbs behave like ordinary transitive verbs; their atypical conceptual structure is not accessible to the syntax.

LMT provides us with a monotonic account of the mapping from thematic roles to syntax. Recognizing a-structure as a distinct dimension of linguistic representation, intermediate between the thematic roles of lexical conceptual structure and the grammatical functions of f-structure allows us to express generalizations in terms of it. As we will see in the coming sections, it provides a basis for accounting for unaccusative phenomena and passivization without resorting to grammatical function changing processes. The difference between asymmetrical languages (such as English and Chicheŵa) and symmetrical languages (like Kichaga) will be expressed later in this chapter in terms of a condition on a-structure. And since it is a-structure arguments that are mapped into the syntax proper (more specifically f-structure), it can be used as the basis for the analysis of complex predicate constructions like the causative, in which two predicates are merged, even when the merger takes place syntactically as in French. The

---

[11]Technically, the PRED value as represented in the verb's lexical entry should reflect the underspecified values of a-structure, but we will follow the usual LFG practice and use fully specified functions. Thus, instead of the more accurate (ia), we will continue to use (ib).

(i) a. (↑ PRED) = 'place ⟨(↑ [− *o*]) (↑ [− *r*]) (↑ [− *o*])⟩'

b. (↑ PRED) = 'place ⟨(↑ SUBJ) (↑ OBJ) (↑ OBL_{Loc})⟩'

appendix to this chapter outlines an account of French causatives in terms of LMT.

4.3.3 Unergatives and unaccusatives

It is widely accepted that languages display a contrast between two types of intransitive verbs, generally called unergatives and unaccusatives. Unergatives are (roughly) verbs whose sole argument is agentive or volitional, while intransitives whose sole argument is (roughly) patientive are unaccusative. In many languages, the arguments of unaccusative verbs display behavior that is otherwise typical of OBJ rather than SUBJ.

While English is not blessed with "unaccusative phenomena," evidence of unaccusativity has been claimed. One construction that seems to distinguish between unaccusatives and unergatives is the resultative construction. Resultatives in transitive clauses refer to the OBJ, not the SUBJ.

(36) a. They wiped the table clean.
b. *They wiped the table tired.

As expected, the SUBJ of an unergative verb cannot control a resultative.

(37) a. *We danced into a frenzy.
b. *We laughed under the table.

Surprisingly, the SUBJ of an unaccusative can control a resultative.

(38) a. The river froze solid.
b. My son grew tall.

It is generally believed that this requires a syntactic explanation in terms of the unergative/unaccusative distinction. This view is not universal; it has been argued that the correct explanation is semantic (Jackendoff 1990, Rappaport-Hovav and Levin 2000). It is premature to tell whether a semantic analysis will turn out to be superior; we will pursue the more conventional syntactic analysis here.

Derivational and multistratal theories account for such facts by treating unaccusatives as taking initial OBJ rather than SUBJ (Perlmutter 1978, Burzio 1986). Such treatment makes it possible to account for OBJ-like behavior of the unaccusative argument. Unaccusative phenomena have therefore been considered evidence for derivational theories. However, LMT allows us to

express the properties of unaccusatives without hypothesizing an additional stratum of grammatical functions, by distinguishing the arguments of unergatives and unaccusatives at a-structure. The argument of an unergative is [− *o*] at a-structure, while the argument of an unaccusative is [− *r*].

(39) a. *dance* ⟨ Agent ⟩

a-structure:	[− *o*]	non-Patient/Theme
	|	
f-structure:	S	SUBJ Mapping 1

b. *grow* ⟨ Theme ⟩

a-structure:	[− *r*]	Patient/Theme
	|	
f-structure:	S	SUBJ Mapping 2

The resultative construction can be formalized as an operation on a-structure.

(40) Add a predicative argument with resultative semantics predicated of [− *r*].

Unaccusative phenomena thus provide evidence for a-structure as a separate level of representation in LFG. This is the only syntactic level at which unaccusative and unergative arguments can be distinguished.

4.3.4 Passives and ditransitives

We now return to where we started this chapter: the analysis of the passive construction. As we have seen, passivization involves a change in the mapping of arguments to syntax. Therefore, it is natural to treat is as an a-structure operation.

Passivization can be characterized in terms of a-structure as follows.

(41) Do not map $\hat{\theta}$ to the syntax. (Often written: $\hat{\theta} \mid \varnothing$)

This nonmapping of an argument can be called **suppression**. Suppression of $\hat{\theta}$ is associated in the lexicon with whatever passive morphology is characteris-

tic of the language. In English, this is participial morphology. Note the effect of suppression on the mapping to f-structure.

(42)

	place ⟨	Agent,	Patient/Theme,	Location ⟩
a-structure:		[− *o*]	[− *r*]	[− *o*]
		∅		
			|	|
f-structure:			S	OBL_{θ}

The resulting lexical form is the intended passive one. The [− *r*] argument becomes the SUBJ, even though it is not $\hat{\theta}$, as a result of the Subject Condition; if it were to be realized as an OBJ, there would be no SUBJ. The a-structure status of the argument is the only condition on being mapped to SUBJ; it can have any thematic role or (as in the case of idiom chunks) no thematic role at all.

An interesting question that arises concerning the LMT analysis of passive is the interaction between passivization and the ditransitive construction, as discussed earlier in this chapter. Consider again the following alternation in English, with thematic roles marked following Jackendoff's analysis mentioned earlier.

(43) a. $Quark_{Agent}$ handed a $drink_{Patient/Theme}$ to $Morn_{Goal}$.
b. $Quark_{Agent}$ handed $Morn_{Beneficiary/Goal}$ a $drink_{Theme}$.

The monotransitive version (43a) is unremarkable. But consider the ditransitive (43b). Recall that a Beneficiary is similar to a Patient, and thus higher on the thematic hierarchy than the Theme. The θ-structure thus consists of an Agent, a Beneficiary/Goal, and a Theme, in that order. The LMT mapping principles would then operate to derive an a-structure. The Agent, being neither a Patient nor a Theme, maps to [− *o*]. The Beneficiary, being a kind of Patient, maps to [− *r*]. And the Theme can, apparently, map either to [− *r*] (by virtue of being a Theme) or, less preferably, to [+ *o*] (by virtue of being a "secondary Theme," defined for English as a non-Patient Theme). It appears, then, that two a-structures are possible in principle.

(44) a. preferred

	hand ⟨	Agent,	Ben/Goal,	Theme ⟩
a-structure:		[− *o*]	[− *r*]	[− *r*]

b. dispreferred

	hand ⟨	Agent,	Ben/Goal,	Theme ⟩
a-structure:		[− *o*]	[− *r*]	[+ *o*]

However, only the second of these can map to a well-formed f-structure lexical form. The [− *o*] argument is $\hat{\theta}$, and therefore must be mapped to SUBJ ([− *o*, − *r*]). This leaves only one additional [− *r*] function (OBJ). There can be no grammatical mapping of the a-structure with two [− *r*]'s (44a). The other a-structure (44b) will map successfully, with the [− *r*] Beneficiary/Goal argument marked [+ *o*] (OBJ) and the [+ *o*] Theme marked [+ *r*] (OBJ$_\theta$).

Consider, however, what would happen if we were to try passivizing the each of the a-structures in (45).

(45) a.

	hand ⟨	Agent,	Ben,	Theme ⟩	
a-struc.		[− *o*]	[− *r*]	[− *r*]	
		∅			
f-struc			S	O	SUBJ Mapping 2 NonSUBJ Mapping

	hand ⟨	Agent,	Ben,	Theme ⟩	
a-struc.		[− *o*]	[− *r*]	[− *r*]	
		∅			
f-struc.			O	S	SUBJ Mapping 2 NonSUBJ Mapping

b.

	hand ⟨	Agent,	Ben,	Theme ⟩	
a-struc.		[− *o*]	[− *r*]	[+ *o*]	
		∅			
f-struc.			S	O_θ	SUBJ Mapping 2 NonSUBJ Mapping

As we can see, the double-[− *r*] a-structure can project not one, but two f-structure lexical forms. Specifically, we predict that either "object" of the

verb can be SUBJ under passivization. As we have seen, this is wrong for most varieties of English.

(46) a. Morn was handed a drink.
b. *A drink was handed Morn.

The correct result for English is obtained if we take the less preferred a-structure in which the Theme is $[+o]$. On the other hand, as mentioned earlier, this result is correct for some other languages, such as Kichaga. This seems to be a parameter of cross-linguistic variation. A natural place to isolate this variation is in a-structure. Asymmetrical languages rule out a-structures with two unrestricted ($[-r]$ arguments, while symmetrical languages do not.

(47) **Asymmetric Object Parameter**
$*\langle \ldots [-r] \ldots [-r] \ldots \rangle$

This must be a condition on a-structure; it cannot be a condition on fully specified argument functions, since SUBJ is also $[-r]$. It also cannot be a condition on thematic roles, since what is at stake here is not the thematic roles but rather their mapping to syntax.

We thus end this chapter where we started, with the passive. The properties of the passive mark it as an a-structure phenomenon. We now have a theory of a-structure in which to place it.

4.4 Appendix: Romance causatives

In justifying the representation of a-structure, we referred to the Romance causative construction, exemplified by the following examples from French. (Examples (48a) and (48b) are repeated from (12a,b); we are introducing (48c) here for the first time.)

(48) a. On fera parler Jean de son dernier voyage.
one CAUS.FUT talk John of his last journey
'One will make John talk of his last journey.'

b. Elle a fait visiter la ferme à ses parents.
she has CAUS.PART visit the farm DAT her relatives
'She made her relatives visit the farm.'

c. Elle a fait visiter la ferme par ses parents.
she has CAUS.PART visit the farm by her relatives
'She made her relatives visit the farm.'

Such constructions do not exist in English, and thus do not enter into the grammar of English that we are formulating. We will sketch an analysis of the Romance causative construction in this appendix. The analysis is largely based on the influential work of Alsina (1996a). However, Alsina assumes a nonstandard approach to object functions (rather than the OBJ and OBJ_θ of standard LFG, he rejects the Uniqueness Condition and allows multiple OBJs) and a nonstandard feature system for LMT. Our analysis will therefore be an adaptation of Alsina's.

We need to make certain assumptions about the syntax of Romance languages.

- We begin with an observation about objects in Romance languages. As we have seen, Theme/Patient arguments are generally mapped to [− *r*], but those classified as "secondary" on a language-particular basis can be [+ *o*] as a marked option. In English, "secondary" means a Theme which is not also a Patient. In the Romance languages the definition of "secondary" is different. In verbs with meanings similar to *give*, the Theme is the OBJ and the Goal is the OBJ_θ (marked with the particle *à* in French).[12] We hypothesize provisionally that a "secondary" Theme/Patient in Romance languages is a non-Theme Patient.
- The VP sister of V can be a cohead (i.e. can be annotated "↑ = ↓"). This is necessary to account for auxiliary constructions, as analyzed in Butt, King, Niño, and Segond (1999).
- The Romance languages are asymmetrical.

We assume (following Alsina) that the causative verb in Romance takes three arguments: the entity causing the action (Agent), an entity acted on by the Agent (Patient), and the action caused. The action caused is expressed by the VP in the structural complement position of the causative verb (*faire* in French), and the a-structure of its head verb ultimately fuses with the a-structure of the causative verb. The Patient of the causative verb is identical

[12] Alsina argues convincingly that the *à*-marked nominal is an object and not an oblique. Note for example that when pronominal it is marked with a clitic on the verb.

to one of the arguments of the other verb.

We can postulate a-structures for the French verbs in the examples. The causative verb's Patient is not a Theme, so it is a secondary Patient/Theme. We use the notation "…" to indicate that the causative verb's a-structure must fuse with that of another verb.

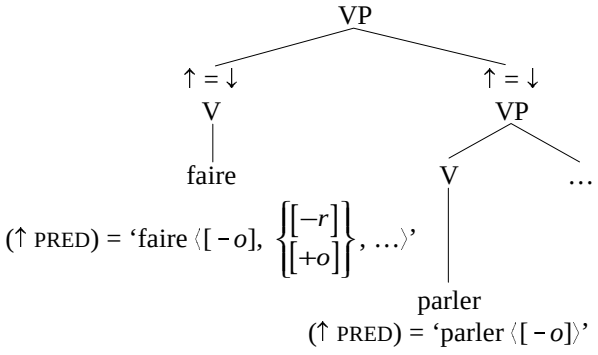
(49) a. *faire* $\langle [-o], \left\{ \begin{bmatrix} -r \end{bmatrix} \atop \begin{bmatrix} +o \end{bmatrix} \right\}, \ldots \rangle$

b. *parler* $\langle [-o] \rangle$

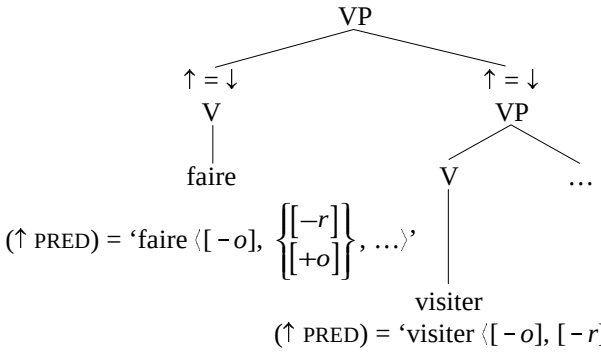
c. *visiter* $\langle [-o], [-r] \rangle$

Using the cohead annotation on the VP in structural complement position, the following partial annotated c-structures result.

(50) a.



b.



Given this as background, we need to examine what happens to the PRED

features in the f-structures. Generally, identical features which end up in the same f-structure unify, but PRED is different. Since the PRED feature is an indication of the meaningfulness, two PRED features are independently meaningful and therefore cannot unify. What is proposed by Alsina and others who have researched complex predicates is that when one of the PREDs belongs to a light verb, like the French *faire*, predicate composition takes place. This is conceptually similar to unification, in that it combines the properties of the two PRED features, but it is formally different. Since the Patient argument of *faire* binds an argument of the other verb, the bound argument does not project into the syntax. If the verb is intransitive, there is only one possible argument with which the Patient of *faire* can be identified, but if it is transitive, there are two possibilities.

(51) a. *faire* $\langle [-o], \left\{ \begin{matrix} [-r] \\ [+o] \end{matrix} \right\}_i, \ldots \rangle$ + *parler* $\langle [-o]_i \rangle$

The sole argument of *parler* is bound by the Patient argument of *faire*. Since the $[-r]$ mapping is the unmarked one for the second argument of *faire*, it is the one that is used. The resulting complex predicate has the following a-structure:

faire-parler $\langle [-o], [-r]_i \langle \_\_\_{}_i \rangle \rangle$

This will map into grammatical functions with our existing a-structure–f-structure mapping rules with the causer as SUBJ and the speaker as OBJ, as in (48a)

b. *faire* $\langle [-o], \left\{ \begin{matrix} [-r] \\ [+o] \end{matrix} \right\}_i, \ldots \rangle$ + *visiter* $\langle [-o]_i, [-r] \rangle$

The Patient of *faire* is identified with the first ("visitor") argument of *visiter*, which is not independently mapped to f-structure. Since French is an asymmetrical language, the Patient of *faire* is $[+o]$. The composite a-structure is:

faire-visiter $\langle [-o], [+o]_i \langle \_\_\_{}_i, [-r] \rangle \rangle$

The mapping to grammatical functions results in the causer as SUBJ, the visitor as OBJ_θ, and the visitee as OBJ, as in (48b).

c. *faire* $\langle [-o], \left\{\begin{matrix}[-r]\\[+o]\end{matrix}\right\}_i, \ldots\rangle$ + *visiter* $\langle [-o], [-r]_i\rangle$

This time the Patient of *faire* is identified with the second ("visitee") argument of *visiter*, which is not independently mapped to f-structure. The composite a-structure is:

faire-visiter $\langle [-o], [-r]_i \langle [-o], \_\_\_{}_i\rangle\rangle$

The a-structure–f-structure mapping maps the causer to SUBJ, the visitor to OBL_θ, and the visitee (also the Patient of *faire*) to OBJ, as in (48c).

The patterns of Romance causativization are thus accounted for.

There are two important observations to be made about this analysis. In the first place, the composition of a-structures in the syntax is made possible by a monotonic account like LMT. The realization of the SUBJ argument of *visiter* as the OBL_θ of *faire-visiter* (in (48c)) does not require any function-changing processes; it is inherent in the a-structure mapping of the Agent argument to $[-o]$ ($\text{SUBJ}/\text{OBL}_\theta$). The same composition of a-structures takes place in the lexicon in languages like Turkish. The second observation is that typological differences between languages in the causative construction can be attributed to various aspects of the analysis: the nature of object functions in a language, the identification of the Patient of 'cause' with an argument of the other verb, the existence of a Patient argument for 'cause'. More work is required to determine the full extent of crosslinguistic variation, but the LFG/LMT account provides a rich basis for such study.

Additional readings

LMT originates in the work of Levin (1988), as developed in Bresnan and Kanerva (1989). Naturally, the passive construction played a central role in the development of LMT. Arguments against a Case-based theory of passive are given by Zaenen and Maling (1982), and arguments against argument remapping (and in favor of LMT) are given by Bresnan (1990). Our brief discussion of Bantu ditransitive constructions glossed over many interesting details; for more see Bresnan and Moshi (1990), Alsina and Mchombo (1990), Alsina and Mchombo (1993), Alsina (1994), and Alsina (1996b).

There has been much work on complex predicates, which the Appendix to this chapter barely begins to cover: Butt (1995), Frank (1996), the papers in Alsina, Bresnan, and Sells, eds. (1997), Ackerman and Webelhuth (1998), Alsina (1996a), Andrews and Manning (1999), and

others. On causatives, see Alsina (1992), Alsina and Joshi (1991), and Matsumoto (1998).

The LMT analysis of "deep unaccusativity" was first given in Bresnan and Zaenen (1990). Slightly different views of a-structure and LMT than those presented here have been proposed by Alsina (1996a) and Manning (1996). On the analysis of a language in which unergatives and unaccusatives behave the same, see Lødrup (2000) on Norwegian.

On lexical mapping in nominals, see Laczkó (2000).

Exercises

1. For each of the following sentences, show how the grammatical functions of the arguments are derived from the a-structure of the verb by Lexical Mapping Theory.

   a. The phonologist spoke.

   b. The transformationalist moved unintentionally.

   c. Books about pragmatics read easily.

   d. We bought a sentence from the syntactician.

   e. The morphologist was handed a word.

   f. The semanticist has a dictionary.

   g. The relational grammarian went to her brother-in-law's house.

2. Explain the ungrammaticality of each of the following using Lexical Mapping Theory.

   a. We bought the syntactician a sentence. (* with the same meaning as (d) in Question 1, ✓ on the (irrelevant) reading with *the syntactician* as Beneficiary.)

   b. *A word was handed the morphologist.

   c. *The relational grammarian went her brother-in-law's house.

5

Control: Equi and Raising Constructions

In the previous three chapters we have developed the formalisms of the three syntactic levels of representation and correspondences between them. We now use this theoretical architecture to analyze several central constructions of English.

5.1 Preliminary survey

In this chapter, we will examine a group of constructions that can be jointly called **control** constructions. In LFG, the term "control" is used to refer to any construction in which there is a (in most languages[1]) nonfinite verb form with no overt subject, with particular grammatical constraints on the reference of the missing subject. The following are all examples of control.

(1) a. The geneticist wishes [to clone dinosaurs].
b. [To clone dinosaurs] would please the geneticist.
c. [Cloning dinosaurs] pleases the geneticist.
d. The geneticist tried [to clone dinosaurs].
e. The geneticist kept [cloning dinosaurs].
f. The millionaire persuaded the geneticist [to clone dinosaurs].
g. The geneticist seemed [to clone dinosaurs].
h. They believed the geneticist [to clone dinosaurs]

In each of these cases, the bracketed constituent is a clause-like phrase that has an unexpressed SUBJ (the **controllee**). The DP *the geneticist*, an element of the main clause (the **controller**), is (or may optionally be) understood as the unexpressed SUBJ.

[1]Zec (1987) argues that in Serbo-Croatian finite clauses can also be controlled.

It is customary in generative syntax to distinguish between two different types of control constructions, based on whether the controller bears a thematic role in its clause. In (1g), for example, *the geneticist* is not a thematic argument of *seems*. Because of the theoretical assumptions of transformational theories (such as the Θ Criterion), the thematic status of the controller dictates the analysis. The assumptions of LFG are different, but the terminological difference is a useful taxonomic distinction. We therefore refer to control constructions in which the controller bears a thematic relation to its verb as **equi**[2] and those in which it does not as **raising**.

5.2 Equi: Anaphoric control

We begin with equi constructions that do not involve complements. Consider a sentence like (1b), repeated here.

(1b) [To clone dinosaurs] would please the geneticist.

It has been realized since Postal (1970) that there is some sort of pronounlike element serving as subject in this kind of clause. In GB, this pronounlike element is called *PRO*, and is taken to be an element of c-structure occupying the position [SPEC, IP], which is the usual structural position for subjects.

In LFG we need to ask whether this pronounlike element is an element of c-structure or of f-structure. To answer this question, we need to consider the reasons for hypothesizing it. First, it is needed because the verb in the infinitive selects a subject argument. This is an f-structure consideration. If the f-structure of *to clone dinosaurs* lacked a SUBJ, the f-structure would be incomplete:

(2)
$$\begin{bmatrix} \text{PRED} & \text{`clone } \langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ}) \rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix}$$

[2]The original transformational term "equi" (an abbreviation of "Equivalent NP Deletion") was replaced around 1980 by "control" because the accepted analysis no longer involved a deletion rule. In LFG, as pointed out in the text, the names are purely descriptive and have no relevance to the analysis. The usual usage in LFG is to use the term "control" in the wider sense for which we are using it here, and "equi" in the narrower sense in which "control" is meant in the GB tradition.

Another argument for the syntactic presence of the understood subject is that it interacts with Binding Theory; it can serve as the antecedent for a reflexive in its clause.

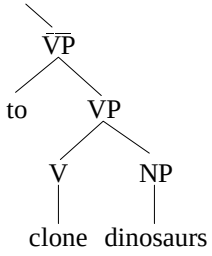
(3) [To clone himself] would please the geneticist even more.

As we will see in Chapter 7, the syntactic dimension of anaphora involves f-structure, so again we have evidence for the unexpressed pronominal SUBJ to be represented at f-structure. More specifically, since pronouns are distinguished from other nouns by the feature [PRED 'PRO'], the infinitive has to have SUBJ [PRED 'PRO'] in its f-structure.

(4)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \vdots & \end{bmatrix} \\ \text{PRED} & \text{'cloned } \langle (\uparrow \text{SUBJ}) \, (\uparrow \text{OBJ}) \rangle \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix}$$

On the other hand, there seems to be no evidence for an empty c-structure subject. By the principle of Economy of Expression, unnecessary c-structure nodes do not exist. Therefore, there is no *PRO* in c-structure.

(5)



The notation "$\bar{\bar{V}}P$" is an informal notation sometimes used in the LFG literature for the *to* infinitive. We will return at the end of the chapter to a discussion of the identity of "$\bar{\bar{V}}P$" in $\bar{X}$ theory. For now, we will use the label "$\bar{\bar{V}}P$".

The control relation here is an anaphoric one, so this is called **anaphoric**

control. In this case, there is an anaphoric link[3] between the unexpressed subject of the infinitive and *the geneticist*.

(6)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{INDEX} & i \end{bmatrix} \\ \text{PRED} & \text{`clone } \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix} \\ \text{TENSE} & \text{CONDITIONAL} \\ \text{PRED} & \text{`please } \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`geneticist'} \\ \text{NUM} & \text{SG} \\ \text{INDEX} & i \end{bmatrix} \end{bmatrix}$$

The analysis is an f-structure equivalent of the GB analysis.

In English, the unexpressed pronoun must be the SUBJ of a nonfinite verb. The lexical entries of nonfinite verbs include an optional equation introduced by the following rule.

(7) Add the optional equation (↑ SUBJ PRED) = 'PRO' to the lexical entry of a verb without the feature TENSE.

In GB, the limitation of unexpressed pronominals to the subjects of nonfinite verbs is supposed to be explained by the "PRO Theorem," which stipulates *PRO* to be a "pronominal anaphor," i.e. to have the properties of both pronouns and reflexives. Since pronouns and reflexives have contradictory binding properties, so the explanation goes, the unexpressed pronoun can only occupy a position that is exempt from binding theory, an "ungoverned" position. Other stipulations of the theory result in subject position of a nonfinite clause being an ungoverned position. This approach is based on the dubious idea that the unexpressed pronoun simultaneously has the properties of pronouns and reflexives. Bresnan (1982b) disputes this, and shows that it has no reflexive properties.

---

[3]We provisionally indicate this anaphoric link by "coindexing", understanding INDEX as an f-structure attribute. We will discuss the status of index further at the end of Chapter 7.

Bresnan suggests that the limitations on the distribution of the unexpressed pronoun are language-specific. Many languages allow pronominal arguments to be unexpressed; this phenomenon is sometimes called "pronoun drop" or "pro drop." While GB distinguishes this empty pronoun from the one in subject position of nonfinite verbs by calling it *pro*, there is no reason to consider it a distinct entity. Bresnan's proposal is that the English rule (7) is a specific instance of a universal parameterized rule.

(8) Add the optional equation (↑ GF PRED) = 'PRO' to the lexical entry of a verb <in some languages: without the feature TENSE>.

The LFG analysis of both control and pro-drop thus differs from familiar structurally-based analyses in that the unexpressed pronouns are not represented at c-structure. There is evidence for this in Malayalam, which has a condition on anaphora that does not allow a pronoun to precede its antecedent, independent of any structural considerations. Precedence is a c-structure concept; elements that exist only in f-structure cannot precede or follow other elements. Unexpressed (pro-dropped) pronouns do not display precedence effects.

(9) a. [awan aanaye ɳuḷḷiyaṭinə šeesam] kuṭṭi uraŋŋi.
he elephant.ACC pinched.it after child.NOM slept
'The child slept after someone else/*the child pinched the elephant.'

b. [aanaye ɳuḷḷiyaṭinə šeesam] kuṭṭi uraŋŋi.
elephant.ACC pinched.it after child.NOM slept
'The child slept after the child/someone else pinched the elephant.'

This can be captured using the notion of f-precedence. As we have seen, elements which are not part of c-structure either do not enter into f-precedence relations or both f-precede and are f-preceded by every other element in the f-structure. It would be difficult to account for this contrast in a theory in which all syntactic elements, including those that are unexpressed, are represented in c-structure.

Unexpressed pronouns differ in some properties from expressed pronouns. Note the following contrast.

(10) a. People who know the geneticist often discuss his cloning dinosaurs. (*his* = *the geneticist* possibly)
b. People who know the geneticist often discuss cloning dinosaurs. (understood SUBJ of *cloning* ≠ *the geneticist*)

The f-structure of both these sentences is approximately as follows (with the "…" under [PRED 'PRO'] filled in with the different properties of the expressed and unexpressed pronoun).

(11)
$$\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix}
\text{PRED} & \text{'person'} \\
\text{NUM} & \text{PL} \\
\text{ADJ} & \left\{ \begin{bmatrix}
\text{SUBJ} & [\text{"who"}] \\
\text{TENSE} & \text{PRES} \\
\text{PRED} & \text{'know } \langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ}) \rangle\text{'} \\
\text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{'geneticist'} \\ \text{NUM} & \text{SG} \end{bmatrix}
\end{bmatrix} \right\}
\end{bmatrix} \\
\text{ADJ} & \left\{ [\text{PRED} \quad \text{'often'}] \right\} \\
\text{PRED} & \text{'discuss } \langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ}) \rangle\text{'} \\
\text{OBJ} & \begin{bmatrix}
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \vdots & \end{bmatrix} \\
\text{PRED} & \text{'clone } \langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ}) \rangle\text{'} \\
\text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix}
\end{bmatrix}
\end{bmatrix}$$

In this sentence, the unexpressed pronoun cannot be controlled by (corefer with) *the geneticist*, while the overt pronoun can. *The geneticist* is deeper in the f-structure than the pronoun. This relative deepness is similar to the c-structural concept of c-command, so it has been dubbed **f-command**.

(12) a. The unexpressed pronoun can only be coindexed with an f-commanding function.
b. For f-structures α, β, α f-commands β iff α does not contain β and every f-structure that contains α contains β.

Note that the restriction cannot be stated in terms of c-command; in the

following sentence the controller f-commands the unexpressed pronoun but does not c-command it.

(13) Cloning dinosaurs will amuse the geneticist. (unexpressed SUBJ of *cloning* = *the geneticist*, possibly)

F-structure provides the proper configuration; c-structure does not.

5.3 Raising: Functional control

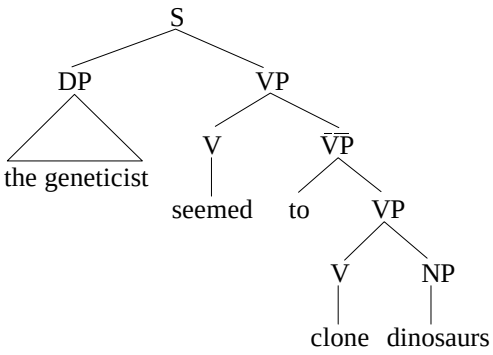
5.3.1 Raising-to-subject

We turn now to raising constructions. Consider the following sentence.

(14) The geneticist seemed to clone dinosaurs.

The c-structure of this sentence is unremarkable.

(15)



What is problematic is the status of the matrix SUBJ, *the geneticist*. On the one hand, it is not a semantic argument of *seem*. Only the infinitival complement is semantically related to the main verb. On the other hand, this SUBJ is a syntactic argument of *seemed*, as well as of *cloned*. Therefore, it must be part of *seemed*'s a-structure and PRED feature value. Following the notation introduced in Chapter 1 for idiomatic arguments, we place nonthematic arguments outside of the angle brackets.

(16) 'seem ⟨(↑ COMP)⟩ (↑ SUBJ)'

With the above as background, let us consider the options for analyzing raising in LFG. As a first attempt, we might try the same anaphoric control analysis as we used for equi.

(17)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`geneticist'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{`seem} \left\langle (\uparrow \text{COMP}) \right\rangle (\uparrow \text{SUBJ})\text{'} \\ \text{COMP} & \begin{bmatrix} \text{SUBJ} & [\text{PRED} \quad \text{`PRO'}] \\ \text{PRED} & \text{`clone} \left\langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \right\rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

However, this cannot be the analysis of raising constructions such as this, because the resulting f-structure is ill-formed. Specifically, it violates the Coherence Condition, which stipulates that all meaningful argument functions (i.e. ones with a PRED feature) be assigned a thematic role. The SUBJ in the outermost f-structure, which includes the feature [PRED 'geneticist'], does not receive a thematic role.

Let us consider the situation. The infinitival complements of raising verbs lack their own SUBJ argument. Instead, the SUBJ argument comes from the outside, in our example from the SUBJ of the main clause. This outside element is the controller. The relation between the infinitival clause and the controller is one of predication: the infinitival clause can be seen as a kind of predicate, predicated of the SUBJ of *seem*. This is different from the anaphoric relation we posited for equi constructions. Raising involves not coreference between two elements, but rather a single element shared by two clauses.

To formalize this intuition concerning the analysis of raising, we need to add machinery to the theory. First, we need to enrich our inventory of grammatical functions with a predicational function. Such a function can also be called an **open** function, because the SUBJ argument is left open to be predicated of an external entity. Second, we need to be able to establish the predication relationship.

First, we will deal with the open grammatical function. Up until now, all the grammatical functions we have seen have been **closed** functions,

functions in which all the argument functions come from within. In LFG, the open function is called XCOMP. That is the function of the infinitival complement to *seem*. We can update the lexical form of *seem* accordingly.

(18) 'seem ⟨(↑ XCOMP)⟩ (↑ SUBJ)'

It might be expected that if the XCOMP is a predicate, it would not be limited to infinitives. Other categories can be predicative as well. This expectation is fulfilled.

(19) a. The geneticist seems happy.
b. The geneticist seems out of his league.

This categorial flexibility is the original source of the name XCOMP. The "X" represents a generalization over a variety of categories.[4]

Interestingly, not all raising predicates accept XCOMPs of all categories. For example, in Standard American English *seem* cannot take an NP (or DP) complement. In some of the LFG literature, categorially restricted XCOMPs are called VCOMP, NCOMP, ACOMP, and PCOMP. However, this is formally illicit, since it expresses c-structure (category) information in f-structure. A better approach would be to include the following specification in the lexical entry of *seem*: (20a) states it in prose; (20b) expresses it formally, where "λ" is the "(category) label function" and ϕ^{-1} is the mapping from f-structure to c-structure. Under LFG's projection architecture, an f-structure-based constraint of this kind can refer to correspondence relationships between levels.

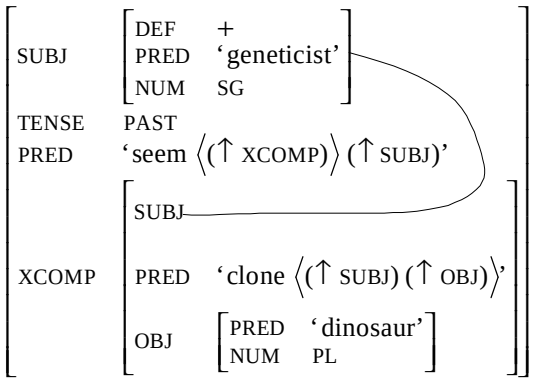
(20) a. N (or D, which is a functional version of N) cannot be the category label of any c-structure constituent in the set of nodes corresponding to the XCOMP.
b. $N \notin \lambda (\phi^{-1} (\uparrow \text{XCOMP}))$

Next, we need some way to establish the relation of predication. Within the existing formalism of f-structure, this can be done by establishing a relation of token identity between the controller and the infinitival SUBJ. That

[4]A historically less accurate, but nevertheless mnemonic, interpretation of the name XCOMP would be that the SUBJ of the XCOMP is open, i.e. *x*.

is to say, the same entity simultaneously fills two functions: SUBJ of *seem* and SUBJ of *clone*. This is actually not very different from the transformational analysis of raising, which suggests that the relation between controller and controllee is identity. The only formal tool available to transformational theories to model identity between two positions is movement from one position to the other. The LFG formalism allows a more direct approach: allowing the two functions to have the same f-structure as their value.[5] This relation of identity which is the LFG analysis of predicational constructions is called **functional control**. The usual notation is to draw a line connecting the two functions.

(21)



5.3.2 Raising-to-object

The analysis of sentences of the following type has been controversial in generative syntax.

(22) a. I believe the geneticist to clone dinosaurs.
b. I expect the geneticist to clone dinosaurs

What is universally agreed is that *the geneticist* is the subject of the subordinate clause and is not semantically related to the main verb *believe* or *expect*. For this reason, expletives and idiom chunks can occur in that position.

[5]In HPSG, this kind of relationship is called structure sharing.

(23) a. I believe there to be dinosaur embryos in the can.
b. I expect tabs to be kept on the velociraptors.

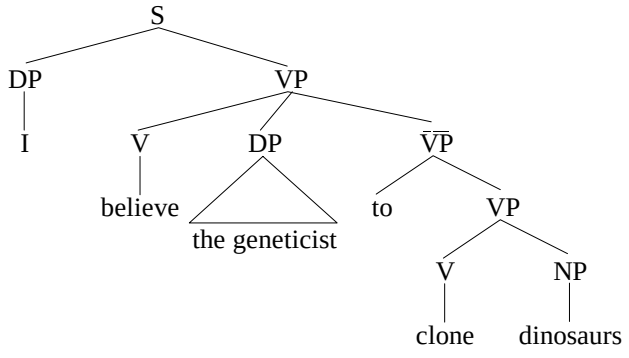
The dispute concerns the question of whether *the geneticist* is also a syntactic (surface) argument, object, of the main verb. If it is, we have a second kind of raising construction, one in which the controller is an OBJ rather than a SUBJ. The taxonomic name for this construction, based on the older derivational perspective, is Raising-to-Object.

The analysis of this kind of construction as Raising-to-Object was the first analysis proposed in generative grammar. However, since Chomsky (1973) there has been a second analysis available in which the nominal following verbs like *believe* and *expect* is not analyzed as a part of the main clause, but simply as the subordinate subject. This alternative analysis has come to be known as the Exceptional Case Marking (or ECM) analysis.

In Government/Binding theory, Raising-to-Object is ruled out on theory-internal grounds. It is stipulated that the nominal in object position must be assigned a thematic role by the verb whose object it is. Given this stipulation, the post-*believe* nominal cannot be the object of *believe*. Formally, the Raising-to-Object analysis violates GB's Projection Principle. Therefore, the ECM analysis is forced.

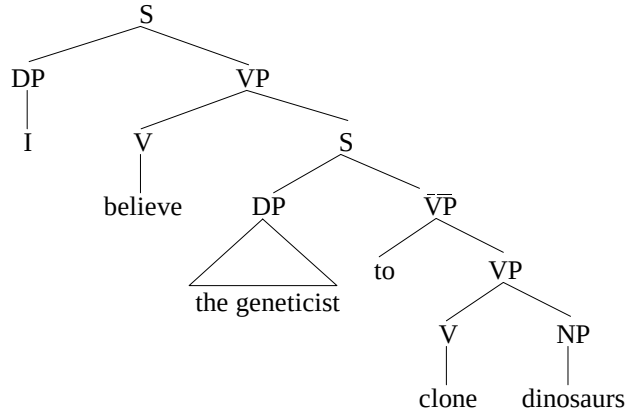
The two analyses can be realized in LFG as follows.

(24) Raising-to-Object analysis



$$\begin{bmatrix} \text{SUBJ} & [\text{“I”}] \\ \text{PRED} & \text{‘believe} \langle (\uparrow \text{SUBJ}) (\uparrow \text{XCOMP}) \rangle (\uparrow \text{OBJ})\text{’} \\ \text{OBJ} & [\text{“the geneticist”}] \\ \text{XCOMP} & \begin{bmatrix} \text{SUBJ} & \\ \text{PRED} & \text{‘clone} \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle\text{’} \\ \text{OBJ} & [\text{“dinosaurs”}] \end{bmatrix} \end{bmatrix}$$

(25) “ECM” analysis



$$\begin{bmatrix} \text{SUBJ} & [\text{“I”}] \\ \text{PRED} & \text{‘believe} \langle (\uparrow \text{SUBJ}) (\uparrow \text{COMP}) \rangle\text{’} \\ \text{COMP} & \begin{bmatrix} \text{SUBJ} & [\text{“the geneticist”}] \\ \text{PRED} & \text{‘clone} \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle\text{’} \\ \text{OBJ} & [\text{“dinosaurs”}] \end{bmatrix} \end{bmatrix}$$

There are both theoretical and empirical reasons to prefer the Raising-to-Object analysis over ECM. Both kinds of arguments are important, but the empirical arguments more so because they outweigh any theoretical considerations. We will provide both kinds of arguments.

We begin with a theory-internal argument: passivization. The nominal after *believe* becomes the subject of *believe* under passivization.

(26) a. The geneticist is believed to clone dinosaurs.
b. There are believed to be dinosaur embryos in the can.

As we have seen, passivization in LFG is an operation on a-structure. As a result of the suppression of the $\hat{\theta}$ argument, the [– *r*] argument is mapped to SUBJ. The nominal that follows *believe* in the active must be the [– *r*] argument in *believe*'s a-structure, realized at f-structure as OBJ in the active and SUBJ in the passive. The lexical status of passivization thus forces the Raising-to-Object analysis in LFG. Of course, as a theory-internal argument, this would not convince a GB theorist any more than the Projection Principle argument would convince an LFG theorist. This is the weakness of a theory-internal argument.

We can go one step farther, and show that the basis for GB's disallowing of the Raising-to-Object analysis is flawed. Recall that the central problem for Raising-to-Object in GB is that the object must be assigned a thematic role. The Projection Principle excludes nonthematic objects. However, as Postal and Pullum (1988) observe, expletives (which are by nature nonthematic) are possible in object position.

(27) a. I never gave it a thought that geneticists would clone dinosaurs.
b. I take it that velociraptors are dangerous.
c. I regret it deeply that the T-rex got loose.
d. I will see to it that the dinosaurs will behave.
e. When dinosaurs want to eat the same tree, they fight it out.

The grammaticality of sentences such as these shows that there is no problem with positing nonthematic OBJs.

There are several empirical arguments that can be raised in support of the Raising-to-Object analysis, as noted originally by Postal (1974), and since in many studies such as Bresnan (1982b) and Pollard and Sag (1994). For example, consider Heavy NP Shift. This is the construction that allows "heavy" objects to appear at the end of the sentence instead of in normal object position.

(28) a. The geneticist put [a gene to make the dinosaur lysine-dependent] in the embryo.
b. The geneticist put in the embryo [a gene to make the dinosaur lysine-dependent].

Heavy NP Shift can apply to the nominal after verbs like *believe*.

(29) a. I believe [the geneticist who was looking for a job last year] to be cloning dinosaurs.
b. I believe to be cloning dinosaurs [the geneticist who was looking for a job last year].

However, Heavy NP Shift does not apply to subjects.

(30) *I believe is cloning dinosaurs [the geneticist who was looking for a job last year].

The nonapplication of Heavy NP Shift to the subject cannot be attributed to the finiteness of the subordinate clause. Gerunds can have subjects, and their subjects are also immune to Heavy NP Shift.

(31) a. I disapprove of [the geneticist who was looking for a job last year] cloning dinosaurs.
b. *I disapprove of cloning dinosaurs [the geneticist who was looking for a job last year].

In undergoing Heavy NP Shift, then, the post-*believe* nominal behaves like an OBJ.

Another argument in favor of treating the nominal as an element of the higher clause relates to the scope of adverbials. An adverbial following the nominal can have the main clause as its scope; this is not possible when the adverbial follows an element that is uncontroversially in the subordinate clause.

(32) a. I believe dinosaurs to this day to be intelligent.
b. *I believe (that) dinosaurs to this day are intelligent.

This contrast is explained by an analysis that places the nominal in the main clause.

The f-command condition on anaphoric control provides another argument. The SUBJ of a subordinate clause cannot anaphorically control the SUBJ of another clause subordinate to the main clause. However, the Raising-to-Object nominal can control.

(33) a. Cloning dinosaurs will prove the geneticist to have no scruples.
(SUBJ of *cloning* = *the geneticist*, possibly)
b. Cloning dinosaurs will prove that the geneticist has no scruples.
(SUBJ of *cloning* ≠ *the geneticist*)

That is to say, the Raising-to-Object nominal f-commands the unexpressed pronoun. This is possible only if the Raising-to-Object nominal is in the main clause.

Object pronouns, especially *it*, tend to cliticize onto the preceding verb. One way in which this is manifested is in the reduction of the vowel to [ə]. This does not happen to subject pronouns. A Raising-to-Object pronoun can reduce.

(34) a. I believe ət to be true that the geneticist cloned dinosaurs.
b. *I believe ət is true that the geneticist cloned dinosaurs.

Last, but certainly not least, traditional constituency tests show that the Raising-to-Object nominal does not form a constituent with the infinitive. This can be shown by constructions like pseudoclefting and Right Node Raising.

(35) a. What I believe is that the geneticist is cloning dinosaurs.
b. *What I believe is the geneticist to be cloning dinosaurs.
c. What I dislike is the geneticist cloning dinosaurs.

(36) a. I believe, but you don't believe, that the geneticist is cloning dinosaurs.
b. *I believe, but you don't believe, the geneticist to be cloning dinosaurs.

We therefore conclude that the Raising-to-Object analysis is the correct one, on both theoretical and empirical grounds. In LFG, this means that it is a functional control structure.

5.3.3 Licensing functional control

To summarize thus far, the infinitival complements of verbs like *seem* and verbs like *believe* bear the predicative (open) function XCOMP. The SUBJ of this XCOMP is the SUBJ of *seem* and the OBJ of *believe*. This is achieved formally through feature sharing. In this section we will fill in some of the

technical details.

This sharing of feature structure is a lexical property of the raising verb. It must therefore be licensed by the verbs' lexical entries. The lexical entries of *seem* and *believe* will have the following **control equations**:

(37) a. *seem* (↑ SUBJ) = (↑ XCOMP SUBJ)
b. *believe* (↑ OBJ) = (↑ XCOMP SUBJ)

These are specific instances of the general functional control equation, which we assume is specified by a lexical rule.

(38) **Functional Control Rule**
If (↑ XCOMP) is present in a lexical form, add the equation:
(↑ CF) = (↑ XCOMP SUBJ)

The controller must be a core function because noncore functions are essentially grammaticalized thematic relations, and cannot have relations to other predicates.

The lexical entries for these verbs are as follows.

(39) a. *seem* (↑ PRED) = 'seem ⟨(↑ XCOMP)⟩ (↑ SUBJ)'
(↑ SUBJ) = (↑ XCOMP SUBJ)

b. *believe* (↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ XCOMP)⟩ (↑ OBJ)'
(↑ OBJ) = (↑ XCOMP SUBJ)

These lexical entries make it clear that the control equations in the lexical entries of these verbs are predictable. The controller in each case is a semantically meaningful argument function which does not receive a thematic role. If they were not functional controllers, the resulting f-structures would be incoherent.

Many raising verbs also take nonraising complements.

(40) a. It seems that the geneticist clones dinosaurs.
b. I believe that the geneticist clones dinosaurs.

This means that each of these verbs projects two a-structures from its θ-structure. A full account of control and complementation must include a-structure

rules. Since propositional Theme arguments do not map to OBJ in English,[6] we need to augment the θ-structure–to–a-structure mapping principles. This means we must add the functions COMP and XCOMP to the feature system. It has been suggested by Zaenen and Engdahl (1994) that they are variants of OBL_θ: [+*r*, −*o*] functions. We will take a more cautious position here, and assume that COMP and XCOMP have a special feature [+*c*]; all other functions (including SUBJ) will be [−*c*].

(41) **θ-structure to a-structure mapping** (revised)
nonpropositional Patients and Themes map to [−*r*]
"secondary" nonpropositional Patients and Themes (in English, non-Patient Themes) map to [+*o*]
propositional arguments map to [+*c*].
non-Theme/Patient arguments map to [−*o*]

The nonraising version of *seem* has the following mapping of arguments.

(42)

θ-structure		[Propositional Theme]
a-structure	[−*o*]	[+*c*]
GFs	SUBJ	COMP

(↑ PRED) = 'seem ⟨(↑ COMP)⟩ (↑ SUBJ)'
(↑ SUBJ FORM) = it

It is derived by an a-structure rule that adds a nonthematic argument with the [FORM it] feature. Following tradition, we call this rule Extraposition.

(43) **Extraposition**
⟨… [+*c*] …⟩ ⇒ [−*o*] ⟨… [+*c*] …⟩
(↑ [−*o*] FORM) = it

Like all a-structure rules, this rule is strictly monotonic: it adds information to the a-structure without changing what is already there.

The raising version has a slightly different mapping.

[6]For a different view, see Dalrymple and Lødrup (2000).

(44)

θ-structure		[Propositional Theme]
a-structure	[− *o*]	[+ *c*]
GFs	SUBJ	XCOMP

(↑ PRED) = 'seem ⟨(↑ XCOMP)⟩ (↑ SUBJ)'
(↑ SUBJ) = (↑ XCOMP SUBJ)

The a-structure rule responsible for this lexical entry can be called "Raising-to-Subject." Unlike Extraposition, Raising-to-Subject does not supply any content for the nonthematic argument introduced.

(45) **Raising-to-Subject**
⟨[+ *c*]⟩ ⇒ [− *o*] ⟨XCOMP⟩

One of these rules must apply in the mapping of *seem*. If neither rule applied, we would expect a mapping in which the propositional argument is the SUBJ, but this is ungrammatical.

(46) *That the geneticist clones dinosaurs seems.

In general, clausal SUBJs of verbs in English express Causes, not Themes. Themes can be SUBJs of other categories, such as adjectives, but not verbs.

(47) a. That the geneticist clones dinosaurs scares me.
b. That the geneticist clones dinosaurs is obvious.

Our a-structure mapping principles bar the mapping of propositional Theme arguments of verbs to SUBJ. Since every verb must have a SUBJ, either Extraposition or Raising-to-Subject must apply in the mapping to produce a grammatical lexical form.

Extraposition can also apply to the predicates in (47). The a-structure mapping principles provide alternative mappings for non-Theme propositional arguments, either [− *o*] or [+ *c*]. When the [+ *c*] mapping is chosen, Extraposition has the same results as with verbs like *seem*.

(48) a. It scares me that the geneticist clones dinosaurs.
b. It is obvious that the geneticist clones dinosaurs.

Raising-to-Subject is more restricted lexically, but it sometimes applies to

arguments of adjectives. The adjective *likely*, for example, has three different mappings.

(49) a. That the geneticist will clone dinosaurs is likely.
b. It is likely that the geneticist will clone dinosaurs.
c. The geneticist is likely to clone dinosaurs.

Transformational theories derive these sentence types from each other in the syntax. In a lexical theory, these are treated as different forms of argument mapping.

Raising-to-Object is a little more complicated.

(50) a. I believe the story.
b. I believe that the geneticist clones dinosaurs.
c. I believe the geneticist to clone dinosaurs.

The verb *believe* has three argument mappings.

(51)

θ-structure	[Experiencer]	[Theme]
a-structure	[− *o*]	[− *r*]
GFs	SUBJ	OBJ

(↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ OBJ)⟩'

(52)

θ-structure	[Experiencer]	[Propositional Theme]
a-structure	[− *o*]	[+*c*]
GFs	SUBJ	COMP

(↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ COMP)⟩'

(53)

θ-structure	[Experiencer]		[Prop Theme]
a-structure	[− *o*]	[− *r*]	[+*c*]
GFs	SUBJ	OBJ	XCOMP

(↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ XCOMP)⟩ (↑ OBJ)'
(↑ OBJ) = (↑ XCOMP SUBJ)

Raising-to-Object verbs are typically ones that, like *believe*, can take either an ordinary Theme (OBJ) or a propositional Theme (COMP). We make this

part of the a-structure rule of Raising-to-Object.

(54) **Raising-to-Object**

$$\langle [-o] \left\{ \begin{matrix} [-r] \\ [+c] \end{matrix} \right\} \rangle \Rightarrow \langle [-o] \text{ XCOMP} \rangle \ [-r]$$

5.4 Equi complements

Consider sentences such as the following.

(55) a. The geneticist agreed to clone dinosaurs.
b. The geneticist tried to clone dinosaurs.

These sentences have main verbs that take equi complements: the main clause SUBJ is a thematic argument of the verb, and is also the SUBJ of the complement. As complements, they resemble functional control constructions, but as equi constructions one might think that they involve anaphoric control. In transformational theory, the analysis is forced by the architecture of the theory. If the subject of the main clause and the (understood) subject of the subordinate clause are both thematic arguments of their respective verbs, they must be two distinct D-structure elements. Thus, sentences like these must be analyzed as having a PRO subject in the subordinate clause, the equivalent of LFG's anaphoric control. However, the architecture of LFG is different. These could be anaphoric control, with a COMP argument that has an unexpressed pronoun SUBJ. It could also be functional control, a predicative construction with an XCOMP argument in which the matrix SUBJ and the complement SUBJ are functionally identified. Nothing precludes such an analysis. Unlike the Θ Criterion of GB, there is no principle of LFG that disallows one subsidiary f-structure from filling two different thematic positions, nor is there any reason that there should be. So these sentences could involve either functional or anaphoric control.

(56) a.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`geneticist'} \end{bmatrix} \\ \text{PRED} & \text{`try / agree } \langle (\uparrow \text{SUBJ}) (\uparrow \text{XCOMP}) \rangle \text{'} \\ \text{TENSE} & \text{PAST} \\ \text{XCOMP} & \begin{bmatrix} \text{SUBJ} & \\ \text{PRED} & \text{`clone } \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

b.

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`geneticist'} \\ \text{INDEX} & i \end{bmatrix} \\ \text{PRED} & \text{`try / agree } \langle (\uparrow \text{SUBJ}) (\uparrow \text{XCOMP}) \rangle \text{'} \\ \text{TENSE} & \text{PAST} \\ \text{XCOMP} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{INDEX} & i \end{bmatrix} \\ \text{PRED} & \text{`clone } \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle \text{'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

To determine the correct analysis, let us consider some of the properties of anaphoric and functional control. The clearest properties that could be used to distinguish them pertain to the controller: in functional control, the controller must be present (in the f-structure) and it must be a core function. If it were absent, the subordinate clause would be incomplete because it would lack a SUBJ, and, as discussed earlier, only core functions can be specified by a control equation. On the other hand, neither of these is necessary for anaphoric control: pronouns need not have antecedents and there is no restriction on the grammatical function of an antecedent of a pronoun. On the other hand, anaphoric control should allow split controllers, because pronouns can take split antecedents; while a functional controller is the single element specified by the control equation. Testing these properties on *try* and *agree*, we find the following.

(57) a. It was agreed to clone dinosaurs.
b. *It was tried to clone dinosaurs.

(58) a. It was agreed by the geneticists to clone dinosaurs.
b. *It was tried by the geneticists to clone dinosaurs.

(59) a. The geneticist said that the paleontologist agreed to clone dinosaurs. (SUBJ of *clone* = *the paleontologist* and *the geneticist* possibly)
b. The geneticist said that the paleontologist tried to clone dinosaurs. (SUBJ of *clone* = *the paleontologist* only)

This evidence suggests that the complement of *agree* is an anaphorically controlled COMP, while the complement of *try* is a functionally controlled XCOMP. None of the properties of *try* would be explained by an anaphoric control analysis.

A less conclusive distinction between anaphoric control and functional control concerns the extent to which the governing verb occurs obligatorily in an equi construction. The rough generalization is that obligatory control constructions involve functional control and nonobligatory control constructions involve anaphoric control. The basic idea is that functional control is a lexical property of the governing verb, while anaphoric control is unrelated to the nature of the governing verb. Similarly, since anaphoric control involves an ordinary COMP function, an overt SUBJ should be an option. With *try* and *agree*, these properties confirm the analysis of *agree* as anaphoric control and *try* as functional control.

(60) a. The paleontologist agreed that the geneticist would clone dinosaurs.
b. *The paleontologist tried that the geneticist would clone dinosaurs.

(61) a. The paleontologist agreed for the geneticist to clone dinosaurs.
b. *The paleontologist tried for the geneticist to clone dinosaurs.

However, the relation between obligatoriness and functional control is not absolute. Obligatoriness of control is due to the semantics of the verb, and there is no reason that this obligatoriness at the semantic level should necessarily be mirrored by the syntax. That is to say, "obligatory anaphoric

control" is possible.[7] Conversely, since verbs can have a COMP/XCOMP alternation, being a functional control verb does not necessarily rule out an alternative with a lexical SUBJ. Obligatoriness of the control relation is thus a less reliable guide.

In some languages, distinguishing it between functional control and anaphoric control is easier. Since functional control is complete identity, features like CASE are shared. In languages with extensive morphological Case marking, the effects of this identity are easier to see. Examples in Icelandic are discussed by Andrews (1982; 1990).

The crucial point is that while in transformational theory, equi and raising have to have different analyses, in LFG either one could be functional control; equi can be either functional or anaphoric. One has to look at the properties of each control verb to figure out which kind of control is involved. However, the empirical facts point to the correctness of this kind of approach; both kinds of control constructions exist. A theory which forces equi constructions to be anaphoric cannot account for the properties of functional-control equi.

5.5 C-structure

We return now to the question of the category of *to*, and of the infinitival clause. The LFG literature has generally been neutral on this question, with occasional suggestions (as in Bresnan 2000) that anaphoric control infinitives are CP and functional control infinitives are VP. In GB, *to* is treated as an infl. In this section, we will outline an approach that is empirically supported and is compatible with LFG. It does not reflect a consensus among LFG researchers.

First, we note that the *to* infinitival has the distribution of CP. It can be the SUBJ of a clause and a complement to V, N, and A, but not P.

(62) a. [To clone dinosaurs] would please the geneticist.
b. The geneticist tried [to clone dinosaurs].
c. [$_{DP}$ the geneticist's attempt [to clone dinosaurs]]
d. [$_{AP}$ very happy [to clone dinosaurs]]
e. *[$_{PP}$ before [to clone dinosaurs]]

This distribution is the same as CP, and different from either IP or VP.

[7]Zec (1987) shows that Serbo-Croatian has obligatory anaphoric control.

Furthermore, *to* infinitivals can be introduced by *wh* phrases, which occupy the [SPEC, CP] position.

(63) a. He doesn't know [which dinosaur to clone].
b. [$_{DP}$ a species [with which to experiment]]

We therefore conclude that "$\overline{V}\overline{P}$" is to be identified with CP.

If the *to* infinitive is a CP, the most natural analysis of *to* is as a complementizer. It is the element that converts what is otherwise a VP into a CP. Furthermore, *to* does not exhibit particularly infl-like properties, such as being followed by a marker of polarity[8] or being stressable when the following VP is ellipted.

(64) a. The geneticist will not/so clone dinosaurs.
b. The geneticist wants to ??not/*so clone dinosaurs.

The most natural analysis is that *to* is a complementizer.

We need to extend the ID rule for C′, to licence a VP complement as an alternative to IP and S.

(65) $$C' \rightarrow \underset{\uparrow = \downarrow}{C} \quad , \quad \underset{\uparrow = \downarrow}{\left\{ \begin{array}{c} S \\ IP \\ VP \end{array} \right\}}$$

As we saw in Chapter 3, the lexical entries for the complementizers *that* and *to* will specify that the former requires the feature TENSE (i.e. a finite clause) while the latter disallows it.

(66) a. *that* C (↑ TENSE)
b. *to* C ¬(↑ TENSE)

Given the CP status of *to* infinitivals, XCOMPs can be any category other than IP or S, or CP containing IP or S. What distinguishes these categories is the

[8] *Not* is not entirely ungrammatical, because it can be adjoined to VP for VP negation. However, this is not the normal place for *not* in clausal negation of the infinitive. Furthermore, the emphatic positive polarity item *so* is totally ungrammatical after *to*.

presence of the feature TENSE. We therefore add the following term to the ID rule for VP.

(67) $VP \rightarrow \ldots,\ \begin{matrix} XP \\ (\uparrow \text{XCOMP}) = \downarrow \\ \neg\,(\downarrow \text{TENSE}) \end{matrix},\ \ldots$

XCOMPs provide interesting confirmation of our decision to separate ID rules and LP rules, instead of using traditional phrase structure rules. The LP rules for English place PPs and CPs (and IPs and Ss) in final position (in that order, if both are present). This leads us to expect that in a VP with both an XCOMP and a PP, the XCOMP will precede the PP unless it is a *to* infinitive (i.e. CP). This is correct.

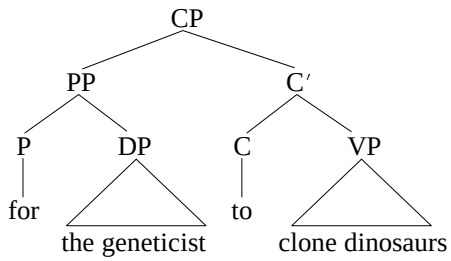
(68) a. The dinosaur seems $[_{\text{AP (XCOMP)}}$ cloned] $[_{\text{PP}}$ to me].
b. The dinosaur seems $[_{\text{PP}}$ to me] $[_{\text{CP(XCOMP)}}$ to be cloned].

Finally, although it is not a control construction, we also need to consider the structure of the "*for* infinitive", such as the bracketed clause in the following sentence.

(69) [For the geneticist to clone dinosaurs] would be a disaster.

The generative tradition has considered *for* to be a complementizer, and researchers in LFG have generally followed this tradition. If *to* is a complementizer, however, it is not clear that *for* would also be a complementizer. If it were, *for* infinitives would have two complementizers. What *for* actually seems to be is a marker for the SUBJ of the infinitive, since it is present when the SUBJ is represented in c-structure and absent when it is not. This is the position taken by Jespersen (1940: 308). The c-structure of the infinitival would then be:

(70)



The ID rule for CP has to license a SUBJ in specifier position for an infinitive, and specify that it be PCASE-marked with *for*.[9]

(71) $$\mathrm{CP} \rightarrow \left\{ \begin{array}{c} \mathrm{XP} \\ (\uparrow \mathrm{FOC}) = \downarrow \end{array} \right. , \quad \mathrm{C}' \\ \left\{ \begin{array}{c} (\uparrow \mathrm{SUBJ}) = \downarrow \\ \neg (\uparrow \mathrm{TENSE}) \\ (\downarrow \mathrm{PCASE}) =_{c} \mathrm{OBL}_{\mathrm{Ben}} \end{array} \right\} \uparrow = \downarrow$$

Additional readings

The LFG analysis of control was first laid out in Bresnan (1982b), and compared one-on-one in Mohanan (1983) with the GB analysis. The analysis has not appreciably changed, and studies of other languages have supported the LFG distinction between anaphoric and functional control as superior to the transformationalist distinction between control/equi and raising. Among these studies are Andrews (1982; 1990) on Icelandic, Neidle (1982) on Russian, Kroeger (1993) on Tagalog, Mohanan (1983) on Malayalam, and Arka and Simpson (1998) on Balinese.

The a-structure rules proposed here are original, and the c-structure analysis of *to* infinitives as CP is based on Falk (in preparation).

Exercises

1. For each of the following sentences, determine whether the bracketed clause is functionally controlled or anaphorically controlled. Explain. Draw f-structures. (There may not be a clear analysis for some of the cases.)

   a. Jake persuaded Nog [to attend Starfleet Academy].
   b. [To show emotion] would be difficult for Spock.
   c. The tribbles kept [eating grain].
   d. Kirk asked Gillian [to drive him to the ship].
   e. Troi wants [to ignore her mother].
   f. Picard expects [to boldly go where no one has gone before].

---

[9]We assume that *for* lexically bears the PCASE feature $\mathrm{OBL}_{\mathrm{Ben}}$, since it normally marks benefactives.

g. Sisko promised Kira [to protect Bajor].
h. Uhura signaled to Kirk to talk.

2. In Chapter 4, we proposed an a-structure rule adding a resultative argument. The rule can be divided into two parts:
   (i) Add a predicative argument with resultative semantics
   (ii) The resultative is predicated of [– *r*].

   a. What is the grammatical function of the resultative?
   b. How can we formalize (ii)?
   c. How does this analysis of resultatives account for the underlined DPs in sentences like:
      She ran <u>the soles</u> off her shoes.
      He sneezed <u>his head</u> off.

6

Long Distance Dependencies

6.1 Overview

We turn now to sentences like the following.

(1) a. Which book do you think I put on the shelf?
b. That theory, she told me she had never heard of.

In these sentences, the initial phrase can be conceptualized as belonging in two different clauses simultaneously, with a different function in each. Because the number of clauses between the two positions is unlimited,[1] these constructions are often called **long distance dependencies** or **unbounded dependencies**. In the transformational tradition the analysis of these sentences involves movement, so the construction can also be called "*wh* movement" or "Ā movement." We will not be using these derivationally oriented names here, although a name with a derivational flavor that is sometimes used in the nonderivational literature is **extraction**. The top end of the long distance dependency can be called the **filler** and the lower end can be called the **gap**.

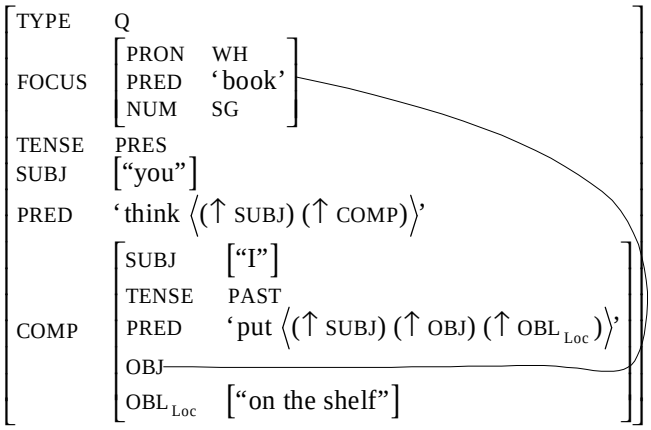
In English, long distance dependencies involve a missing element at the gap position (whence the name *gap*). The content of the filler has two grammatical functions, one typical of the position in which the filler is located, and one typical of the gap. For example, in (1a) the filler is *which book* and the gap is the would-be DP or NP after *put*. The DP *which book* therefore has the functions typical of the position that it occupies and of the postverbal nominal position. The latter is clearly OBJ of the subordinate

[1] In fact, it can be zero. The two functions can be in the same clause, as in (i).

(i) Which book did you read?

clause, while the former bears the grammaticized discourse function FOCUS. In the previous chapter, we introduced the notation of a curved line connecting two f-structure positions to indicate that one subsidiary f-structure fills two functions. Using the same notation here, the f-structure of (1a) is (2).

(2)



Assigning the filler position a discourse function is the LFG equivalent of calling it an $\bar{A}$ (nonargument) position in structural theories. However, the exact nature of the discourse function is not directly dependent on the structural position of the filler. There are two structural positions for fillers in English, [SPEC, CP] for *wh* phrases, and adjoined to IP for "topicalized" phrases. Elements in either position can have the function of either TOPIC or FOCUS. A TOPIC represents old information, while a FOCUS represents contrast (and thus new information). In [SPEC, CP] position, a question phrase has the function FOCUS while a relative pronoun has the function TOPIC. A constituent adjoined to IP can also be either FOCUS or TOPIC. These structural positions and their functions are licensed by the following ID rules.

(3) a. $\text{CP} \rightarrow \underset{(\uparrow \text{DF}) = \downarrow}{\text{XP}}, \underset{\uparrow = \downarrow}{\text{C}'}$
$(\downarrow \text{PRON}) =_{c} \text{WH}$

b. $\text{IP} \rightarrow \underset{(\uparrow \text{DF}) = \downarrow}{\text{XP}}, \underset{\uparrow = \downarrow}{\text{IP}}$
$(\downarrow \text{PRON}) \neq \text{WH}$

By the Extended Coherence Condition (p. 62), an discourse function must be linked to an argument function; a FOCUS or TOPIC not identified with an argument function is ungrammatical.

6.2 Licensing the dependency

6.2.1 Functional uncertainty

Long distance dependencies, like everything in f-structure, must be licensed by functional equations. Consider the following sentences with long distance dependencies. (The position of the gap is represented by a line.)

(4) a. Who did you see ___?
b. Who do you think ___ saw you?
c. Who do you think you saw ___?
d. Who did the hamster claim it thought ___ saw you?
e. Who did the hamster claim it thought you saw ___?
f. Who did the hamster claim it thought that the dinosaur said that the pterodactyl believes ___ saw you?

In each of these cases, some clause-internal grammatical function is identified with FOCUS. This is similar to functional control, in that it involves feature sharing. But it is different, in that the relationship between the two functions cannot be expressed as a finite expression. Consider what sorts of functional equations we would need to express these; we express these both as outside-in expressions starting from the clause of the filler and as inside-out expressions starting from the clause of the gap.

(5) a. Outside-in
(↑ FOCUS) = (↑ OBJ)
(↑ FOCUS) = (↑ COMP SUBJ)
(↑ FOCUS) = (↑ COMP OBJ)
(↑ FOCUS) = (↑ COMP COMP SUBJ)
(↑ FOCUS) = (↑ COMP COMP OBJ)
(↑ FOCUS) = (↑ COMP COMP COMP COMP SUBJ)

b. Inside-out
(↑ OBJ) = (↑ FOCUS)
(↑ SUBJ) = ((COMP ↑) FOCUS)
(↑ OBJ) = ((COMP ↑) FOCUS)
(↑ SUBJ) = ((COMP COMP ↑) FOCUS)
(↑ OBJ) = ((COMP COMP ↑) FOCUS)
(↑ SUBJ) = ((COMP COMP COMP COMP ↑) FOCUS)

However, as first observed by Kaplan and Zaenen (1989), the relationship can be expressed if we use the Kleene star operator.

(6) a. Outside-in (↑ DF) = (↑ COMP* GF)
b. Inside-out (↑ GF) = ((COMP* ↑) DF)

Unlike the equations we have seen until now, here there is more than one possible f-structure configuration that will satisfy the equation; any number of COMPs can intervene. There is an infinite number of possible solutions to an equation such as this. Consequently, this kind of functional designation is called **functional uncertainty**. As we can see, it can take the form of outside-in functional uncertainty or inside-out functional uncertainty.

Under the functional uncertainty formalism, long distance dependencies are licensed locally, one f-structure at a time. This brings long distance dependencies into line with the observation (Chapter 1) that all relations in syntax are local. In transformational theories, *wh* movement is not intrinsically local, although it may be restricted to near-locality by principles such as subjacency.[2]

By analyzing long distance dependencies in terms of functional uncertainty, LFG claims that c-structure properties are irrelevant to the behavior of the construction. This is in direct contrast with the movement

[2]The analysis of long distance dependencies in Kaplan and Bresnan (1982) is also not local. In that analysis, called "constituent control" the formal system of c-structure–f-structure mapping is extended. In addition to the metavariables ↑ and ↓, where ↑ receives the same variable as the immediately dominating ↓, Kaplan and Bresnan use metavariables ⇑ and ⇓, which are paired up with the same variable at a distance. Kaplan and Zaenen's functional uncertainty formalism has superseded the constituent control formalism. This local licensing of long distance dependencies is also a property of the HPSG analysis, in which the SLASH feature is propagated through the tree one node at a time.

analysis in transformational theories. The evidence favors the functional approach. For example, the filler and gap need not be the same category.

(7) a. [$_{CP}$That the hamster might like falafel], we didn't talk about ___.
b. *We didn't talk about [$_{CP}$that the hamster might like falafel].
c. We didn't talk about [$_{DP}$the hamster's fondness for falafel].

The ungrammaticality of (7b) is due to the fact that the preposition *about* takes an OBJ, not a COMP. In English, the ID rules assign the OBJ function to the DP/NP position only. In a movement theory of long distance dependencies, the grammaticality of (7a) is surprising given the ungrammaticality of (7b). In LFG, it is unproblematic. The f-structure is:

(8)

$$\begin{bmatrix}
\text{TOPIC} & \begin{bmatrix}
\text{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`hamster'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\
\text{TENSE} & \text{POSSIBILITY} \\
\text{PRED} & \text{`like} \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle\text{'} \\
\text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`falafel'} \end{bmatrix}
\end{bmatrix} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{PERS} & 1 \\ \text{NUM} & \text{PL} \end{bmatrix} \\
\text{POL} & \text{NEG} \\
\text{TENSE} & \text{PAST} \\
\text{PRED} & \text{`talk} \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBL}_{\text{about}}\ \text{OBJ}) \rangle\text{'} \\
\text{OBL}_{\text{about}} & \begin{bmatrix} \text{PCASE} & \text{OBL}_{\text{about}} \\ \text{OBJ} & \end{bmatrix}
\end{bmatrix}$$

There is nothing to rule out this f-structure. The fact that the constituent with the TOPIC function could not have been generated in the c-structure position normally associated with $\text{OBL}_{\text{about}}$ OBJ is irrelevant. Since the sentence is grammatical, a theory which does not rule it out is preferable.

6.2.2 Direction of licensing

As we saw in the previous section, long distance dependencies can be licensed either by outside-in functional uncertainty or inside-out functional uncertainty. In the LFG literature, both approaches have been proposed: Kaplan and Zaenen (1989) use the outside-in approach, while Bresnan

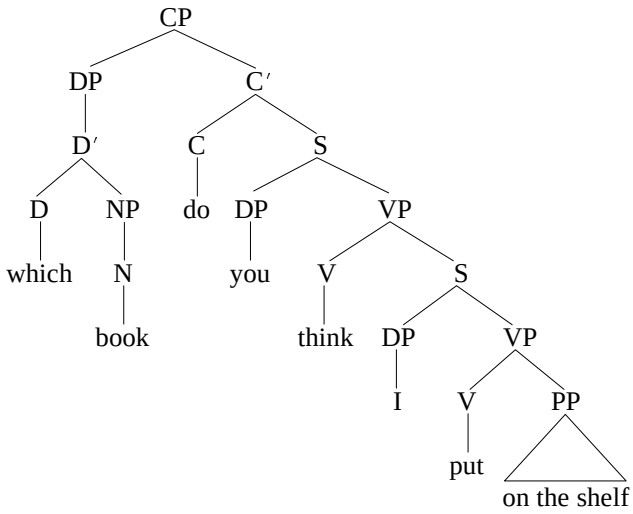
(1995a; 2000) argues for inside-out licensing. In this section, we will examine the details of the two approaches. We will reconcile them in the next section.

We begin with outside-in licensing. Under this approach, any clause that has a FOCUS or TOPIC will include the following equation.

(9) (↑ DF) = (↑ COMP* GF)

The c-structure of (1a) would be:

(10)



Ignoring the effect of the functional uncertainty equation, the f-structure associated with this c-structure is:

(11)
$$\begin{bmatrix} \text{TYPE} & \text{Q} \\ \text{FOCUS} & \begin{bmatrix} \text{PRON} & \text{WH} \\ \text{PRED} & \text{`book'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{SUBJ} & [\text{"you"}] \\ \text{PRED} & \text{`think} \langle (\uparrow \text{SUBJ})\ (\uparrow \text{COMP}) \rangle\text{'} \\ \text{COMP} & \begin{bmatrix} \text{SUBJ} & [\text{"I"}] \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{`put} \langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ})\ (\uparrow \text{OBL}_{\text{Loc}}) \rangle\text{'} \\ \text{OBL}_{\text{Loc}} & [\text{"on the shelf"}] \end{bmatrix} \end{bmatrix}$$

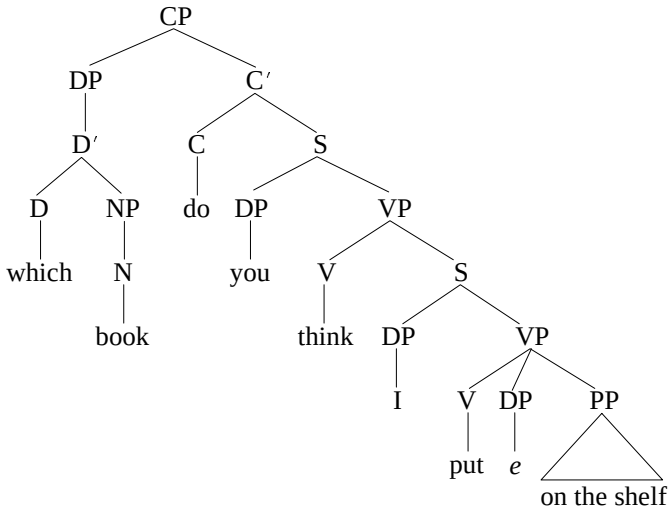
This f-structure is both incomplete and incoherent. It is incomplete because the verb *put* selects an OBJ, and there is none present in its local f-structure. It is incoherent because the discourse function FOCUS is not linked to an argument function, violating the Extended Coherence Condition. The functional uncertainty equation solves both problems, linking the FOCUS to *put*'s missing OBJ.

This analysis has several properties. First, there is no c-structure marking of the gap, analogous to the *wh* trace of movement theories. The only such analog is the argument function in f-structure. A c-structure gap is not needed; by Economy of Expression this means that there is none. Since c-structure is a model of the overt expression of syntax, and empty categories are, by definition, not overt, this is generally taken to be an advantage. Second, there are no constraints on identifying the gap. It can be anything any number of COMPs down. We will return to this presently. Third, it is unclear what node in the c-structure to associate the outside-in functional uncertainty equation with. Kaplan and Zaenen annotate it to the DF node itself, but this assumes that the DF is always present in the c-structure. This assumption is incorrect; in English relative clauses the DF need not be overt.

(12) the book [you think I put on the shelf]

The picture is slightly different under the inside-out approach. An inside-out equation has to be associated with the gap. One straightforward way to do this would be to make the gap a c-structure element (a "trace").

(13)



The verb *put* has an OBJ by virtue of being followed by a DP in normal OBJ position. The following rule licenses the empty category.

(14) $XP \rightarrow \quad \overset{e}{\uparrow = ((\text{COMP} * \text{GF} \uparrow) \text{DF})}$

This has the opposite advantages and disadvantages of the outside-in analysis. The gap end of the long distance dependency is marked, constraining the link, but at the expense of postulating an empty category.

6.2.3 Subjects vs. nonsubjects

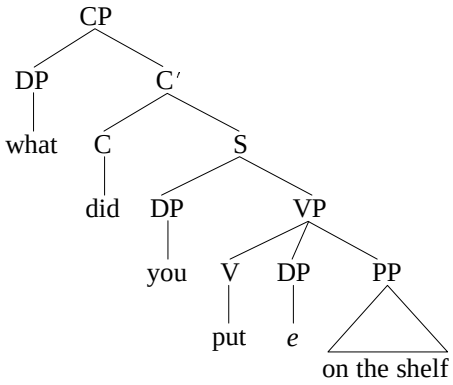
Any account of long distance dependencies needs to consider differences between extraction of subjects and nonsubjects, both in English and crosslinguistically. We will discuss some of these differences here, and suggest that they are relevant to determining the direction of the licensing of the long distance link. More precisely, we will distinguish between three cases: local linking of the DF to SUBJ (not involving functional uncertainty) long distance linking to SUBJ (outside-in functional uncertainty) and nonsubjects (inside-out functional uncertainty).

We begin by contrasting main clause subject questions from main clause nonsubject questions.

(15) a. Who put the book on the shelf?
b. What did you put on the shelf?

When a nonsubject is questioned, English is subject to "Subject-Aux Inversion", generally analyzed with the auxiliary in complementizer position instead of infl.

(16)



Formally, this can be achieved by lexically specifying auxiliaries as belonging ambiguously to either category.[3]

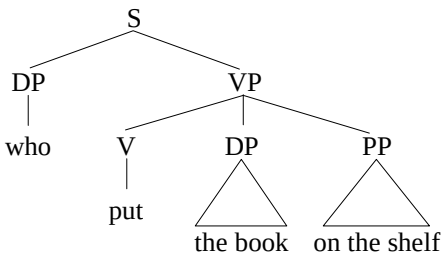
(17) *did* I or C (↑ TENSE) = PAST
C ⇒ (↑ TYPE) = Q

What is puzzling is the lack of Subject-Aux Inversion effects when the subject is questioned. Instead, a subject question resembles an ordinary declarative sentence.

[3]With "X⇒Y" meaning "if X then Y". This is slightly oversimplified, since auxiliaries can be C in some other constructions, too, such as:

(i) Never in his life had he read such an interesting book.

(18)



This structure has been proposed in GB by Grimshaw (1995). In a transformational theory, this would mean that subjects do not undergo *wh* movement.

We will adopt the "no movement" analysis as the best way to explain the lack of Subject-Aux Inversion. Recall that SUBJ is an overlay function, just like the discourse functions, and is the default topic. More generally, SUBJ seems to have an affinity for being identified with discourse functions. This can be expressed by annotating the subject DP with the following optional equation.

(19) $(\uparrow \text{DF}) = \downarrow$

Given this equation, *who* will be assigned some discourse function (such as FOCUS) in addition to SUBJ without need for the complementizer projection. The Economy of Expression principle prohibits unnecessary c-structure, so a complementizer projection will be ungrammatical.

Next, there are languages in which the gap of a long distance dependency can only be a SUBJ, such as many Austronesian languages. This observation, originally due to Keenan and Comrie (1977), suggests that SUBJ has a special status in long distance dependency constructions. In English, this special status is manifested, ironically, in what appears to be a special restriction on SUBJ extraction: the "*that*-trace" effect.

(20) a. Which shelf do you think I put the book on ___?
b. Which shelf do you think that I put the book on ___?
c. Who do you think ___ put the book on the shelf?
d. *Who do you think that ___ put the book on the shelf?

Finally, note the similarity between the outside-in equation for long distance dependencies and the functional control equation.

(21) a. (↑ DF) = (↑ COMP* GF)
b. (↑ AF) = (↑ XCOMP GF); GF = SUBJ

In the functional control equation, the controllee is constrained by Universal Grammar to be SUBJ. One possible explanation for this is that SUBJ, uniquely among the argument functions, is an overlay function and therefore not exclusively related to its governing predicate.[4] Under the same logic, one would expect the gap end of the long distance equation to be SUBJ.

We will account for the special status of SUBJ extraction by adopting a mixed analysis, under which both outside-in and inside-out licensing of long distance dependencies are possible. Outside-in licensing is constrained to cases where the gap is SUBJ, while inside-out licensing involves (for English, at least) an empty category in c-structure. Languages that only allow extraction of SUBJ only allow outside-in licensing of long-distance dependencies. This approach is similar to one taken in early constraint-based theories; in both Generalized Phrase Structure Grammar (Gazdar 1981) and early LFG (Falk 1983b) it was proposed that SUBJ extraction is unique in not involving a structural empty category.

To conclude, long distance dependencies are licensed in three different ways.[5]

(22) a. matrix subjects: (↑ DF) = ↓ annotated to the SUBJ
b. embedded subjects (↑ DF) ≠ (↑ SUBJ) ⇒ (↑ DF) = (↑ COMP$^+$ SUBJ)
c. nonsubjects: $XP \rightarrow \quad e \quad \uparrow = ((COMP * GF \uparrow) DF)$

As we observed above, the outside-in equation cannot be associated with the structural position of the DF, since the DF need not be overt. We will analyze it as an optional lexical specification in the lexical entries of verbs.

As for the *that*-trace effect, despite the tradition that sees it as a structural property it seems to be a lexical property of the complementizer. In some languages, the effect is triggered by some complementizers and not others.

---

[4]This is essentially the position of Falk (2000).

[5]In the equation for embedded SUBJs we have replaced the Kleene star with the Kleene plus, because no COMPs would be the same as the matrix SUBJ equation. It is also limited to apply in cases where the DF and SUBJ are not locally identified.

Sobin (1987) reports that this is even true for some speakers of English, who allow *that*-trace violations when the complementizer is *that* but obey the constraint when the complementizer is *if* or *whether*. The ultimate source of the constraint may be the larger independence of finite complements as compared with nonfinite complements (Givón 1990: 517). This independence is reinforced by explicitly marking the complement with the finite complementizer. Complementizers that trigger the *that*-trace effect would be ones that explicitly mark their complements for conceptual independence from the clauses in which they are embedded. Since the SUBJ is the one argument that can be related to a higher clause (because SUBJ is an overlay function), this independence is expressed in the syntax by prohibiting the identification of the SUBJ with something higher.

(23) $(\uparrow \text{SUBJ}) \neq ((\text{GF}^{+} \uparrow) \text{GF})$

In standard English, *that* will include this lexical specification.

6.2.4 On empty categories

The analysis in the previous section requires us to recognize an empty element in c-structure. While this is familiar (and uncontroversial) in transformational theories, it runs counter to the trend in constraint-based theories of syntax. There seem to be two basic reasons to question the existence of an empty category. First, it is not overt. C-structure represents the overt expression of a syntactic object. Representing covert elements in c-structure violates the defining characteristic of c-structure. Furthermore, it poses potential problems for language comprehension, as the hearer would not know to parse an empty position. Second, there is no evidence for empty elements. They represent an unprovable (and unfalsifiable) theoretical construct. A theory that can do without them is therefore preferable. One theory that can do without them is the version of LFG in which all licensing of long distance dependencies is achieved through outside-in designation. In this section, we will argue that, at least for English, these objections to empty categories, while not without merit, are overstated.

We begin with the question of whether empty categories can be said to be "overt." Surprisingly, the answer is yes, at least for languages like English. An empty category is a position in the c-structure in which something should be present but is not. English is a language in which word order is very rigid and complements are rarely omissible. In the VP *put on the shelf* something clearly is missing. The verb *put* takes an OBJ argument, and OBJ in English

is invariably realized by an NP or DP in postverbal position within the VP. Unlike other languages, English does not allow the OBJ of *put* to be omitted and understood elliptically, even when the discourse context is clear. Unlike other languages, English does not allow the OBJ to scramble to other positions in the clause. Given these properties of English, it is clear that something is missing *in the c-structure*. An empty element is simply the formal device to represent this: the something missing is there structurally, but unfilled or empty. In this sense, the empty element can be said to be overt, and thus a legitimate element of c-structure.

Possible evidence for empty categories has been discussed by Bresnan (1995a, 2000). We will outline it here. The argument is based on the "weak crossover" effect.

(24) a. Who loves his mother? (*who* = *his*, possibly)
b. *Who does his mother love? (*who* ≠ *his*)

Bresnan argues that in English the following constraint holds.

(25) An operator O may not bind a pronoun P if the rightmost part of the c-structure correspondent of P precedes the rightmost part of the c-structure correspondent of O.[6]

This constraint accounts for the ungrammaticality of (24b), but only if we postulate an empty category in c-structure. The c-structure correspondent of the pronoun consists of one part: the [SPEC, DP] of the DP *his mother*. The operator in f-structure has two functions: FOCUS and OBJ. In a c-structure with no empty category, this operator has a one-part c-structure correspondent as well, [SPEC, CP]. However, [SPEC, CP] precedes the position of the c-structure correspondent of the pronoun, which should lead us to believe that coreference is possible. On the other hand, if there is an empty category in the post-VP position it is also part of the c-structure correspondent of the operator. It is the rightmost part of this correspondent, and it follows the c-structure correspondent of the pronoun, correctly disallowing coreference.

6.3 Islands and pied piping

It is well known that there are restrictions on the relation between filler and

[6]Or, if P f-precedes O, under the definition of f-precedence that Bresnan assumes.

gap in long distance dependency constructions. These restrictions have come to be known collectively as island constraints. A major contribution to the understanding of these constraints is Kaplan and Zaenen's (1989) observation that they are based on grammatical functions, not structure.

The long distance dependency equations as we have stated them specify that only the function COMP may occur on the path between filler and gap. The inside-out equation associated with the c-structure gap is:

(26) $\uparrow = ((\text{COMP}^* \text{ GF } \uparrow) \text{ DF})$

Since NPs and DPs cannot bear the function COMP, the Complex NP Constraint follows.

(27) *What did you deny the claim that you put *e* on the shelf?
would require *e* to be annotated $\uparrow = ((\text{OBJ COMP OBJ } \uparrow) \text{ DF})$

Similarly, a clause functioning as SUBJ is an island (the Subject Condition).

(28) *What do you think that [to put *e* on the shelf] would be a good idea?
would require *e* to be annotated $\uparrow = ((\text{COMP SUBJ OBJ } \uparrow) \text{ DF})$

Extraction from adjuncts is generally ungrammatical, as shown by the following contrast.

(29) a. Which table did he put a book on?
b. *Which table did he use a computer on?

This cannot be expressed directly in a c-structure-based approach, which instead has to talk about adjuncthood indirectly through claims about distinct structural positioning. This can be made to work in some languages (such as English), but Kaplan and Zaenen claim that there are languages (Icelandic is their example) in which the islandhood facts are the same but no structural distinction between complements and adjuncts can be motivated. A functional account of islands can refer to complements and adjuncts explicitly.

However, restricting the path to COMPs is too restrictive. The path can also include the functions XCOMP and OBL_{θ}.

(30) a. What did she seem to put *e* on the table?
↑ = ((XCOMP OBJ ↑) DF)

b. Which shelf did they take the book out of *e*?
↑ = ((OBL$_\theta$ OBL$_\theta$ OBJ ↑) DF)

We can accommodate this by altering the functional uncertainty equations.

(31) a. embedded subjects

$$(\uparrow \mathrm{DF}) \neq (\uparrow \mathrm{SUBJ}) \Rightarrow (\uparrow \mathrm{DF}) = (\uparrow \left\{ \begin{array}{c} \mathrm{COMP} \\ \mathrm{XCOMP} \\ \mathrm{OBL}_\theta \end{array} \right\}^{+} \mathrm{SUBJ})$$

b. nonsubjects:

$$\mathrm{XP} \rightarrow \quad e$$
$$\uparrow = ((\left\{ \begin{array}{c} \mathrm{COMP} \\ \mathrm{XCOMP} \\ \mathrm{OBL}_\theta \end{array} \right\}^{*} \mathrm{GF} \uparrow) \mathrm{DF})$$

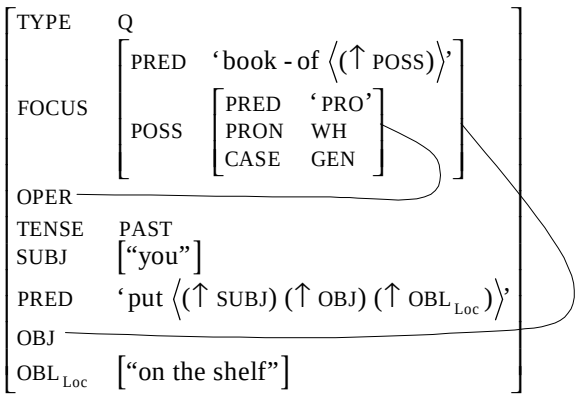
Island phenomena thus provide evidence for a functional approach to long-distance dependencies.

Related to islands is the phenomenon of "pied piping," where the FOCUS or TOPIC includes more than just the operator. This is related to islands because this is often used as a strategy to avoid island violations.

(32) a. *Whose did you put book on the shelf?
b. Whose book did you put on the shelf?

In this case, the FOCUS is *whose book*, but the operator is just *whose*. (Here we use OPER as the name of the grammatical function of the operator.)

(33)



Our ID rule for CP, repeated below, does not allow this.

(34)

$$\text{CP} \rightarrow \underset{(\uparrow \text{DF}) = \downarrow}{\text{XP}} \quad , \quad \underset{\uparrow = \downarrow}{\text{C}'}$$
$$(\downarrow \text{PRON}) =_{c} \text{WH}$$

This ID rule does not distinguish the DF (FOCUS or TOPIC) from the operator. It is the operator that must have the [PRON WH] feature, not the DF. The operator is embedded at some undetermined depth within the DF (including potentially being identical); functional uncertainty can be used to model this.[7]

(35)

$$\text{CP} \rightarrow \underset{(\uparrow \text{DF}) = \downarrow}{\text{XP}} \quad , \quad \underset{\uparrow = \downarrow}{\text{C}'}$$
$$(\uparrow \text{OPER}) = (\downarrow \text{GF}^{*})$$
$$(\uparrow \text{OPER PRON}) =_{c} \text{WH}$$

Similarly, the non-*wh* ID rule needs to be updated.

[7]This analysis is based on Kaplan and Bresnan (1982). Kaplan and Bresnan use the function name Q for our OPER.

(36) $\text{IP} \rightarrow$ XP $(\uparrow \text{DF}) = \downarrow$ $(\downarrow \text{GF}^* \text{PRON}) \neq \text{WH}$, IP $\uparrow = \downarrow$

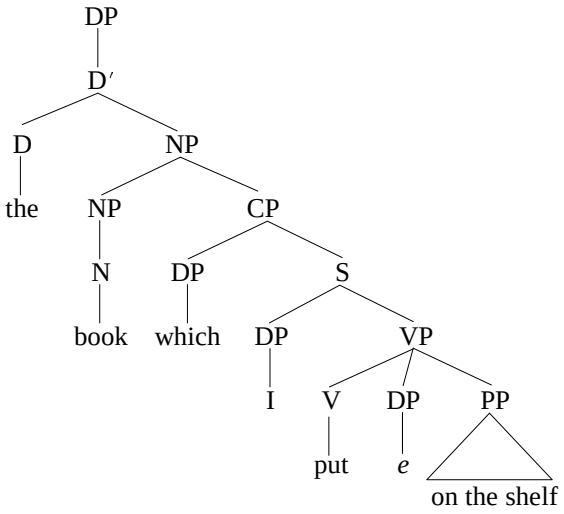
6.4 Relative clauses

We conclude this chapter with discussion of restrictive relative clauses in English.[8] There are interesting intricacies in the structure of relatives, which we will describe using the analysis of long distance dependencies developed in this chapter.

The simplest type of relative clause is one with a relative pronoun. Relative pronouns, like interrogative pronouns, are *wh* elements and therefore occupy [SPEC, CP] position. The discourse function held by the relative pronoun is TOPIC.

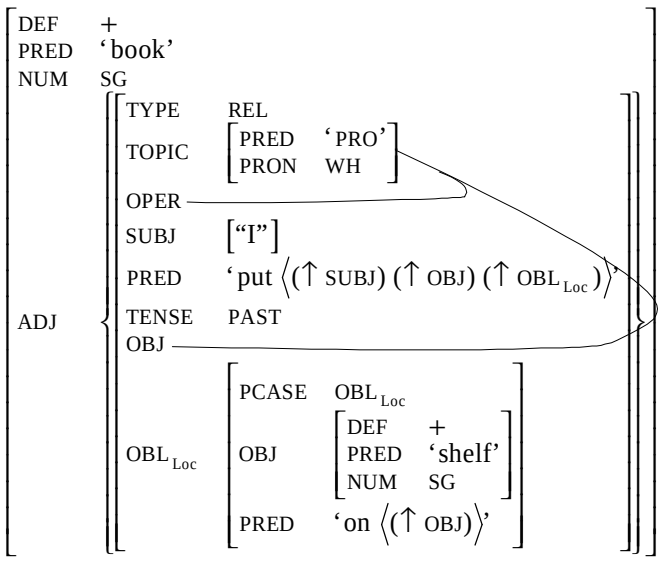
(36) a. the book which I put on the shelf

b.



[8]We will not discuss other kinds of relative clauses, such as nonrestrictive relatives or free relatives.

c.



Our existing rules generate such relative clauses, with the addition of an ID rule allowing CP to be adjoined to NP.

The status of the relative pronoun as TOPIC can be shown by the interaction of relative clauses with other constructions. TOPIC is a syntacticized discourse function. As such, it represents part of the interface between syntax and discourse. A topic in discourse grammar is old information; it is therefore incompatible with constructions that serve to introduce new entities into the discourse (see Bresnan and Mchombo 1987). One such construction is the *there* construction. Note:

(38) *the book which there is *e* on the table

This confirms the analysis of *which* as TOPIC.

However, this is not the only type of relative clause in English. There are also relative clauses that lack a relative pronoun. Such relatives can be either CPs introduced by the complementizer *that* or IPs (or Ss) introduced by no complementizer.

(39) a. the book that I put on the shelf
b. the book I put on the shelf

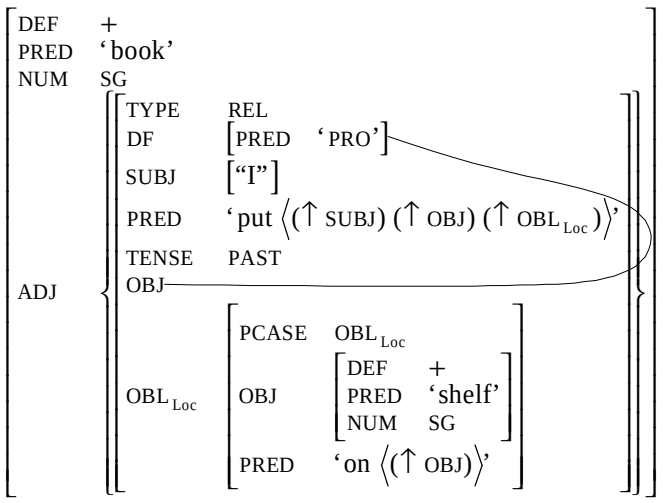
Unlike the construction introduced by a relative pronoun, these do not appear to involve a *wh* element. However, they do include a gap, just like the *wh* relative. Furthermore, it can be shown that they involve long distance dependencies: they obey island constraints.

(40) a. *the book that I denied the claim that I put *e* on the shelf
b. *the book that to put *e* on the shelf would be a mistake

It is for this reason that derivational theories have postulated the movement of an "empty operator" or a deleted relative pronoun.

The f-structure of *wh*-less relatives must be something like:

(41)



Such an f-structure embodies the analysis of *wh*-less relatives as long distance dependencies.

Next, we must determine the nature of the DF in this f-structure. In *wh* relatives, the functions TOPIC and OPER are both involved because, as we have seen, sometimes the operator is embedded within the fronted phrase. In *wh*-less relatives, there is no evidence for two distinct elements. Furthermore, there is evidence against analyzing the unexpressed filler as TOPIC. Unlike the fronted *wh* relative pronoun, the unexpressed relative pronoun can be used in

the *there* construction.[9]

(42) a. *the book which there is *e* on the table
b. the book that there is *e* on the table

We therefore conclude that the unexpressed filler in the *wh*-less relative clause has the function OPER. In a sense, this is similar to the notion of "empty operator" in derivational theories, with "empty" reinterpreted as meaning "present in f-structure but not c-structure." Since this "empty operator" is a property of the relative clause construction, the most natural source for it would be the ID rule that licenses the relative clause.

(43)

$$\text{NP} \rightarrow \quad \text{NP} \quad , \qquad \left\{ \begin{matrix} \text{CP} \\ \text{IP} \\ \text{S} \end{matrix} \right\}$$
$$\uparrow = \downarrow \qquad\qquad \downarrow \in (\uparrow \text{ADJ})$$
$$\left((\downarrow \text{OPER PRED}) = \text{`PRO'}\right)$$

Not surprisingly, infinitives (which are CPs) can also be relative clauses. However, the details are interesting. Note the following.

(44) a. a shelf [on which to put the books]
b. a shelf [to put the books on]
c. *a shelf [which to put books on]

What these examples show is that an overt *wh* element is possible only in the pied piping construction. From the perspective of the LFG analysis developed here, this means only when the TOPIC (if there is one) is not identical to OPER. The Economy of Expression principle provides a way to rule out (c): if the same f-structure can result from a c-structure with fewer phrasal nodes, the more complex c-structure is ruled out. It appears that in infinitival relatives, unlike finite ones, an "empty operator" can also have the function of TOPIC,

[9]This contrast was noted in a post by Joan Bresnan on 21 March 2000 to the LFG List, and Bresnan's judgment matches that of the author of this textbook. There are apparently people for whom the contrast does not hold. It is possible that there are idiolectal differences, and that for some people the understood relativized element *is* TOPIC.

thus blocking an overt TOPIC. This can be achieved by associating an optional equation with the infinitival complementizer *to*.

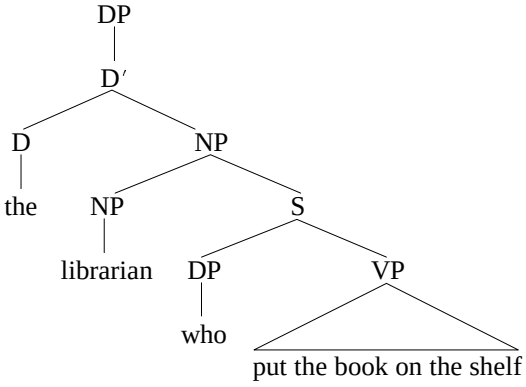
(45) (↑ OPER) = (↑ TOPIC)

6.5 Subjects revisited

In light of the addition of relative clauses to our analysis of long distance dependencies, we need to reconsider the "extraction" of matrix SUBJs. Our analysis is that matrix SUBJs receive a DF in situ. Note how this applies to a relative clause.

(46) a. the librarian [who put the book on the shelf]

b.



c.
$$\begin{bmatrix}
\text{DEF} & + \\
\text{PRED} & \text{`librarian'} \\
\text{NUM} & \text{SG} \\
\text{ADJ} & \left\{ \begin{bmatrix}
\text{TYPE} & \text{REL} \\
\text{DF} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{PRON} & \text{WH} \end{bmatrix} \\
\text{SUBJ} & \\
\text{PRED} & \text{`put} \left\langle (\uparrow \text{SUBJ})\ (\uparrow \text{OBJ})\ (\uparrow \text{OBL}_{\text{Loc}}) \right\rangle\text{'} \\
\text{TENSE} & \text{PAST} \\
\text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`book'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\
\text{OBL}_{\text{Loc}} & \begin{bmatrix}
\text{PCASE} & \text{OBL}_{\text{Loc}} \\
\text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`shelf'} \\ \text{NUM} & \text{SG} \end{bmatrix} \\
\text{PRED} & \text{`on} \left\langle (\uparrow \text{OBJ}) \right\rangle\text{'}
\end{bmatrix}
\end{bmatrix} \right\}
\end{bmatrix}$$

The DF in the relative clause could be either TOPIC or OPER.

However, the OPER/TOPIC of a relative clause can also be identified with the SUBJ when there is a complementizer and no relative pronoun.

(47) a. the librarian that put the book on the shelf
b. a librarian to put the book on the shelf

On the other hand, in the absence of either a *wh* relative pronoun or a complementizer, the relativized element cannot be understood as SUBJ of the relative clause.

(48) *[$_{\text{DP}}$ the librarian put the book on the shelf]

It therefore seems plausible to identify this ability as a lexical property of the complementizers. In the case of *that*, the SUBJ is identified with OPER (since, as we saw above, there is no TOPIC in *that* relatives). With *to*, on the other hand, the identification must be with TOPIC. This is because there is a contrast between relative clauses and interrogatives: FOCUS/OPER and SUBJ cannot be identified in an infinitival interrogative.

(49) *I asked [who to put the books on the shelf].

These facts concerning SUBJs are quite intricate. They clearly show that SUBJ extraction is different from the extraction of other elements. They also show the need to develop careful detailed analyses.

Additional readings

As mentioned in the text, Kaplan and Zaenen (1989) originated the analysis of long-distance dependencies in terms of functional uncertainty. The formal implications are also discussed in Kaplan and Maxwell (1988a).

The earlier "constituent control" formalism of Kaplan and Bresnan (1982), although superseded by functional uncertainty, was the basis for some early studies such as Zaenen (1983) and Falk (1983b). The former discussed syntactic effects along the extraction path in some languages ; the latter dealt with the *that*-trace effect, and proposed an analysis which shares some features with the analysis given here.

Discourse functions and structure in Russian are discussed within both GB and LFG by King (1995). Empty categories in German are considered in Berman (1997). The restriction of extraction to SUBJ in Tagalog is covered in Kroeger (1993). Bresnan (1998) examines crosslinguistic variation in weak crossover effects. As noted in the text, the discourse effects of the status of the relativized element as TOPIC were originally noted by Bresnan and Mchombo (1987).

Much of the detail in the analysis of English long-distance dependencies in this chapter is original, and some (including the subject/nonsubject distinction and the analysis of the *that*-trace effect) is drawn from Falk (2000).

Exercises

1. Explain the ungrammaticality of each of the following:

   a. *I think this book that you should read.
      (cf. ✓I think that this book you should read.)

   b. *What do you think I read the book?

   c. *Who did you deny the claim that I saw *e*?

   d. *This book, [that he read *e*] is most surprising.

   e. *What did he buy the newspaper [after he ate *e*]?

f. *[$_{DP}$ a shelf on which for you to put the book]
(cf. ✓a shelf for you to put the book on)

2. The analysis of the English possessive *'s* has always been a problem. The traditional analysis treats it as an inflectional morpheme (a Case marker). However, it does not always appear on the ostensibly genitive noun. For example, in a DP like *a friend of my wife's daughter* (meaning 'the daughter of a friend of my wife'), the noun that "should" be genitive is *friend*. An alternative analysis would be to treat *'s* as a syntactic head which takes a DP complement. However, heads in English precede their complements, and *'s* follows its ostensible complement. In LFG, the Lexical Integrity Principle forces an analysis of *'s* as an affix (i.e. the more traditional analysis). Show how inside-out functional uncertainty can be used to overcome the problems with this analysis.

7

Anaphora

7.1 Overview

We turn now to the theory of anaphora, the equivalent of the "Binding Theory" of Government/Binding. In some versions of transformational theory, Binding Theory is central because it relates not only to actual anaphoric relations, but also to movement constructions. In LFG, Binding theory is not needed for regulating the operation of movement. But, of course, any theory needs to be able to account for the facts of anaphora.

The conventional binding theory divides anaphoric NPs (or DPs) into two groups: reflexives/reciprocals and pronouns. These two types of anaphoric elements are generally in complementary distribution, and many theories of anaphora treat this bifurcation as central. In GB and MP, this is taken even further. The idea is that reflexives and reciprocals are elements that are grammatically assigned an antecedent, while pronouns are not. Consequently, the term *anaphor* is used in GB with the restricted meaning of "reflexive/reciprocal," and pronouns are excluded from the class of anaphors.

A second claim of the conventional approach to anaphora in structural/derivational theories is the centrality of the notion c-command.[1] The one condition that a grammatical antecedent (binder) must meet is c-command; not even linear precedence is recognized as a condition on anaphoric relations.

As we will see in this chapter, both of these assumptions are open to question. The LFG approach to anaphora rejects both of them in favor of an approach that is lexical (in that properties of anaphoric elements are lexical properties) and functional (in that anaphoric relations are defined in terms of

[1]A c-structure node A c-commands a c-structure node B iff the first branching node dominating A also dominates B.

f-structure properties).

7.2 Prominence

We will begin by considering the c-command restriction of conventional theories of anaphora, which we state informally.

(1) A binds B iff A and B are coindexed and A c-commands B.

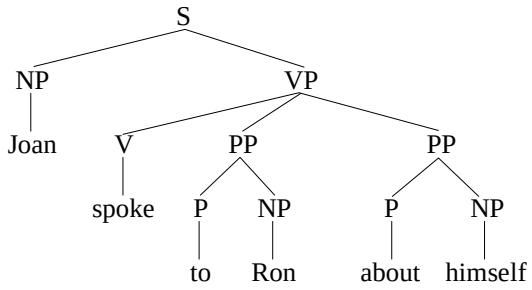
The fundamental idea behind the c-command restriction is that an antecedent has to be more prominent than the anaphor, in some sense. C-command is a hypothesis as to the nature of this prominence. The question is whether it is the right concept. We will attempt to answer that question by looking at the evidence for prominence conditions.

First, it has been known for a long time that there is a precedence condition on anaphora. Although this is clearer in some other languages, it can also be seen in English. Note the following contrast:

(2) a. Joan spoke [to Ron_i] [about himself_i].
b. *Joan spoke [about himself_i] [to Ron_i].

Under the simplest assumptions about c-structure, which we have adopted, the two PPs are sister nodes, both immediately dominated by the same VP node.

(3)



The PPs thus mutually c-command each other, and neither NP c-commands the other. The ordering constraint cannot be derived from c-command in terms of either the NPs or the PPs. Of course, more complex c-structures can be proposed for such sentences. For example, if one limits c-structures to binary branching structures, the outermost PP might be adjoined to the VP

node. However, this would also give us the wrong result, as the later PP would then asymmetrically c-command the earlier PP.

We conclude, then, that one aspect of prominence relevant for the theory of anaphora is linear precedence. While one is always free to represent the relation of linear precedence in terms of a tree structure that resembles constituent structure, as is sometimes done, and then identify linear precedence with asymmetrical c-command, the fact remains that what is being modeled is linear order, and that it is linear order that is relevant here. In LFG, linear order is modeled as a relation between c-structure nodes, represented graphically as left-to-right positioning, or as a relation between f-structure elements based on this left-to-right c-structure positioning (f-precedence). Since anaphoric elements are represented as such at f-structure, we will state the condition in terms of f-precedence. However, as we have already discussed, f-precedence is the f-structure projection of a c-structure relation.

(4) **Linear Precedence Condition on Anaphora**
A binds B iff A and B are coindexed and A f-precedes B.

In a theory in which linear precedence is represented in terms of asymmetric c-command at some level of "linear precedence structure", this condition would take on the superficial appearance of the c-command condition. However, this would not be c-command in the sense of c-structure.

There are other dimensions of prominence as well. Compare the sentences above with the following.

(5) *Joan spoke [about Ron$_i$] [to himself$_i$].

This sentence is ungrammatical even though the intended antecedent precedes the anaphor. Identifying the relevant dimension turns out to be somewhat more complicated, however. Again, it clearly is not c-command. However, the literature on anaphora has cited both thematic (Jackendoff 1972) and functional (Bresnan 2000) dimensions.

(6) a. **Thematic Hierarchy Condition on Anaphora**
A binds B iff A and B are coindexed and A is higher than B on the thematic hierarchy.

b. **Relational Hierarchy Condition on Anaphora**
A binds B iff A and B are coindexed and A is higher than B on the relational hierarchy.

The distinction between the two hierarchies is (hopefully) ultimately an empirical question. The problem is that it is hard to tease apart because (by virtue of the nature of linking, LMT or its equivalent in other theories) the hierarchies usually coincide. Cases that seem to argue for one or the other can be reanalyzed. For example, our sample sentence seems to be an argument for a thematic hierarchy condition, since both arguments are OBL_{θ}. But it could also be taken as evidence that not all members of the OBL_{θ} family share the same position on the relational hierarchy. Conversely, consider:

(7) a. We sold the slave to himself.
b. *We sold himself to the slave.
c. We sold the slave himself.[2]
d. *We sold himself the slave.

In the version with the PP, the OBJ (Theme) can bind the OBL_{Goal} (Goal) but not conversely. In the ditransitive case, the OBJ (Goal) can bind the OBJ_{θ} (Theme) but not conversely. This looks like a situation where grammatical functions are relevant. However, our analysis of the ditransitive alternation hypothesizes a difference in the thematic roles: the Goal is also a Beneficiary in the ditransitive, thus outranking the Theme.

The clearest cases that show that grammatical functions must play a role are ones like the following.

(8) a. The derivationalist contradicted himself.
b. *Himself was contradicted by the derivationalist.

This contrast distinguishes the two hierarchies. Thematic roles cannot explain the difference, since in both cases *the derivationalist* is the Agent and *himself* is the Patient. On the other hand, the relational hierarchy predicts the contrast, since SUBJ can bind OBJ but ADJ (or OBL_{by}) cannot bind SUBJ. Interestingly, the c-command condition would also achieve the correct results

[2]This sentence also has an irrelevant reading with *himself* functioning as an intensifier. The intended reading is the one where *himself* is the Theme.

in English. However, similar facts hold in languages with different configurational properties from English; since grammatical functions are more uniform across languages, this contrast should be attributed to the relational hierarchy of grammatical functions.

The kind of prominence relevant for anaphora thus seems to be defined independently at three distinct levels of representation:

- c-structure:[3] (f-)precedence
- f-structure: relational hierarchy
- a-structure:[4] thematic hierarchy

We define a cross-level notion of **rank**.

(9) a. A outranks B at c-structure iff A (f-)precedes B.
b. A outranks B at f-structure iff, for some C containing B, A and C are in the same local f-structure and A is higher than C on the relational hierarchy.[5]

---

[3]The statement of the linear precedence condition, which we are now restating as c-structure rank, in terms of f-precedence rather than directly in terms of (c-)precedence is, as noted above, a formal matter, the result of the fact that anaphors are represented as such at f-structure. We leave open the possibility that rank may be used in other situations where direct c-structure precedence is enough, and thus define c-structure rank to include both precedence and f-precedence.

Note also that a-structure prominence is also represented in derivative fashion at f-structure, in the PRED value of the verb. Thus all three kinds of prominence, even though they represent prominence at different levels, are accessible at f-structure. This is a result of the projection architecture of LFG.

[4]We are assuming here that thematic prominence is defined (or at least checked) at the (syntactic) level of a-structure rather than the semantic/conceptual level we are modeling informally as θ-structure. While we have no direct evidence for this (and it is unclear what would constitute direct evidence), it is interesting to note that while relative position on the thematic hierarchy is relevant for anaphora, actual thematic role is not. Since a-structure is organized by relative hierarchical position and θ-structure by actual thematic roles (or conceptual relations), this may be indirect evidence for the relevance of a-structure.

[5]This definition follows Bresnan (2000)

c. A outranks B at a-structure iff A is higher than B on the thematic hierarchy.
d. A outranks B iff A outranks B at one or more levels of representation, as specified on a language-specific basis.
e. A binds B iff A and B are coindexed and A outranks B.

In English, f-structure rank is of greatest importance. The effects of the thematic hierarchy and linear order are only visible in a small set of cases involving two elements with grammatical function of identical rank.

This sensitivity to rank at different levels of representation is to be expected in a parallel structure theory like LFG. To the extent that it is justified on empirical grounds, it is evidence for a parallel structure theory. The GB attempt to reduce rank to c-command is empirically deficient.

We have found no need for the concept of c-command in our theory of anaphora. The job of c-command is taken by rank, which is empirically superior. However, there is no reason in principle to banish c-command; empirical evidence for c-command could be easily accommodated by LFG.

7.3 Types of anaphors

7.3.1 English

The other assumption of conventional binding theories is the bifurcation of anaphoric elements into reflexives/reciprocals and pronouns. Related to this is the idea that only reflexives/reciprocals are anaphors. In LFG, reflexives, reciprocals, and pronouns are all **anaphors**; they are all elements with syntactic requirements on what can antecede them.

The bifurcation of anaphors into two classes has some intuitive appeal with languages like English. We will begin by assuming it to be essentially correct, and determine the nature of binding conditions for English. As a preliminary characterization, we state Conditions A and B as follows.

(10) A: Reflexives and reciprocals must be bound locally.
B: Pronouns must be free locally.

Where binding theories differ is in the characterization of "locally." The most common characterization is based on argument functions. For example, Chomsky (1986) defines it in terms of "complete functional complex," a piece of structure in which all arguments selected by the head predicate are specified. In LFG, the natural level at which to specify such a segment of the

sentence is f-structure. It is called **(clause) nucleus**.

(11) A (clause) nucleus is the subpart of an f-structure consisting of a PRED feature and all the argument functions it selects.

We can restate binding conditions A and B in terms of the concept of nucleus.

(12) A: A reflexive or reciprocal must be bound in the minimal nucleus containing it.
B: A pronoun must be free in the minimal nucleus containing it.

Consider the following examples. The elements of the minimal nucleus containing the anaphor are bolded.

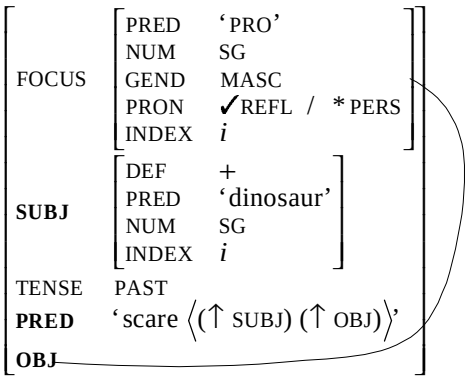
(13) a. The dinosaur$_i$ scared himself/*him$_i$.

b.
$$\begin{bmatrix} \textbf{SUBJ} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{`dinosaur'} \\ \text{NUM} & \text{SG} \\ \text{INDEX} & i \end{bmatrix} \\ \text{TENSE} & \text{PAST} \\ \textbf{PRED} & \text{`scare} \langle (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) \rangle' \\ \textbf{OBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{MASC} \\ \text{PRON} & \checkmark\text{REFL} \ / \ *\text{PERS} \\ \text{INDEX} & i \end{bmatrix} \end{bmatrix}$$

c. The antecedent outranks the anaphor (at all levels) and is located within the minimal nucleus containing the anaphor.

(14) a. Himself/*him$_i$, the dinosaur$_i$ scared.

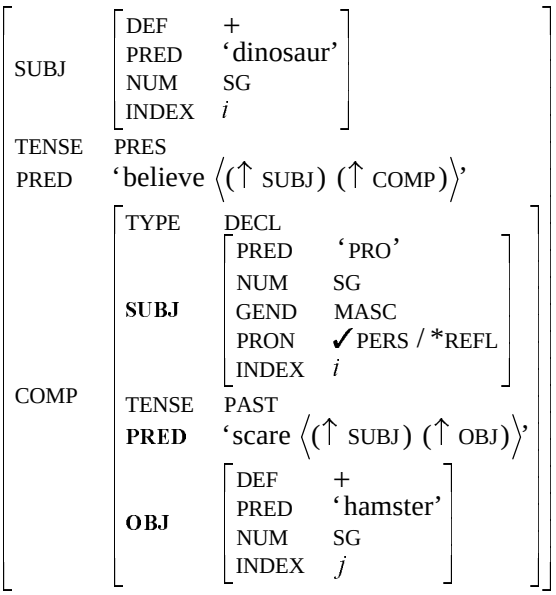
b.



c. The anaphor is the value of both FOCUS and OBJ. As OBJ, it is outranked by its antecedent and the antecedent is located within the minimal nucleus containing the anaphor.

(15) a. The dinosaur$_i$ believes that he/*himself$_i$ scared the hamster.

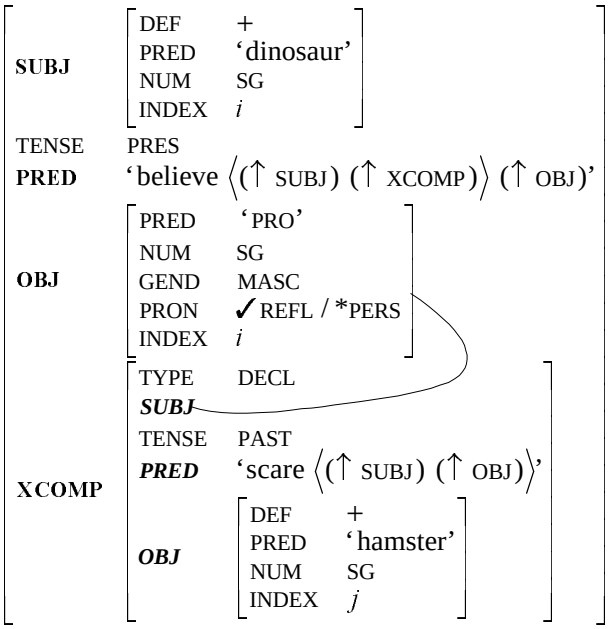
b.



c. The antecedent is outside the minimal nucleus containing the anaphor. The anaphor is therefore free in its minimal nucleus and must be realized as a personal pronoun, not a reflexive.

(16) a. The dinosaur$_i$ believes himself/*him$_i$ to have scared the hamster.

b.



c. The anaphor is both SUBJ and OBJ, in different nuclei. Both nuclei are marked (the lower/inner one is italicized). The intended antecedent is in the same nucleus as the OBJ, and outranks it. For this reason, the anaphor is a reflexive.

Because the nucleus defines the binding domain for reflexives and reciprocals, they can be called **nuclear** anaphors, while personal pronouns can be called **nonnuclear**.

Our binding conditions predict complete complementary distribution. However, this is not quite right. Note the following.

(17) a. Kirk and Picard admire their officers.
b. Kirk and Picard admire each other's officers.

Here we see a pronoun and a reciprocal in the same position, contrary to the prediction of complementarity. Note the f-structure.

(18)
$$\begin{bmatrix} \text{SUBJ} & [\text{``Kirk and Picard''}] \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`admire}\ \langle(\uparrow \text{SUBJ})\ (\uparrow \text{OBJ})\rangle\text{'} \\ \text{OBJ} & \begin{bmatrix} \text{DEF} & + \\ \textbf{POSS} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \vdots & \vdots \end{bmatrix} \\ \textbf{PRED} & \text{`officer - of}\ \langle(\uparrow \text{POSS})\rangle\text{'} \\ \text{NUM} & \text{PL} \end{bmatrix} \end{bmatrix}$$

Again, we have bolded the minimal nucleus containing the anaphor. This makes it clear that the antecedent is outside the minimal nucleus, and the anaphor is therefore free within its minimal nucleus. Our binding theory therefore predicts that only a nonnuclear anaphor should be grammatical. The key seems to be that this nucleus differs from others we have seen in lacking the function SUBJ. It has been hypothesized that the nucleus relevant for the binding of nuclear anaphors must contain a SUBJ. A (clause) nucleus that contains a SUBJ is called a **complete (clause) nucleus**. We revise our binding conditions accordingly.

(19) A: A nuclear anaphor (reflexive or reciprocal) must be bound in the minimal complete nucleus containing it.
B: A nonnuclear anaphor (personal pronoun) must be free in the minimal nucleus containing it.

It is this asymmetry between conditions A and B that accounts for the contrast.

This is still not all there is to anaphora in English. Dalrymple (1993) discusses more complex cases. First, note that we have said nothing about picture NPs, which tend to pose problems for all theories of anaphora. It has been argued by some, including Pollard and Sag (1994) in HPSG and Dalrymple in LFG, that reflexives in picture NPs are governed not by syntactic conditions but by discourse conditions. Other uses ("logophoric" reflexives, "intensive" reflexives) are also governed by discourse conditions. In general, anaphora seems to be governed not only by conditions at f-structure, a-structure, and c-structure, but also at information structure. We will not pursue this further here, however. Instead we will continue to focus on f-structure.

7.3.2 Other languages

The major problem with the standard binding theory, upon which we based the discussion in the previous section, is that the bifurcation into reflexives/reciprocals and personal pronouns is too restrictive crosslinguistically. Many languages have anaphoric elements whose properties cannot be characterized in terms of this categorization. We therefore must reject it, as we rejected the hypothesis that the c-command relation is central to anaphora. The material in this section is drawn from the seminal work of Dalrymple (1993).

We begin with informal characterizations of the properties of anaphors in three Indo-European languages: English (based on the previous section), Marathi (an Indic language spoken primarily in Maharashtra state, India), and Norwegian (a Northern Germanic language, spoken primarily in Norway).

(20) a. English:
he—(pronoun) may not be bound in the minimal nucleus.
himself—(reflexive) must be bound in the minimal complete nucleus

b. Marathi
to—(pronoun) may not be bound to a coargument (i.e. within its minimal nucleus).
swataah—("local reflexive") two dialects:
•must be bound by a SUBJ in the minimal complete nucleus.
•must be bound by a SUBJ in the minimal finite clause.
aapaṇ—("long-distance reflexive") must be bound by something that is not a coargument (i.e. not in the minimal nucleus). The binder must also be the most prominent element in its a-structure; i.e. $\hat{\theta}$ or "logical subject."

c. Norwegian
ham—(pronoun) may not be bound within its minimal nucleus
seg—("long-distance reflexive") must be bound by a SUBJ within the minimal finite clause but not within its minimal nucleus.[6]
seg selv—must be bound by a SUBJ within its minimal nucleus.

[6]There is also an intransitivizing ("lexical reflexive") use of *seg*. This is irrelevant.

ham selv—must be bound by a nonSUBJ within the minimal complete nucleus.
sin—(possessive reflexive) must be bound by a SUBJ within the minimal finite clause.
hans—(possessive pronominal) may not be bound by the next higher SUBJ (i.e. a SUBJ in the minimal complete nucleus).
hverandre—(reciprocal) must be bound in minimal complete nucleus.

This array of facts makes it clear that a simple reflexive/pronoun bifurcation is empirically inadequate. Note especially Norwegian, which has seven distinct types of anaphors.

The conclusion that Dalrymple draws from these facts is that the properties of anaphors are lexical properties. Each anaphor is lexically marked, not with features like GB's [±anaphoric] and [±pronominal], but with a more detailed specification of coreferential possibilities. This specification takes the form of two types of constraints: domain constraints (where the antecedent can be relative to the anaphor) and antecedent constraints (antecedent must/must not be a SUBJ). Constraints can be positive or negative. The array of anaphoric properties above shows that the domain can be characterized in terms of nuclei (i.e. the feature PRED), complete nuclei (PREDs containing the function SUBJ), and the feature TENSE.

Dalrymple categorizes the following possible conditions:

(21) Antecedent conditions
a. Subject Binding Condition– the anaphor must be bound by a SUBJ.
b. Subject Disjointness Condition– the anaphor may not be bound by a SUBJ.
c. GF Binding Condition– the anaphor must be bound by some element, unrestricted in terms of GF.
d. GF Disjointness Condition– the anaphor may not be bound by any element (i.e. bearing any GF).

(22) Domain conditions
a. Coargument Binding Condition– the anaphor must be bound to an element in the minimal nucleus containing it.
b. Coargument Disjointness Condition– the anaphor may not be bound to an element in the minimal nucleus containing it.

c. Minimal Complete Nucleus Binding Condition– the anaphor must be bound to an element in the minimal complete nucleus containing it.
d. Minimal Complete Nucleus Disjointness Condition– the anaphor may not be bound to an element in the minimal complete nucleus containing it.
e. Minimal Finite Domain Binding Condition– the anaphor must be bound to an element in the minimal f-structure containing it that includes the attribute TENSE.
f. Minimal Finite Domain Disjointness Condition– the anaphor may not be bound to an element in the minimal f-structure containing it that includes the attribute TENSE.
g. Root S Binding Condition– the anaphor must be bound by some element in the sentence.
h. Root S Disjointness Condition– the anaphor may not be bound by some element in the sentence.

The anaphors of English and Marathi can be specified in terms of these constraints.

(23) a. English:
he—(pronoun) GF Disjointness Condition + Coargument Disjointness Condition
himself—(reflexive) GF Binding Condition + Minimal Complete Nucleus Binding Condition

b. Marathi
to—(pronoun) GF Disjointness Condition + Coargument Disjointness Condition.
swataah—("local reflexive") two dialects:
•SUBJ Binding Condition + Minimal Complete Nucleus Binding Condition.
•SUBJ Binding Condition + Minimal Finite Domain Binding Condition.
aapaṇ—("long-distance reflexive") GF Binding Condition + Root S Binding Condition + Coargument Disjointness Condition + a-structure condition restricting antecedent to $\hat{\theta}$.

This system of constraints is more complex than the conventional

binding theory. However, it is also empirically more adequate. As noted at the beginning of this chapter, this account is lexical and functional.

7.4 Formalization

If the properties of anaphors are lexically specified, this specification must take the form of a functional equation in the lexical entry of the anaphor. From the anaphor, we have to go outward to find the f-structure containing the antecedent. Since we have to go outward, an inside-out equation is required. Since the distance is potentially infinite, an inside-out functional uncertainty equation is required. Once we have reached the function containing the antecedent, an outside-in path of length 1 is required to actually get the antecedent. Schematically, the specification of a possible antecedent is:

(24) ((PathOut ↑) PathIn)
where PathOut is a functional uncertainty expression and PathIn is a single GF specification.

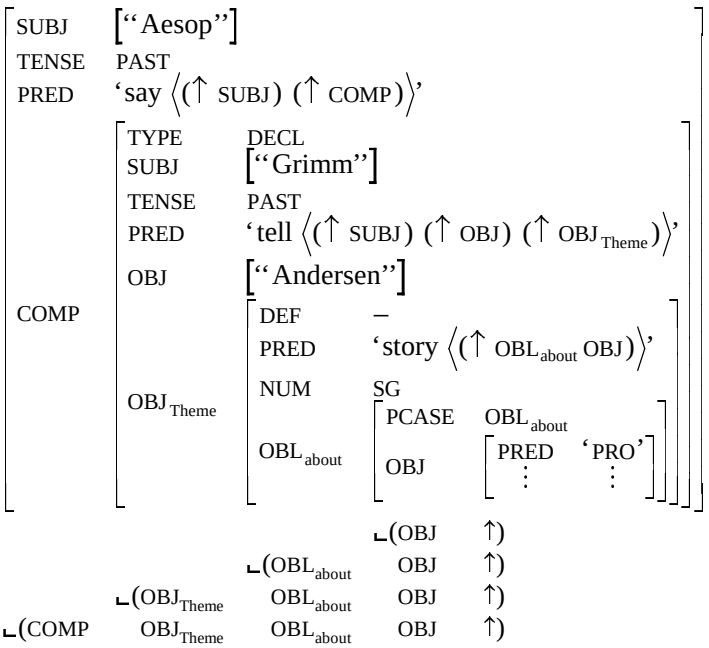
As we have seen, PathIn is universally limited to being either SUBJ or (unspecified) GF. PathOut is GF*.

To further develop this formal approach to anaphora, let us consider a concrete example, one involving the English reflexive *himself*.

(25) Aesop said that Grimm told Andersen a story about himself.

In this sentence, *himself* can be coreferential with either *Grimm* or *Andersen*, but not with *Aesop*. Consider the f-structure of this sentence, with the various inside-out paths from the anaphor designated as shown.

(26)



Each inside out designator refers to the subsidiary f-structure the left bracket of which is shown by the symbol "⌞".

Each of these inside-out paths (the PathOut of the schema above) represents a potential domain within with *himself* might be able to find an antecedent, what we might call a Domain Path. If all were actually possible, the lexical specification for the antecedent of *himself* would be simply:

(27) ((GF* ↑) GF)

However, the longest of these paths represents an ungrammatical domain for the antecedent of *himself*. This is because the antecedent of *himself* must be within the minimal complete nucleus containing *himself*. In other words, the problem is that the f-structure which is the value of the function COMP contains the attribute SUBJ. It is important to note that this is different from islands in long distance dependencies. Islands are generally due to constraints on functions on the path itself. With anaphora, the constraints do not relate to which attributes can appear on the path; instead, they are **off-path constraints**. After the GF of the anaphor itself, each step along the path must

be checked to ensure that the value of the GF named does not itself contain the attribute SUBJ. Similarly, if the restriction had been "minimal nucleus" instead of "minimal complete nucleus," the attribute PRED would have to be checked for. The notation for these off-path constraints, which apply to all but the innermost function (the function of the anaphor), is:

(28) $\neg (\rightarrow \text{ATTRIB})$

This is read "does not include the attribute ATTRIB". So the lexical specification for the English reflexive will be:

(29) $((\text{GF}^* \text{GF} \uparrow) \text{GF INDEX}) = (\uparrow \text{INDEX})$
$\neg (\rightarrow \text{SUBJ})$

This is a formal specification of Principle A, or of a combination of the GF Binding Condition and the Minimal Complete Nucleus Binding Condition. It requires the reflexive to share an INDEX with something higher, as long as the path does not cross something with a SUBJ. What is not included is the notion that the antecedent must outrank the anaphor, which we will assume is an independent condition on the instantiation of anaphoric equations.

English (nonnuclear) personal pronouns will have a negative specification, preventing them from sharing an INDEX with any outranking element in their minimal nucleus.

(30) $((\text{GF}^* \text{GF} \uparrow) \text{GF INDEX}) \neq (\uparrow \text{INDEX})$
$\neg (\rightarrow \text{PRED})$

This lexical specification correctly allows a pronoun to be coindexed with something outside the minimal nucleus or with nothing at all.

This formalism also allows us to specify lexical entries for the Marathi anaphors considered earlier.

(31) a. *to*– (pronoun)
$((\text{GF}^* \text{GF} \uparrow) \text{GF INDEX}) \neq (\uparrow \text{INDEX})$
$\neg (\rightarrow \text{PRED})$

b. *swataah*—("local reflexive"; complete nucleus dialect)
$((\text{GF}^* \text{GF} \uparrow) \text{SUBJ INDEX}) = (\uparrow \text{INDEX})$
$\neg (\rightarrow \text{SUBJ})$

: c. *swataah*—("local reflexive"; finite domain dialect)

$$((\ \text{GF}^*\ \ \text{GF}\uparrow)\ \text{SUBJ INDEX}) = (\uparrow \text{INDEX})$$
$$\neg(\rightarrow \text{TENSE})$$

d. *aapaṇ*—("long-distance reflexive")[7]

$$((\ \text{GF}^*\ \ \text{GF}\uparrow)\ \text{GF INDEX}) \neq (\uparrow \text{INDEX})$$
$$\neg(\rightarrow \text{PRED})$$
$$((\text{GF}^*\uparrow)\ \gamma(\hat{\theta})\ \text{INDEX}) = (\uparrow \text{INDEX})$$

7.5 On INDEX

We have been following a long tradition in generative syntax in treating referential nominals as having a feature (or feature structure) called an index, which is a representation of its referentiality. This is the approach taken in Government/Binding theory (where the index is an arbitrary label annotated to the DP) and in HPSG (where INDEX is a feature structure which is part of SYNSEM|CONTENT). We have formalized this by treating INDEX as an f-structure attribute.

There is a certain improbability to this approach. Anaphora is about referential relations. Since reference is not part of syntax, it would need to be shown that referentiality is represented syntactically. In a theory based on the concept of parallel corresponding structures, there is no reason to reject a theory in which syntax constrains representations at some other level.

The alternative, then, is that Binding Theory is a theory of syntactic constraints on the representation of reference and coreference, presumably a representation of meaning (semantics) or discourse (pragmatics/information structure). This is the approach taken in the Minimalist Program (Chomsky 1995), for example. In LFG, this could be realized in terms of a projection of f-structure: the referential structure that results from applying some mapping function ρ to f-structure. Using a subscripted ρ to indicate the referential projection, we would replace our anaphoric equations in terms of INDEX with ones expressed in terms of the projection.

(32) a. ((PathOut GF $\uparrow$) PathIn INDEX) = ($\uparrow$ INDEX)
b. ((PathOut GF $\uparrow$) PathIn)$_\rho$ = $\uparrow_\rho$

[7]We call the a-structure to f-structure mapping γ, so "$\gamma(\hat{\theta})$" means "the GF corresponding to $\hat{\theta}$."

This is the approach taken by Dalrymple (1993), who identifies the referential projection with a level of semantic structure, the σ projection.

This alternative approach deals with coreference directly instead of filtering it through a syntactic construct of indices. A thorough analysis along these lines would require a theory of referentiality and a more detailed view of the projection architecture of LFG. Given the scope of this textbook, we are using the simpler account in terms of indices. However, an account in terms of a nonsyntactic level of representation is preferable.

7.6 Anaphora and levels of representation

Anaphora confirms the correctness of the parallel architecture of LFG. The establishment of an anaphoric relationship between two elements involves identifying two elements at f-structure as being identical at referential/ semantic structure on the basis of properties defined independently at f-structure, c-structure, a-structure, and (if we take other uses of reflexives into account) information structure. It is striking to what extent a purely c-structural theory of anaphora has to propose otherwise unjustified constituencies in order to reduce the different notions of rank to c-command; a perusal of arguments for c-structure configurations in GB and MP reveals that many of them rest on the need to establish c-command for anaphora (or other types of binding, such as quantification).

It is important to note that this sensitivity of a construction to independent properties at distinct levels of representation is more than just consistent with LFG. The architecture of LFG predicts that linguistic constructions should have this property. Any construction will, of necessity, have representations at each of the autonomous levels posited by the parallel architecture; it is reasonable to expect that several of these levels will impose its own requirements. The interaction is predicted (and made possible) by the projection architecture.

Additional readings

The basic concepts of the LFG theory of anaphora, such as nucleus, originate in early work on anaphora, summarized in Sells (1985). The formal analysis given here is, as noted in the text, taken from Dalrymple (1993). The idea that the analysis of anaphora involves relating elements at distinct levels goes back at least to Jackendoff (1972), whose theory of anaphora involved c-structure, referential structure, and θ-structure.

The typology of reflexives is discussed in Bresnan (2000)., where nuclear, subjective, and

logophoric reflexives are discussed. Bresnan also has references to studies on anaphora in various languages.

Exercises

1. A PP complement to a verb can fill an argument position in one of three ways.

   - The PP can be an XCOMP argument of the verb. As discussed by Bresnan (2000), this is most commonly the case with semi-idiomatic expressions like *at odds with*, *in love with*, *in awe of*. (For the purpose of this exercise, you can consider these to be single complex lexical items, with transitive lexical forms).

   - The PP can be an OBL_θ argument of the verb. Locative arguments are most frequently of this kind.

   - The object of the preposition can be the argument of the verb (bearing the function OBL_θ OBJ vis-à-vis the verb). When only one preposition is possible in the complement to the verb, it is generally an indication that the preposition is not a meaningful argument-taking item but simply a Case marker.

   Show how LFG's theory of anaphora explains the different behavior of the objects in each of these three types of PPs. They are exemplified below.

   (i) Max_i kept the computer at odds with him/*himself$_i$.
   (ii) Max_i put the computer near him/himself$_i$.
   (iii) Max_i gave a computer to himself/*him$_i$.

2. Characterize the Norwegian anaphors described in this chapter in terms of domain and antecedent conditions and in terms of anaphoric equations.

8

Conclusion

8.1 Summary

In this book we have outlined the components of the theory of Lexical-Functional Grammar, and we have seen that there are reasons to prefer it to derivational and c-structural theories of syntax. We have seen the centrality of grammatical functions to the workings of syntax and the nonnecessity of derivational operations.

The view of syntax that emerges from LFG is at once similar to and different from the conventional transformationalist view. This is because while LFG rejects the formal mechanisms of derivational theories it often manages to capture the basic insights reached by earlier researchers. Most of the analyses we have discussed are of this nature, building on earlier work rather than rejecting it completely.

On the other hand, LFG rejects many of the conclusions that earlier generative researchers have reached concerning the nature of Universal Grammar. Many such conclusions, such as the idea that grammatical functions can be represented structurally, were based on a typologically impoverished base of languages. These models of Universal Grammar are often claimed to be more "explanatory" than models like LFG. However, it is important to realize that one cannot explain phenomena without an adequate description.

Generative linguistics faces an important challenge. By placing the search for linguistic universals (and thus Universal Grammar) at the center of its concerns, it makes a bold claim that all languages share a certain degree of linguistic (and particularly syntactic) structure. Theories of generative syntax must thus be able to analyze any language in the world.

LFG, as a generative theory, faces the same challenge as other generative theories. If anything, the challenge to LFG is greater than that to some other theories, because one of the basic arguments for LFG's parallel structure

architecture including a direct representation of grammatical functions is the inability of c-structural theories to account for nonconfigurational languages like Warlpiri and Wambaya. This is a typological argument. It is therefore important for LFG to show that it is typologically plausible in other realms.

Researchers in LFG have been very aware of this challenge, and have studied many typologically diverse languages. Among the languages that have been studied in LFG are Warlpiri, Wambaya, Navajo, Plains Cree, Greenlandic (Inuit), Welsh, Irish, German, Dutch, West Flemish, Icelandic, Norwegian, Finnish, Russian, Serbo-Croatian, Chicheŵa, Ewe, Moroccan Arabic, Malayalam, Hindi-Urdu, Japanese, and, of course, English (list from Bresnan 2000). Many of the concepts discussed in this book developed as a result of problems raised by a construction in some language for an earlier formulation. At the same time, the formalism of LFG has not changed significantly since it was first introduced in Kaplan and Bresnan (1982), thus proving itself to have been adaptable to a wide variety of languages.

8.2 LFG and other constraint-based theories

LFG shares the distinction of being a constraint-based theory with other frameworks, most notably Head-driven Phrase Structure Grammar and Construction Grammar. Naturally, there are many points of contact between LFG and these other frameworks, which are in some sense variations on the same theme. However, there are also differences between them. In this section, we will briefly discuss some differences between LFG and HPSG.

Perhaps the biggest difference between LFG and HPSG is the LFG notion of parallel representations. In HPSG, all linguistic information about a linguistic element (phonological, structural, functional, semantic, pragmatic) is located in a single AVM called a sign. HPSG thus makes the implicit claim that all linguistic structure is of the same type. LFG, on the other hand, treats different dimensions of linguistic structure as having different formal properties.

Grammatical functions play a major role in LFG, as we have seen. In HPSG, while they are recognized in some sense, they do not play the same central role. For example, long distance dependencies are analyzed by passing information about the filler down through the sign to the phrase with the gap. This passing of information is based on syntactic category; there is no functional component. Argumenthood is based more on category than function, with argument function status derived from the ARG-ST ("argument structure") list. Relative prominence for such principles as Binding

Theory is based purely on the ARG-ST list.[1]

On the other hand, in HPSG much of the descriptive power lies in the notion of a feature system. Feature structures belong to types and subtypes, which identify different constructions and word classes. In LFG, on the other hand, features are limited to the role of one kind of attribute in f-structure,[2] without the rich structure to the system that HPSG posits.

LFG shares with other constraint-based theories a certain basic set of assumptions about linguistic structure, but it is important to keep in mind that the different theories are not notational variants of each other. It is to be hoped that further research will help to pinpoint the strengths and weaknesses of each approach.

8.3 Optimality Theory

Another line of research which is "constraint-based," but in a different way, is Optimality Theory (OT). OT claims that Universal Grammar consists of a set of constraints on structures, and that grammaticality is determined by how well a structure obeys the constraints. Since the constraints can contradict each other, no structure meets all of them. OT posits a language-specific ranking of the constraints; a grammatical structure is one whose highest constraint violation is farther down on this ranking than potential competing structures. OT has proven to be most promising in providing generative grammar with a theory of markedness, traditionally a problem for generative theories. Many researchers in LFG have examined the possibility of a hybrid LFG/ OT, in which the parallel levels of LFG are governed by OT constraints and constraint interaction.

OT's view of constraints is particularly well suited to LFG's parallel architecture. The existence of autonomous levels of representation, each with its own constraints, can easily lead to conflicting constraints. LFG treats certain types of linguistic variation as the result of differential ranking of constraints at different levels. One example of this is the notion of rank in the theory of anaphora: the antecedent of an anaphor must outrank the anaphor, but rank is defined at several levels of representation. Different languages give different importance to each of these levels. OT (and, in particular,

---

[1]I will not take a position on what HPSG's ARG-ST list corresponds to in LFG; it shares certain properties with a-structure and with f-structure.

[2]and, to some extent, in c-structure, such as categorial features.

LFG/OT) thus provides an interesting alternative to the popular notion of parameters in accounting for linguistic variation.

In this textbook, we have not discussed LFG/OT. To do so would require a separate exposition of the principles and formalisms of OT. In addition, the exact nature of the fit between OT and LFG is not entirely clear yet, nor is OT in its current form universally recognized by practitioners of LFG as a useful tool.

8.4 Formalism vs. functionalism

Contemporary approaches to the study of language tend to be classified as being either "formalist" or "functionalist". Formalist approaches treat language as having a formal structure, characterized by rules which relate specifically to language. Syntax is generally taken by formalist approaches to be autonomous, and to have a central place in the linguistic system. Explanation is taken to be based on "internal" formal properties of the syntactic system and syntactic representations. Functionalist approaches, on the other hand, see syntactic patterns as the result of the communicative functions of the various constructions; syntax is either nonautonomous or nonexistent in such approaches. Explanation is thus "external" to the syntax.

This bifurcation, although ubiquitous, is artificial. Detailed analysis of innumerable constructions shows that formal syntactic tools are required for the description of language. On the other hand, these constructions have communicative functions which must also be part of the linguistic description. The properties of linguistic constructions are thus most sensibly a result of both internal and external properties. The (primarily) sociologically induced distinction between the two types of theories leads functionalists to propose forced functional accounts of constituency and grammatical functions and formalists to represent different communicative functions with different c-structure configurations.

LFG provides a way to bridge this gap. LFG, like all varieties of generative grammar, is a formalist theory. As we have seen, LFG syntax has a rich formal language-internal structure, of the kind we would expect from a formal theory. However, the parallel, projection-based architecture calls into question two elements of a typical formalist theory: the autonomy of syntax and the reliance on internal explanation.

It is difficult to characterize syntax (or anything else) in LFG as autonomous or nonautonomous. As we have seen, the architecture of the theory posits the syntactic modules as being governed by module-internal principles, such as $\bar{X}$ theory and Economy of Expression in c-structure and

Completeness, (Extended) Coherence, and Uniqueness in f-structure. On the other hand, the modules are related to each other by projection functions, thus allowing c-structure to (for example) be influenced by f-structure. The autonomy that these modules have is thus a limited autonomy. Similarly, we have made reference to additional potential dimensions of language, such as information structure. Information structure, whatever it looks like, will be related to one or more syntactic level of representation, and rules of syntax can thus potentially be governed by information structure (communicative) properties.

Similarly, LFG allows both internal and external explanation. The existence of correspondence between syntactic structures and information structure opens up the possibility of explaining some syntactic patterns by reference to communicative functions. We have already seen one example of this in our discussion of relative clauses, where the syntactically assigned grammatical function TOPIC maps to a representation with certain informational properties (such as givenness), resulting in its incompatibility with presentational constructions (i.e those presenting elements that are new to the discourse) such as the *there* construction.

LFG thus opens the way for linguistics to grow beyond the formalist/functionalist bifurcation, with a functionally-sensitive formalism. It allows a single theory to express the generalizations which have been discovered by linguists from both sides of the great divide.

8.5 ILFGA and Internet resources

Research in LFG is supported by the International Lexical-Functional Grammar Association (ILFGA). ILFGA sponsors, directly or indirectly, several Internet-based resources in LFG. These are listed below, with the URL of the relevant web site and the name(s) of the person or people currently responsible. Naturally, while the names and the URLs are accurate as of the publication of this textbook, both are subject to change at any time.

- There are two general websites on LFG, with links to other on-line LFG resources:

  `http://www-lfg.stanford.edu/lfg/`

  `http://clwww.essex.ac.uk/LFG/`

  The Essex site is maintained by Douglas Arnold and the Stanford site by Miriam Butt and Tracy Holloway King. Many LFG researchers also maintain their own "unofficial" LFG websites; most notably "Joan Bresnan's Unofficial Links and Notes on LFG/OT", located at

`http://www-lfg.stanford.edu/lfg/bresnan/unofficial-links.html`
There is an LFG FAQ ("frequently asked questions"), maintained by Mary Dalrymple, at:
`http://www-lfg.stanford.edu/lfg/lfg-information.html`
The FAQ can also be ordered by e-mail from
`listserv@listserv.linguistlist.org`
by sending a message reading
`get faq.txt`
It can also be retrieved by anonymous FTP from
`ftp-lfg.stanford.edu`
where it can be found in the directory
`/pub/lfg/lfg-information`

- ILFGA also has an organizational website:
  `http://www-lfg.stanford.edu/lfg/ilfga/`
  The ILFGA bulletin is edited by Miriam Butt. The bulletin is available by e-mail from
  `listserv@listserv.linguistlist.org`
  by sending a message reading
  `get lfg-bulletin.txt`

- ILFGA sponsors an annual international conference, organized by Christopher Manning and Rachel Nordlinger. The proceedings of past conferences are available at the CSLI Publications website:
  `http://csli-publications.stanford.edu/hand/miscpubsonline.html`

  Information about the upcoming conference is available at a site reachable from the Essex LFG website and the ILFGA site.

- There is a web-accessible bibliography of work in LFG, maintained by Avery D. Andrews. It can be found in various formats at:[3]
  `http://www-lfg.stanford.edu/lfg/bibliography.html`
  It can also be ordered by e-mail in plain text format by sending an e-mail to
  `listserv@listserv.linguistlist.org`

[3]The bibliography and certain other documents can also be retrieved by anonymous FTP. The FAQ has detailed information.

with the body of the e-mail:
`get lfg-bib.text`

- Many LFG researchers make their papers available on the Web. While some do this through their personal websites, it can also be done through the LFG Archive, also maintained by Avery D. Andrews.
  `http://www-lfg.stanford.edu/lfg/archive/archfront.html`

- Issues in LFG are discussed electronically through the LFG List, an e-mail discussion list sponsored by ILFGA operated through the Listserv at the LINGUIST List, and maintained by Yehuda N. Falk. Many conference announcements are also distributed via the LFG List. Past posts to the list can be read at the list archive:
  `http://listserv.linguistlist.org/archives/lfg.html`
  One can subscribe to the LFG List on the Web as well.
  `http://linguistlist.org/subscribing/sub-lfg.html`
  Alternatively, one can subscribe by sending an e-mail to
  `listserv@listerv.linguistlist.org`
  with the body of the message reading
  `subscribe lfg firstname lastname`
  where you substitute your first name and last name for `firstname` and `lastname`.

Additional readings

While this book has focused on the syntax itself, LFG researchers have also studied the implications of LFG for other components of the grammar. For recent work on semantics and LFG, see Dalrymple, ed. (1999). On morphology, see, inter alia, Börjars, Vincent, and Chapman (1997), Börjars and Vincent (2000), and Spencer (2000).

LFG/OT has been explored in too many places to list here. Two useful Internet resources, both listed above, are the Stanford LFG site and Joan Bresnan's Unofficial Links site.

LFG has also been the basis of many studies on computation and processing. A recent publication from one such project (the ParGram project) is Butt, King, Niño and Segond (1999). A data-oriented parsing approach to language comprehension that uses LFG representations as its basis (LFG-DOP) is discussed in Bod (1999).

On all topics, the on-line proceedings of the LFG conferences are worth checking out. The conferences are the best way to stay abreast of developments in LFG.

Appendix A

Glossary

We summarize here all the technical terms introduced in the text. The number at the end of each definition is the page on which the term is introduced.

a-structure– argument structure, a representation of the syntactic functions selected by a predicate, most of which fill thematic argument positions. (14)

adjoined– a c-structure configuration in which an X and a Y combine to form a larger X (linear order irrelevant). Adjuncts and some elements with discourse functions are typically adjoined. (35)

anaphor– a syntactic element that is referentially dependent on another element, subject to syntactic constraints. In LFG usage, the term anaphor is more inclusive than in recent derivational usage, and includes personal pronouns. (174)

anaphoric control– a control construction in which the "missing" embedded SUBJ is an f-structure pronoun. (119)

annotated c-structure– a c-structure with functional equations associated with the nodes. (69)

argument functions– the grammatical functions expressing the arguments of predicates. (55)

asymmetrical languages– languages in which, as in English, the two objects of a ditransitive exhibit different syntactic behavior. (94)

attribute-value matrix– a tabular representation of attributes (functions and features) and their values. In LFG, f-structure is conventionally drawn as an attribute-value matrix. (12)

AVM– see **attribute-value matrix**.(12)

c-structure– constituent structure, a representation of the units that make up the overt form of a sentence. The primitives of c-structure are categories, immediate dominance relations, and ordering relations. (10)

category– classes of words and phrases, which determine morphological and c-structural properties. (33)

clause nucleus– another name for **nucleus**. (175)

closed– grammatical functions in which all grammatical dependents are supplied within the function (124)

co-heads– two or more c-structure heads which are joint f-structure heads of a constituent, such as the D and N in a DP, or the C, I, and V in a CP. (38)

coherent– description of an f-structure in which all meaningful elements are integrated into the meaning of the sentence. (59)

complement functions– argument functions which are not also overlay functions (i.e. argument functions other than SUBJ). (58)

complement position– a c-structure position which is a sister of the head. Elements in complement position of a lexical head generally have complement functions, while elements in complement position of a functional category are usually co-heads. (35)

complete– description of an f-structure in which all argument functions selected by the PRED feature of the head are present. (59)

complete (clause) nucleus– a (clause) nucleus containing the attribute SUBJ. (178)

complex predicate– a single f-structure predicate composed of more than one lexical predicate. (95)

configurational– languages in which grammatical functions are encoded in c-structure configurations, with the SUBJ outside of the VP and the OBJ inside. (22)

constituent– a string of elements in c-structure that form a unit. (33)

constraining equation– a functional equation that requires an attribute to have a particular value. (76)

constraint-based– a description of theories of linguistic structure in which grammaticality is a result of fulfilling the requirements of multiple simultaneous constraints rather than being a result of a linear derivation. (9)

constructive morphology– the use of inflectional morphemes, such as Case, to require that a particular constituent bear a particular function. Constructive morphology works through the use of inside-out existential requirements. (79)

control– as used in LFG, any construction in which the SUBJ of a nonfinite clause is omitted and understood as being identical to some element higher in the sentence. This corresponds to both “control” (equi) and

raising constructions in GB/MP terminology. (117)

control equation– a lexical specification, part of the lexical entry of a functional control predicate, which identifies an argument with the XCOMP'S SUBJ. (132)

controllee– the missing embedded SUBJ in a control construction. (117)

controller– the expressed element in the higher clause in a control construction with which the missing embedded SUBJ is identified. (117)

core functions– the more strictly syntactic argument functions (SUBJ, OBJ, OBJ_θ), which are not explicitly tied to thematic roles and are involved in syntactic rules. (56)

correspondence– a relation between elements at different levels of representation in a parallel architecture theory. (23)

defining equation– a functional equation that establishes a particular value for an attribute. (75)

discourse functions– grammatical functions that express relations between participants in a sentence and the larger discourse in which the sentence is embedded. (57)

endocentric– a phrasal constituent with a c-structure head. The property of having a c-structure head is called **endocentricity**. (42)

endocentric organization– a form of organization used in some languages where highly hierarchical $\bar{X}$ c-structures are used and grammatical functions are encoded configurationally. (48)

equi– a control construction in which the controller is a thematic argument of its clause. (In GB/MP, this is called "control".) (118)

exocentric– a phrasal constituent with no c-structure head. The property of lacking a c-structure head is called **exocentricity**. (48)

extraction– another name for a **long-distance dependency** construction. (145)

f-command– a structural relation in f-structure, corresponding to the c-structure relation of c-command. (122)

f-description– a series of simultaneous functional equations which express all the functional relations in a sentence. (66)

f-precedence– a relation between f-structure elements based on precedence at c-structure. (64)

f-structure– functional structure, a representation of the grammatical functions and features expressed by a sentence. (11)

filler– the higher function in a long-distance dependency, usually represented in c-structure by an element in a specifier or adjoined position. (145)

functional categories– in $\bar{X}$ theory, categories of words which contribute

formal features (such as TENSE) to their phrases. In LFG, members of functional categories lack the PRED feature. (37)

functional control– a predicational construction, formalized in LFG as a control construction in which the missing embedded SUBJ is functionally identified with an element of the higher clause. (126)

functional equation– an expression, in the form of an equation, of the functional relation between two nodes in c-structure. (67)

functional uncertainty– the licensing of a dependency through a functional equation with an infinite number of potential solutions. (148)

gap– the lower function in a long-distance dependency construction, usually a missing element (or empty category) in c-structure. (145)

generative grammar– see **generative linguistics** (1)

generative linguistics– an approach to the study of linguistics that has its origins in the pioneering work of Noam Chomsky in the 1950s. Generative linguistics attempts to develop precise mathematical models of linguistic knowledge which correspond to the mental representations of language. LFG is a generative theory. (1)

grammatical function– function, or role, of an element of syntax, such as "subject" or "adjunct". (10)

grammaticized discourse functions– a more specific name for **discourse functions**. (57)

head– the word in a phrase which determines the bulk of the phrase's grammatical properties, especially category. By extension, the word which contributes a phrase's PRED feature can be called its f-structure head. (35)

immediate dominance (ID) rules– rules which licence immediate dominance relations. In LFG, they also associate structural positions with grammatical functions. (46)

incoherent– description of an f-structure in which not all meaningful elements are integrated into the meaning of the sentence. (59)

incomplete– description of an f-structure in which at least one argument function selected by the PRED feature of the head is absent. (59)

inside-out– a method of designating a function in f-structure by defining a path outwards through the f-structure. This is the f-structure equivalent of bottom-up. (79)

Kleene plus– the '$^+$' in X^+. X^+ means 'at least one X, maybe more'. (46)

Kleene star– the '*' in X*. X* means 'any number of X, including possibly zero'. (45)

lexical categories– in $\bar{X}$ theory, categories of words (and the phrases

projected from them) which generally contribute semantic content to the syntactic structures of which they form a part. The lexical categories are generally considered to be V, N, P, A, and ADV. In LFG, most members of lexical categories are lexically marked with the PRED feature. (34)

lexical form– the f-structure representation of a predicate and the argument functions it selects (its a-structure). (14)

lexical integrity– an idea which typifies lexicalist theories, according to which the internal structure of words is invisible to rules of the syntactic component. In LFG, the Lexical Integrity Principle constrains c-structure. (4)

Lexical Mapping Theory (LMT)– the LFG theory of the mapping of arguments of a predicate to the syntax. (96)

lexocentric organization– a form of organization used in some languages where flat c-structures are used and grammatical functions are encoded morphologically. (48)

linear precedence (LP) rules– rules which specify the linear order of sister constituents in c-structure. (46)

long-distance dependency– a construction in which a single element has two functions which can be infinitely far apart. In derivational theories, long-distance dependencies are modeled as *wh* movement. (145)

m-structure– possible additional level of representation in which inflectional features are specified. (86)

metavariables– the symbols ↑ and ↓, used in defining the functional relationships between c-structure positions. The metavariables are general references to c-structural positions; in the description of an actual sentence, they are replaced by variables. (69)

monotonic– refers to the building up of information (such as a linguistic representation) by successively adding information, without changing or destroying anything. (9)

nonargument functions– the grammatical functions that express relations other than argumenthood. (56)

nonconfigurational– languages in which grammatical functions are not encoded in c-structure configurations. Such languages typically do not have a VP constituent. (22)

noncore functions– the argument functions (primarily the OBL_θ family) that generally are marked overtly with their thematic role. (56)

nonterm functions– another name for **noncore functions**. (56)

nonnuclear– anaphors that may not be bound in the minimal nucleus (or minimal complete nucleus) containing them. (177)

nuclear– anaphors that must be bound in the minimal nucleus (or minimal complete nucleus) containing them. (177)

nucleus– those elements of an f-structure comprising an argument-taking PRED and all the argument functions it selects. (175)

oblique– an argument whose role (usually thematic role) is explicitly marked, either by a preposition or by a semantic Case. In LFG, the oblique functions are treated as a class of grammatical functions OBL_θ. (14)

off-path constraints– constraints that limit functions on a path by disallowing them from containing certain attributes. (183)

open– a predicative function, one from which the SUBJ is absent and functionally identified with an element of the higher clause (functional control). The open function is called XCOMP in LFG. (124)

overlay functions– grammatical functions which serve clause-external roles, such as linking participants to other clauses in the sentence (SUBJ) or to other parts of the discourse (discourse functions). (57)

outside-in– a method of designating a function in f-structure by defining a path inwards through the f-structure. This is the f-structure equivalent of top-down. (79)

parallel– a description of the architecture of a theory of language in which different dimensions of linguistic information are represented at different levels of representation with different primitives and principles. (23)

ϕ– the mapping from c-structure to f-structure. (64)

phrase– a c-structure grouping of elements (constituent) which is typified by its ability to appear in different positions in a sentence with the same internal structure. (35)

phrase structure rules– a traditional formal device for licensing c-structures. Although much work in LFG uses phrase structure rules, in this book we factor them into ID rules and LP rules. (44)

projection– a c-structure entity consisting of a word (the head) and elements that relate to it (arguments, adjuncts, specifier). A projection can be a phrase or an intermediate "X'" projection. (35)

projection architecture– the property of LFG under which the various representations are related by correspondence. The representations so related are called "projections"; this is a different use of the term from the $\bar{\text{X}}$ theoretic use above. (24)

raising– a control construction in which the controller is not a thematic argument of its clause. (118)

rank– the relative prominence of an element of a clause. Prominence at

f-structure is defined by position on the relational hierarchy, prominence at a-structure is defined by the thematic hierarchy, and prominence at c-structure is defined by linear order. (173)

relational hierarchy– the relative prominence of argument functions, as shown by their respective accessibility to syntactic manipulation. (56)

restricted– an argument function which is specified as to thematic role, either by being a noncore function or (in the case of OBJ_θ) by the mapping rules of the language. (103)

specifier position– a structural position, the daughter of a phrasal node and sister of an intermediate "X′". (39)

suppression– the nonmapping to the syntax of an argument in the a-structure. (107)

symmetrical languages– languages in which, unlike in English, the two objects of a ditransitive behave the same syntactically. (94)

term functions– another name for **core functions**. (56)

thematic hierarchy– relative prominence of arguments based on their thematic roles. (100)

thematic (θ) roles– informal description of conceptual/semantic participant roles, common in generative studies. (97)

theta-hat ($\hat{\theta}$)– the thematically most prominent argument of a predicate, traditionally called "logical subject". (103)

unbounded dependency– another name for a **long-distance dependency** construction. (145)

unification– the merging of features from different sources which relate to the same linguistic entity. Unification is a formalism which is an alternative to movement as a way of accounting for the "displacement" of linguistic elements. (18)

variable– in the c-structure–f-structure mapping, a temporary name which can be used to refer to a corresponding c-structure node and f-structure. (65)

$\bar{X}$ (X-bar) theory– A restrictive theory of endocentric c-structure categories, according to which phrasal categories are projected from lexical categories. (34)

Appendix B

A Minigrammar of English

Throughout this textbook we have developed parts of a grammar of English. In this appendix we collect all the rules and lexical entries we have developed.

ID rules

Functional maximal projections

$$\text{CP} \rightarrow \quad \text{XP} \quad , \quad \text{C}' \qquad \left\{ \begin{array}{c} (\uparrow \text{DF}) = \downarrow \\ (\downarrow \text{PRON}) =_{c} \text{WH} \\ \\ (\uparrow \text{SUBJ}) = \downarrow \\ \neg\,(\uparrow \text{TENSE}) \\ (\downarrow \text{PCASE}) =_{c} \text{OBL}_{\text{Ben}} \end{array} \right\} \uparrow = \downarrow$$

$$\text{IP} \rightarrow \quad \left\{ \begin{array}{c} \text{DP} \\ \text{NP} \\ \text{CP} \\ \text{PP} \end{array} \right\} \quad , \quad \text{I}' \qquad (\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow \qquad \left((\uparrow \text{DF}) = \downarrow \right)$$

$$\begin{array}{lccc} \text{DP} \rightarrow & \text{DP} & , & \text{D}' \\ & (\uparrow \text{POSS}) = \downarrow & & \uparrow = \downarrow \\ & (\uparrow \text{DEF}) = + & & \\ & (\downarrow \text{CASE}) =_{c} \text{GEN} & & \end{array}$$

Functional single-bar projection

$$\begin{array}{lccc} \text{C}' \rightarrow & \text{C} & , & \left\{ \begin{array}{c} \text{S} \\ \text{IP} \\ \text{VP} \end{array} \right\} \\ & \uparrow = \downarrow & & \uparrow = \downarrow \end{array}$$

$$\begin{array}{lcc} \text{I}' \rightarrow & \text{I} \quad , & \text{VP} \\ & \uparrow = \downarrow & \uparrow = \downarrow \end{array}$$

$$\begin{array}{lcc} \text{D}' \rightarrow & \text{D} \quad , & \text{NP} \\ & \uparrow = \downarrow & \uparrow = \downarrow \end{array}$$

Lexical phrases

$$\begin{array}{lccccc} \text{NP} \rightarrow & \text{N} & , & \text{PP}^{*} & , & \text{CP} \\ & \uparrow = \downarrow & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

$$\begin{array}{lccccccc} \text{VP} \rightarrow & & & & & & & \\ & \text{V} & , & \left\{ \begin{array}{c} \text{DP} \\ \text{NP} \end{array} \right\}^{*} & , & \text{XP} \quad , & \text{PP}^{*} \quad , & \left\{ \begin{array}{c} \text{CP} \\ \text{IP} \\ \text{S} \end{array} \right\} \\ & \uparrow = \downarrow & & (\uparrow [+o]) = \downarrow & & (\uparrow \text{XCOMP}) = \downarrow & (\uparrow (\downarrow \text{PCASE})) = \downarrow & (\uparrow \text{COMP}) = \downarrow \\ & & & \neg (\downarrow \text{TENSE}) & & & & \end{array}$$

$$\begin{array}{lccccc} \text{AP} \rightarrow & \text{A} & , & \text{PP*} & , & \left\{\begin{array}{c}\text{CP}\\ \text{IP}\\ \text{S}\end{array}\right\} \\ & \uparrow = \downarrow & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

$$\begin{array}{lccccccc} \text{PP} \rightarrow & \text{P} & , & \left\{\begin{array}{c}\text{DP}\\ \text{NP}\end{array}\right\} & , & \text{PP} & , & \left\{\begin{array}{c}\text{IP}\\ \text{S}\end{array}\right\} \\ & \uparrow = \downarrow & & (\uparrow \text{OBJ}) = \downarrow & & (\uparrow (\downarrow \text{PCASE})) = \downarrow & & (\uparrow \text{COMP}) = \downarrow \end{array}$$

Adjunction ID rules

$$\begin{array}{l} \text{XP} \rightarrow \text{XP} \ , \ \left\{\begin{array}{c}\text{PP}\\ \left\{\begin{array}{c}\text{AP}\\ \text{ADVP}\end{array}\right\}\end{array}\right\}, \text{XP a lexical category} \\ \uparrow = \downarrow \quad \downarrow \in (\uparrow \text{ADJ}) \end{array}$$

$$\begin{array}{lccc} \text{NP} \rightarrow & \text{NP} & , & \left\{\begin{array}{c}\text{CP}\\ \text{IP}\\ \text{S}\end{array}\right\} \\ & \uparrow = \downarrow & & \downarrow \in (\uparrow \text{ADJ}) \\ & & & \left((\downarrow \text{OPER PRED}) = \text{`PRO'}\right) \end{array}$$

$$\begin{array}{lccc} \text{IP} \rightarrow & \text{XP} & , & \text{IP} \\ & (\uparrow \text{DF}) = \downarrow & & \uparrow = \downarrow \\ & (\downarrow \text{GF}^{*} \text{PRON}) \neq \text{WH} & & \end{array}$$

Conjunction rule

$$\begin{array}{lccc} \text{XP} \rightarrow & \text{XP}^{+} & \text{CONJ} & \text{XP} \\ & \downarrow \in \uparrow & & \downarrow \in \uparrow \end{array}$$

Empty category rule

$$XP \rightarrow \quad e$$
$$\uparrow = ((\{\text{COMP} \mid \text{XCOMP} \mid \text{OBL}_\theta\}^* \text{ GF } \uparrow) \text{ DF})$$

LP rules

X^0 initial
DP $\prec$ PP
PP ($\prec$ {CP | IP})final
SPEC initial
{AP | ADVP} $\prec$ {NP | AP | PP}
$\text{OBJ} \prec_f \text{OBJ}_\theta$

Lexical Mapping Theory

θ-structure to a-structure mapping

nonpropositional Patients and Themes map to [$-r$]
"secondary" nonpropositional Patients and Themes (in English, non-Patient Themes) map to [$+o$] as a marked option
propositional arguments map to [$+c$].
non-Theme/Patient arguments map to [$-o$]

a-structure to f-structure mapping

SUBJ Mapping 1: A [$-o$] argument which is $\hat{\theta}$ maps to SUBJ
SUBJ Mapping 2: [$-r$] may map to SUBJ.
NonSUBJ Mapping: Add positive values of features where possible.

Well-Formedness Conditions

Function-argument biuniqueness Each a-structure role corresponds to a unique f-structure function, and each f-structure function corresponds to a unique a-structure role.

The Subject Condition Every verb must have a SUBJ.

Operations on a-structure

Resultative Formation

$\langle \ldots \rangle \Rightarrow \langle \ldots, \text{XCOMP} \rangle \quad (\uparrow [-r]) = (\uparrow \text{XCOMP SUBJ})$
(XCOMP has resultative semantics)

Passivization

Do not map $\hat{\theta}$ to the syntax. (Often written: $\begin{vmatrix} \hat{\theta} \\ \varnothing \end{vmatrix}$)

Extraposition

$\langle \ldots [+c] \ldots \rangle \Rightarrow [-o] \langle \ldots [+c] \ldots \rangle$
$(\uparrow [-o] \text{ FORM}) = \text{it}$

Raising-to-Subject

$\langle [+c] \rangle \Rightarrow [-o] \langle \text{XCOMP} \rangle$

Empty pronoun rule (English-specific version)

Add the optional equation (↑ SUBJ PRED) = 'PRO' to the lexical entry of a verb without the feature TENSE.

Functional Control Rule

If (↑ XCOMP) is present in a lexical form, add the equation:
(↑ CF) = (↑ XCOMP SUBJ)

Long distance dependency rule
Add the optional equation

$$(\uparrow \text{DF}) \neq (\uparrow \text{SUBJ}) \Rightarrow (\uparrow \text{DF}) = (\uparrow \left\{ \begin{array}{c} \text{COMP} \\ \text{XCOMP} \\ \text{OBL}_{\theta} \end{array} \right\}^{+} \text{SUBJ})$$

to the lexical entry of a verb.

Lexical entries

a	D	(↑ DEF) = –
		(↑ NUM) = SG
agree	V	(↑ PRED) = 'agree ⟨(↑ SUBJ) (↑ COMP)⟩'
believe	V	(↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ OBJ)⟩'
believe	V	(↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ COMP)⟩'
believe	V	(↑ PRED) = 'believe ⟨(↑ SUBJ) (↑ XCOMP)⟩ (↑ OBJ)'
		(↑ OBJ) = (↑ XCOMP SUBJ)
did	I/C	(↑ TENSE) = PAST
		C ⇒ (↑ TYPE) = Q
dinosaur	N	(↑ PRED) = 'dinosaur'
		(↑ NUM) = SG
gorilla's	N	(↑ PRED) = 'gorilla'
		(↑ NUM) = SG
		(↑ CASE) = GEN
		(POSS ↑)
green	A	(↑ PRED) = 'green'

hamsters	N	(↑ PRED) = 'hamster' (↑ NUM) = PL
has(aux)	C/I	(↑ ASP) = PERF (↑ TENSE) = PRES C ⇒ (↑ TYPE) = Q (↑ PART) $=_c$ PAST (perhaps stated at m-structure)
have(aux)	V	(↑ ASP) = PERF (↑ PART) $=_c$ PAST (perhaps stated at m-structure)
help	V	(↑ PRED) = 'help ⟨(↑ SUBJ) (↑ OBJ) (↑ XCOMP)⟩' (↑ OBJ) = (↑ XCOMP SUBJ) ¬(↑ XCOMP PART) [– N] ∈ λ (ϕ^{-1} (↑ XCOMP)) (note: [– N] = V or P)
him	N	(↑ PRED) = 'PRO' (↑ NUM) = SG (↑ GEND) = MASC ((GF* GF ↑) GF INDEX) ≠ (↑ INDEX) ¬ (→ PRED)
himself	N	(↑ PRED) = 'PRO' (↑ NUM) = SG (↑ GEND) = MASC ((GF* GF ↑) GF INDEX) = (↑ INDEX) ¬ (→ SUBJ)
house	N	(↑ PRED) = 'house' (↑ NUM) = SG
in	P	(↑ PCASE) = OBL_{Loc}
keep	V	(↑ PRED) = 'keep ⟨(↑ SUBJ) (↑ OBJ) (↑ XCOMP)⟩' (↑ OBJ) = (↑ XCOMP SUBJ) (↑ XCOMP PART) $=_c$ PRES

keep	V	(↑ PRED) = ‘keep-tabs-on ⟨(↑ SUBJ) (↑ OBL_{on} OBJ)⟩ (↑ OBJ)’ (↑ OBJ FORM) $=_c$ ‘tabs’
might	I/C	(↑ TENSE) = MIGHT C ⇒ (↑ TYPE) = Q
put	V	(↑ PRED) = ‘put ⟨(↑ SUBJ) (↑ OBJ) (↑ OBL_{Loc})⟩
seem	V	(↑ PRED) = ‘seem ⟨(↑ COMP)⟩ (↑ SUBJ)’ (↑ SUBJ FORM) $=_c$ it
seem	V	(↑ PRED) = ‘seem ⟨(↑ XCOMP)⟩ (↑ SUBJ)’ (↑ SUBJ) = (↑ XCOMP SUBJ) N ∉ λ (ϕ^{-1} (↑ XCOMP))
sell	V	(↑ PRED) = ‘sell ⟨(↑ SUBJ) (↑ OBJ2) (↑ OBJ)⟩’
sell	V	(↑ PRED) = ‘sell ⟨(↑ SUBJ) (↑ OBJ) (↑ OBL_{Goal} OBJ)⟩’
summer	N	(↑ PRED) = ‘summer’ (↑ NUM) = SG
tabs	N	(↑ FORM) = ‘tabs’ (↑ NUM) = PL
that	C	(↑ TENSE) (↑ SUBJ) ≠ ((GF$^+$ ↑) GF) ((↑ OPER) = (↑ SUBJ)) (↑ TYPE) = DECL | REL
the	D	(↑ DEF) = +
to	P	(↑ PCASE) = OBL_{Goal}
to	C	¬(↑ TENSE) ((↑ OPER) = (↑ TOPIC)) ((↑ TOPIC) = (↑ SUBJ)) (↑ TYPE) = DECL | REL | Q

try	V	(↑ PRED) = ‘try ⟨(↑ SUBJ) (↑ XCOMP)⟩’ (↑ SUBJ) = (↑ XCOMP SUBJ) C ∈ λ (ϕ^{-1} (↑ XCOMP))
what	D	(↑ PRED) = ‘PRO’ ((DF ↑) TYPE) = Q (↑ PRON) = WH
which	D	(↑ PRON) = WH ((DF ↑) TYPE) = REL ⇒ (↑ PRED) = ‘PRO’
will	I/C	(↑ TENSE) = FUT C ⇒ (↑ TYPE) = Q

References

Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect.* Doctoral dissertation, Massachusetts Institute of Technology.

Ackerman, Farrell, and Gert Webelhuth. 1998. *A Theory of Predicates.* Stanford, Calif.: CSLI Publications.

Aissen, Judith L. 1979. *The Syntax of Causative Constructions.* New York: Garland. Originally 1974 doctoral dissertation, Harvard University.

Akmajian, Adrian, and Frank Heny. 1975. *An Introduction to the Principles of Transformational Syntax.* Cambridge, Mass.: MIT Press.

Alsina, Alex. 1992. On the Argument Structure of Causatives. *Linguistic Inquiry* 23: 517–555.

Alsina, Alex. 1994. Bantu Multiple Objects: Analyses and Fallacies. *Linguistic Analysis* 24: 153–174.

Alsina, Alex. 1996a. *The Role of Argument Structure in Grammar: Evidence from Romance.* Stanford, Calif.: CSLI Publications.

Alsina, Alex. 1996b. Passive Types and the Theory of Object Asymmetries. *Natural Language and Linguistic Theory* 14: 673–723.

Alsina, Alex, Joan Bresnan, and Peter Sells (eds.). 1997. *Complex Predicates.* Stanford: CSLI Publications.

Alsina, Alex, and Smita Joshi. 1991. Parameters in Causative Constructions. *CLS* 27: 1–15.

Alsina, Alex, and Sam A. Mchombo. 1990. The Syntax of Applicatives in Chichewa: Problems for a Theta Theoretic Asymmetry. *Natural Language And Linguistic Theory* 8: 493–506.

Alsina, Alex, and Sam A. Mchombo. 1993. Object Asymmetries and the Chichewa Applicative Construction. *Theoretical Aspects Of Bantu Grammar*, ed. Sam A. Mchombo, 17–45. Stanford, Calif.: CSLI Publications.

Andrews, Avery D. 1982. The Representation of Case in Modern Icelandic. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 427–503. Cambridge, Mass.: MIT Press.

Andrews, Avery. 1985. The Major Functions of the Noun Phrase. *Language Typology and Syntactic Description*, Vol. 1: *Clause Structure*, ed. Timothy Shopen,

62–154. Cambridge: Cambridge University Press.

Andrews, Avery D. 1990. Case Structures and Control in Modern Icelandic. *Modern Icelandic Syntax,* ed. Joan Maling and Annie Zaenen, 187–234. New York: Academic Press.

Andrews, Avery D. 1994. *Syntax Textbook Draft 4.0.* ms., Australia National University, Canberra

Andrews, Avery D. 1996. Semantic Case-stacking and Inside-out Unification. *Australian Journal of Linguistics* 16: 1–55.

Andrews, Avery D., and Christopher D. Manning. 1999. *Complex Predicates and Information Spreading in LFG.* Stanford, Calif.: CSLI Publications.

Arka, I Wayan, and Jane Simpson. 1998. Control and Complex Arguments in Balinese. *Proceedings of the LFG98 Conference, University of Queensland, Brisbane*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.

http://csli-publications.stanford.edu/LFG/3/lfg98.html

Austin, Peter, and Joan Bresnan. 1996. Non-configurationality in Australian Aboriginal Languages. *Natural Language and Linguistic Theory* 14: 215–268.

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing.* Chicago: University of Chicago Press.

Barron, Julia. 1998. *Have* Contraction: Explaining 'Trace Effects' in a Theory Without Movement. *Linguistics* 36: 223–251.

Barron, Julia. 2000. Tense and Agreement Morphology—The Irrelevance of Finiteness. *Proceedings of the LFG00 Conference, University of California, Berkeley*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.

http://csli-publications.stanford.edu/LFG/5/lfg00.html

Berman, Judith. 1997. Empty Categories in LFG. *Proceedings of the LFG97 Conference, University of California, San Diego*, ed. Miriam Butt and Tracy Holloway King on-line: CSLI Publications.

http://csli-publications.stanford.edu/LFG/2/lfg97.html

Bittner, Maria, and Ken Hale. 1996. The Structural Determination of Case and Agreement. *Linguistic Inquiry* 27: 1–68.

Bod, Rens. 1999. *Beyond Grammar: An Experience-based Theory of Language.* Stanford, Calif.: CSLI Publications.

Börjars, Kersti, and Nigel Vincent. 2000. Multiple Case and the 'Wimpiness' of Morphology. *Argument Realization,* ed. Miriam Butt and Tracy Holloway King 15–40. Stanford, Calif.: CSLI Publications.

Börjars, Kersti, Nigel Vincent, and Carol Chapman. 1997. Paradigms, Periphrases and Pronominal Inflection: A Feature-Based Account. *The Yearbook of*

Morphology, ed. Geert Booij and Jaap van Marle, 1–26. Dordrecht: Kluwer.

Börjars, Kersti, Erika Chisarik, and John Payne. 1999. On the Justification for Functional Categories in LFG. *Proceedings of the LFG99 Conference, University of Manchester*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
http://csli-publications.stanford.edu/LFG/4/lfg99.html

Brame, Michael K. 1976. *Conjectures and Refutations in Syntax and Semantics*. New York: Elsevier North-Holland.

Brame, Michael. 1982. The Head-Selector Theory of Lexical Specifications and the Nonexistence of Coarse Categories. *Linguistic Analysis* 10: 321–325.

Bresnan, Joan. 1978. A Realistic Transformational Grammar. *Linguistic Theory and Psychological Reality*, ed. Morris Halle, Joan Bresnan, and George Miller, 1–59. Cambridge, Mass.: MIT Press.

Bresnan, Joan. 1982a. The Passive in Lexical Theory. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 3–86. Cambridge, Mass.: MIT Press.

Bresnan, Joan. 1982b. Control and Complementation. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 282–390. Cambridge, Mass.: MIT Press.

Bresnan, Joan (ed.). 1982. *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: MIT Press.

Bresnan, Joan. 1990. Monotonicity and the Theory of Relation-Changes in LFG. *Language Research* 26: 637–652.

Bresnan, Joan. 1994. Locative Inversion and the Architecture of Universal Grammar. *Language* 70: 72–131.

Bresnan, Joan. 1995a. Linear Order, Syntactic Rank, and Empty Categories: On Weak Crossover. *Formal Issues in Lexical-Functional Grammar*, ed. Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, 241–274. Stanford: CSLI Publications.

Bresnan, Joan. 1995b. Lexicality and Argument Structure. Paper presented at the Paris Syntax and Semantics Conference, October 12–14, 1995.
http://csli-www.stanford.edu/users/bresnan/paris.ps

Bresnan, Joan. 1997. Mixed Categories as Head-sharing Constructions. *Proceedings of the LFG97 Conference, University of California, San Diego*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
http://csli-publications.stanford.edu/LFG/2/lfg97.html

Bresnan, Joan. 1998. Morphology Competes With Syntax: Explaining Typological Variation in Weak Crossover Effects. *Is The Best Good Enough?*, ed. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 59–92. Cambridge, Mass.: MIT Press.

Bresnan, Joan. 2000. *Lexical-Functional Syntax.* Oxford: Blackwell.

Bresnan, Joan, and Jonni M. Kanerva. 1989. Locative Inversion in Chicheŵa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20: 1–50.

Bresnan, Joan, and Ronald Kaplan. 1982. Introduction: Grammars as Mental Representations of Language . *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, xvii–lii. Cambridge, Mass.: MIT Press.

Bresnan, Joan, and Sam Mchombo. 1987. Topic, Pronoun, and Agreement in Chicheŵa. *Language* 63: 741–782.

Bresnan, Joan, and Samuel A. Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language And Linguistic Theory* 13: 181–254.

Bresnan, Joan, and Lioba Moshi. 1990. Object Asymmetries in Comparative Bantu Syntax. *Linguistic Inquiry* 21: 147–185.

Bresnan, Joan, and Annie Zaenen. 1990. Deep Unaccusativity in LFG. *Grammatical Relations: A Cross-Theoretical Perspective* , ed. Katarzyna Dziwirek, Patrick Farrell, and Errapel Mejías-Bikandi, 45–57. Stanford: CSLI Publications.

Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach.* Dordrecht: D. Reidel.

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu.* Stanford, Calif.: CSLI Publications.

Butt, Miriam, and Tracy Holloway King. 1999. The Status of Case. University of Konstanz and Xerox PARC

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño, and Frédérique Segond. 1999. *A Grammar-Writer's Cookbook.* Stanford, Calif.: CSLI Publications.

Chomsky, Noam. 1955. *The Logical Structure of Linguistic Theory.* New York: Plenum. 1975.

Chomsky, Noam. 1957. *Syntactic Structures.* The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1970. Remarks on Nominalization. *Readings in English Transformational Grammar*, ed. Roderick Jacobs and Peter Rosenbaum, 184–221. Waltham, Mass.: Ginn and Co.

Chomsky, Noam. 1973. Conditions on Transformations. *A Festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 232–286. New York: Holt Rinehart and Winston.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, Noam. 1986. *Barriers.* Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, Mass.: MIT Press.

Dalrymple, Mary. 1993. *The Syntax of Anaphoric Binding.* Stanford: CSLI Publications.

Dalrymple, Mary (ed.). 1999. *Semantics and Syntax in Lexical-Functional Grammar: A Resource Logic Approach.* Cambridge, Mass.: MIT Press.

Dalrymple, Mary, and Ronald M. Kaplan. 1997. A Set-Based Approach to Feature Resolution. *Proceedings of the LFG97 Conference, University of California, San Diego*, ed. Miriam Butt and Tracy Holloway King on-line: CSLI Publications.
http://csli-publications.stanford.edu/LFG/2/lfg97.html

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen (eds.). 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Dalrymple, Mary, and Helge Lødrup. 2000. The Grammatical Functions of Complement Clauses. *Proceedings of the LFG00 Conference, University of California, Berkeley*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
http://csli-publications.stanford.edu/LFG/5/lfg00.html

Di Sciullo, Anna Maria, and Edwin Williams. 1987. *On the Definition of Word.* Cambridge, Mass.: MIT Press.

Dyvik, Helge. 1999. The Universality of F-structure: Discovery or Stipulation? The Case of Modals. *Proceedings of the LFG99 Conference, University of Manchester*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
http://csli-publications.stanford.edu/LFG/4/lfg99.html

Falk, Yehuda N. 1983a. Constituency, Word Order, and Phrase Structure Rules. *Linguistic Analysis* 11: 331–360.

Falk, Yehuda N. 1983b. Subjects and Long Distance Dependencies. *Linguistic Analysis* 12: 245–270.

Falk, Yehuda N. 1984. The English Auxiliary System: A Lexical-Functional Analysis. *Language* 60: 483–509.

Falk, Yehuda N. 1992. Suppress α. *Linguistics* 30: 999–1030.

Falk, Yehuda N. 1997. Case Typology and Case Theory. Presented at the twelfth annual meeting of the Israel Association for Theoretical Linguistics, Bar-Ilan University, 2–3 June 1997.
http://pluto.mscc.huji.ac.il/~msyfalk/

Falk, Yehuda N. 1999. Philippine Subjects in a Monostratal Framework. Sixth annual conference of the Austronesian Formal Linguistics Association, 16–18 April 1999, University of Toronto.

Falk, Yehuda N. 2000. Pivots and the Theory of Grammatical Functions. *Proceedings of the LFG00 Conference, University of California, Berkeley*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
http://csli-publications.stanford.edu/LFG/5/lfg00.html

Falk, Yehuda N. in preparation. Infinitival *to*. The Hebrew University of Jerusalem.

Fassi-Fehri, Abdelkader. 1981. *Complémentation et anaphore en arabe moderne: Une apprôche lexicale fonctionelle.* Doctoral dissertation, Université de Paris III.

Fassi-Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation.* Rabat, Morocco: Publications de la Faculté des Lettres et des Sciences Humaines de Rabat.

Fillmore, Charles J. 1968. The Case for Case. *Universals in Linguistic Theory*, ed. Emmon Bach and Robert T. Harms, 1–88. New York: Holt Rinehart and Winston.

Frank, Anette. 1996. A Note on Complex Predicate Formation: Evidence from Auxiliary Selection, Reflexivization, and Past Participle Agreement in French and Italian. *Proceedings of the First LFG Conference, Rank Xerox, Grenoble, August 26–28, 1996*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
`http://csli-publications.stanford.edu/LFG/1/lfg1.html`

Frank, Anette. 2000. Syntax and Morphology of Tense in LFG. *Proceedings of the LFG00 Conference, University of California, Berkeley*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
`http://csli-publications.stanford.edu/LFG/5/lfg00.html`

Gazdar, Gerald. 1981. Unbounded Dependencies and Coordinate Structure. *Linguistic Inquiry* 12: 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag. 1984. *Generalized Phrase Structure Grammar.* Oxford: Basil Blackwell.

Givón, T. 1990. *Syntax: A Functional-Typological Introduction, Volume 2.* Philadelphia: John Benjamins.

Greenberg, Joseph H. 1963. Some Universals of Language with Particular Reference to the Order of Meaningful Elements. *Universals of Language*, ed. Joseph H. Greenberg. Cambridge, Mass.: MIT Press.

Grimshaw, Jane. 1982a. Subcategorizaton and Grammatical Relations. *Subjects and Other Subjects: Proceedings of the Harvard Conference on the Representation of Grammatical Relations*, ed. Annie Zaenen, 35–55. Bloomington, Ind.: IULC.

Grimshaw, Jane. 1982b. On the Lexical Representation of Romance Reflexive Clitics. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 87–148. Cambridge, Mass.: MIT Press.

Grimshaw, Jane. 1995. Projection, Heads, And Optimality. ms.,Rutgers University.

Hale, Kenneth, Laverne Jeanne, and Paul Platero. 1977. Three Cases of Overgeneration. *Formal Syntax*, ed. Peter W. Culicover, Thomas Wasow, and Adrian Akmajian. New York. Academic Press

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Jackendoff, Ray. 1977. *$\bar{X}$ Syntax: A Study of Phrase Structure.* Cambridge, Mass.: MIT Press.

Jackendoff, Ray. 1990. *Semantic Structures.* Cambridge, Mass.: MIT Press.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.

Jespersen, Otto. 1940. *A Modern English Grammar on Historical Principles. Part V: Syntax, Fourth Part* . London: George Allen and Unwin.

Johnson, David E., and Paul M. Postal. 1980. *Arc Pair Grammar*. Princeton, N.J.: Princeton University Press.

Kaplan, Ronald M. 1972. Augmented Transition Networks as Psychological Models of Sentence Comprehension. *Artificial Intelligence* 3: 77–100.

Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. *Linguistic Theory and Computer Applications*, ed. P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson, and P. Bennett, 149–181. London: Academic Press. reprinted in Dalrymple, Kaplan, Maxwell, and Zaenen, eds. 1995, 339–367.

Kaplan, Ronald M., and Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 173–281. Cambridge, Mass.: MIT Press.

Kaplan, Ronald M., and John T. Maxwell III. 1988a. An Algorithm for Functional Uncertainty. *Proceedings of COLING-88, Vol. 1.* 297–302. reprinted in Dalrymple, Kaplan, Maxwell, and Zaenen, eds., 1995, 177–197..

Kaplan, Ronald M., and John T. Maxwell III. 1988b. Constituent Coordination in Lexical-Functional Grammar. *Proceedings of COLING-88, Vol. 1.* 303–305. reprinted in Dalrymple, Kaplan, Maxwell, and Zaenen, 1995, pp. 199–210..

Kaplan, Ronald M., and Annie Zaenen. 1989. Long-Distance Dependencies, Constituent Structure, and Functional Uncertainty. *Alternative Conceptions of Phrase Structure*, ed. Mark R. Baltin and Anthony S. Kroch, 17–42. Chicago: University of Chicago Press.

Keenan, Edward L., and Bernard Comrie. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8: 63–99.

King, Tracy Holloway. 1995. *Configuring Topic and Focus in Russian.* Stanford, Calif.: CSLI Publications.

Koster, Jan. 1978. Why Subject Sentences Don't Exist. *Recent Transformational Studies In European Languages*, ed. S. Jay Keyser. Cambridge, Mass.: MIT Press.

Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog.*

Stanford, Calif.: CSLI Publications.

Laczkó, Tibor. 2000. Derived Nominals, Possessors, and Lexical Mapping Theory In Hungarian DPs. *Argument Realization*, ed. Miriam Butt and Tracy Holloway King, 189–227. Stanford, Calif.: CSLI Publications.

Levin, Lori. 1988. *Operations on Lexical Forms: Unaccusative Rules in Germanic Languages.* New York: Garland. Originally 1986 MIT dissertation.

Lødrup, Helge. 2000. Underspecification in Lexical Mapping Theory: The Case of Norwegian Existentials and Resultatives. *Argument Realization*, ed. Miriam Butt and Tracy Holloway King, 171–188. Stanford, Calif.: CSLI Publications.

Manning, Christopher D. 1996. *Ergativity: Argument Structure and Grammatical Relations.* Stanford, Calif.: CSLI Publications.

Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *Proceedings of the 21st Annual Penn Linguistics Colloquium, U Penn Working Papers in Linguistics* 4.2: 201–225. Philadelphia, Pa.: U Penn Linguistics Department.

Matsumoto, Yo. 1998. On the Cross-Linguistic Parameterization of Causative Predicates: Japanese Perspectives. *Proceedings of the LFG98 Conference, University of Queensland, Brisbane*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
`http://csli-publications/stanford.edu/LFG/3/lfg98.html`

Mohanan, K.P. 1982. Grammatical Relations and Clause Structure in Malayalam. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 504–589. Cambridge, Mass.: MIT Press.

Mohanan, K.P. 1983. Functional and Anaphoric Control. *Linguistic Inquiry* 14: 641–674.

Neidle, Carol. 1982. Case Agreement in Russian. *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 391–426. Cambridge, Mass.: MIT Press.

Nordlinger, Rachel. 1998. *Constructive Case: Evidence from Australian Languages.* Stanford, Calif.: CSLI Publications.

Nordlinger, Rachel. 2000. Australian Case Systems: Towards a Constructive Solution. *Argument Realization*, ed. Miriam Butt and Tracy Holloway King, 41–71. Stanford, Calif.: CSLI Publications.

Perlmutter, David M. 1978. Impersonal Passives and the Unaccusative Hypothesis. *BLS* 4.

Perlmutter, David M. (ed.). 1983. *Studies in Relational Grammar 1.* Chicago: University of Chicago Press.

Pinker, Steven. 1996. *Language Learnability and Language Development*, Second Edition. Cambridge, Mass.: Harvard University Press.

Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Stanford, Calif.: CSLI Publications.

Postal, Paul M. 1970. On Coreferential Complement Subject Deletion. *Linguistic Inquiry* 1: 439–500.

Postal, Paul M. 1974. *On Raising: One Rule of English Grammar and Its Theoretical Implications*. Cambridge, Mass.: MIT Press.

Postal, Paul M., and Geoffrey K. Pullum. 1988. Expletive Noun Phrases in Subcategorized Positions. *Linguistic Inquiry* 19: 635–670.

Rappaport-Hovav, Malka, and Beth Levin. 2000. An Event Structure Analysis of English Resultatives. ms., The Hebrew University of Jerusalem and Stanford University.

Sadler, Louisa. 1999. Non-distributive Features in Welsh Coordination. *Proceedings of the LFG99 Conference, University of Manchester*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
`http://csli-publications.stanford.edu/LFG/4/lfg99.html`

Sag, Ivan A., and Thomas Wasow. 1999. *Syntactic Theory: A Formal Introduction*. Stanford, Calif.: CSLI Publications.

Schwartze, Christoph. 1996. The Syntax of Romance Auxiliaries. *Proceedings of the First LFG Conference, Rank Xerox, Grenoble, August 26–28, 1996*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
`http://csli-publications.stanford.edu/LFG/1/lfg1.html`

Sells, Peter. 1985. *Lectures In Contemporary Syntactic Theories: An Introduction To Government-Binding Theory, Generalized Phrase Structure Grammar, and Lexical-Functional Grammar*. Stanford, Calif.: CSLI Publications.

Shieber, Stuart M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Stanford, Calif.: CSLI Publications.

Simpson, Jane. 1983. *Aspects of Warlpiri Morphology and Syntax*. Doctoral dissertation, MIT

Simpson, Jane. 1991. *Warlpiri Morpho-syntax: A Lexicalist Approach*. Dordrecht: Kluwer.

Sobin, Nicholas. 1987. The Variable Status of COMP-Trace Phenomena. *Natural Language and Linguistic Theory* 5: 33–60.

Spencer, Andrew. 2000. The Case for Separationism: WP ('Top-Down') vs. Morphemic ('Bottom-Up') Approaches to Morphology. *Proceedings of the LFG00 Conference, University of California, Berkeley*, ed. Miriam Butt and Tracy Holloway King On-line: CSLI Publications.
`http://csli-publications.stanford.edu/LFG/5/lfg00.html`

Stahl, J.F. 1967. *Word Order in Sanskrit and Universal Grammar*. Dordrecht: D. Reidel.

Wunderlich, Dieter. 1996. Lexical Categories. *Theoretical Linguistics* 22: 1–48.

Zaenen, Annie. 1983. On Syntactic Binding. *Linguistic Inquiry* 14: 469–504.

Zaenen, Annie, and Elisabet Engdahl. 1994. Descriptive and Theoretical Syntax in the Lexicon. *Computational Approaches to the Lexicon*, ed. B.T.S. Atkins and A. Zampolli, 181–212. Oxford: Clarendon Press.

Zaenen, Annie, and Ronald M. Kaplan. 1995. Formal Devices for Linguistic Generalizations: West Germanic Word Order in LFG. *Formal Issues in Lexical-Functional Grammar*, ed. Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, 215–239. Stanford, Calif.: CSLI Publications.

Zaenen, Annie, and Joan Maling. 1982. Passive and Oblique Case. *Papers in Lexical-Functional Grammar*, ed. Lori Levin, Malka Rappaport, and Annie Zaenen, 159–191. Bloomington, Ind.: IULC.

Zec, Draga. 1987. On Obligatory Control in Clausal Complements. *Working Papers in Grammatical Theory and Discourse Structure*, ed. Masayo Iida, Stephen Wechsler, and Draga Zec, 139–168. Stanford, Calif.: CSLI Publications.

Index