LEXICAL RULES IN THE HIERARCHICAL LEXICON

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Daniel Paul Flickinger
September 1987

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Thomas Wasow
(Principal Advisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Ivan A. Sag

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Joan Bresnan

Approved for the University Committee on Graduate Studies:

Dean of Graduate Studies & Research

Preface

In the summer of 1983 I joined a young computational linguistics research group whose members had just successfully completed a prototype of a natural language understanding system. In planning the next version of the system, the group members agreed to employ more sophisticated knowledge representation techniques in building the lexicon, to allow richer lexical entries and a smaller grammar. It was out of this early design work that the notion of word classes emerged as a way of capturing syntactic generalizations within the lexicon, making crucial use of inheritance of information to eliminate redundancy while allowing for idiosyncracy. The hierarchy of word classes that I present in the first part of this thesis draws heavily on the experience gained in constructing the new system, with the implementation often serving to sharpen the theory of word classes now in use.

Like the word class hierarchy, the theory of lexical rules presented in the second part of the thesis was developed gradually, with the division of labor between the two mechanisms emerging out of the experience of adding several familiar lexical rules to the system. It was to make more precise this division of labor that I began to view lexical rules as relations between classes of minimally specified lexical entries, both for inflection and for derivation. Within the framework that I present and illustrate in this work, lexical rules express simply and yet precisely the properties that two related lexical entries have in common, as well as those properties which are not shared. Through exposition and extended analyses of several English constructions, the work of these rules is distinguished on the one hand from that of the word classes, and on the other hand from that of the phrase structure rules, which employ fully-specified lexical entries.

It is my hope that this work will serve as the basis for at least three lines of continued research: first, it should provide a more precise framework for capturing morphological and syntactic generalizations; second, it leaves quite open the representation of lexical semantics; and third, it allows for more explicit proposals about distinctions between language-specific properties and those properties which are shared by all languages.

Acknowledgements

I have had the good fortune of gaining my education as a linguist in three quite different settings, giving rise to a debt of gratitude that I can only begin to acknowledge here. At the University of Iowa I learned the foundations of this discipline from professors who communicated both their sense of wonder and their commitment to craftsmanship in providing explanations; of lasting influence were Andreas Koutsoudas, William Ladusaw, Catherine Ringen, and my advisor Gregory Iverson.

At Stanford the list of influential people grows much longer, including fellow graduate students, post-doctoral fellows, staff, professors, and a bewildering wealth of scholars associated with the Center for the Study of Language and Information, which conveniently got underway when I began work on this dissertation. I single out for special thanks first a few of the Ventura Hall crowd, including Műrvet Enç, Nancy Wiegand, Susan Stucky (the other Mennonite formal linguist), and Kathie Carpenter, Suzanne Kemmer and Michael Barlow, with whom I have happily shared every step of the Stanford graduate pilgrimage. Next, I warmly thank Gina Wein for her competent administrative support and for her friendship. Finally, I gratefully acknowledge the strong shaping influences of the members of the Stanford linguistics faculty, who teach and also model a vibrant and professional approach to linguistic research. Representative of these scholars are the three members of my reading committee, whose work and counsel have had a profound effect on my work; I thank Joan Bresnan, Ivan Sag, and my principal advisor, Thomas Wasow, whose patience, cheerful persistence, unstinting support, solid critique, creative ideas, and common sense made the writing of this thesis possible and enjoyable. Every student should have such an advisor.

The third important collection of influential people are or have been associated with the Natural Language project at Hewlett-Packard Laboratories, where much of the research behind this dissertation took definite shape. Early and influential members of the group included Jean Mark Gawron, Elizabeth Ann Paulson, Geoffrey Pullum, Ivan Sag, Tom Wasow, and especially Carl Pollard, whose thoughtful, inventive suggestions are woven into the fabric of this work, and to whom I am deeply grateful for countless conversations and generous sharing of ideas. I also thank Derek Proudian, Lew Creary, Marilyn Friedman, John Nerbonne, and Dave Goddeau for sustained, helpful interaction in making more precise the formulation of the ideas presented here. And I acknowledge the important roles that Ira Goldstein and Steven Rosenberg have played, both as contributing researchers in knowledge representation and as supervisors of the research done in the Natural Language group.

Several linguists whose effect on this work must be acknowledged do not fall into the three groups above; Gerald Gazdar, Ewan Klein, Annie Zaenen, and Lori Levin have each been influential in shaping parts of the framework explored here, and I am grateful for the conversations and ideas they have provided.

I also acknowledge the generous financial support I received while writing this dissertation, including a National Science Foundation Graduate Fellowship for three years, four years of tuition at Stanford paid by the Hewlett-Packard Company, and office and computer resources provided by the Center for the Study of Language and Information at Stanford University.

Finally, I express my appreciation to my brothers Ron and John, my sisters Susan and Lois, and my parents for their astonishing patience, unwavering confidence, sympathetic ears, and sense of perspective, without which I would not have been willing to find my place in this profession. With love, I dedicate this work

To my parents,

Calvin and Sidonnia Flickinger

Contents

Chapter 1

Introduction

The status of the lexicon as an appropriate repository of generalizations about language has changed dramatically in the last ten years of work in generative grammar. Today enjoying a position of central importance in theoretical linguistics, the lexicon provides the basis for expressing generalizations about phenomena which include not only inflection and derivation, but also constructions once viewed as outside the scope of the lexicon, such as the passive relation and properties of control. Out of this work emerge two basic questions: (1) What mechanisms are necessary for expressing these lexical generalizations precisely enough to make them falsifiable? (2) What aspects of constituent structure within a sentence remain not accounted for once these lexical mechanisms are in place?

Work of the last ten years by Selkirk, Lieber, Anderson, Williams, Kiparsky, and others contains some concrete proposals for lexical representation and word structure as they affect inflection and derivation, but each of these assumes a rich, transformational syntactic component about which few details are provided. On the other hand, proposals about lexical rules found in the work of linguists following Bresnan, Wasow, and Dowty are concerned chiefly with capturing syntactic or semantic generalizations, within a framework where much or all of the work is done without recourse to transformations. But in this second line of research the structure of the lexicon and its connections to phrase structure are only partially specified.

Both lines of research have demonstrated that an understanding of the nature of regularities in the lexicon can yield explanations as well as new avenues of study in both morphology and syntax. What is needed in order to bring the two largely independent efforts together is an explicit articulation of the structure of the lexicon that must be in place to provide the foundation for syntactic explanations being offered in current work in the frameworks of Lexical Functional Grammar, Generalized Phrase Structure Grammar, and most recently, Head-driven Phrase Structure Grammar.

In this work I present a detailed view of the structure of the lexicon and two mechanisms for expressing regularities within the lexicon, then show how these mechanisms interact with, but do not obviate, the phrase structure rules. In the course of developing this view, I address the two questions posed above, at once broadening and constraining the scope of the lexicon as a tool for explaining properties of human language.

Organization of the thesis

In the next chapter I distinguish two ways in which lexical entries can be related, giving rise to two formal mechanisms for representing this sharing of properties. The first kind of relationship is category-based, where a set of lexical items share some syntactic properties, allowing us to predict (in part) the distribution of an item by knowing it belongs to that set, or word class. The second kind of relationship has a morphological basis, where two lexical entries of different categories have a common morphological base (ignoring suppletion), and often exhibit similar idiosyncratic properties.

The rest of Chapter 2 addresses the first of these two kinds of relations, introducing a set of word classes arranged in a hierarchy, with each class defining just those syntactic and semantic properties that are common to members of the class. In the model that I develop here, a given lexical item will often belong to several word classes, with each class contributing part of the definition of that lexical entry. Of course, the notion of word classes is as old as the study of language itself, but

linguists have in general not exploited the considerable power available when these classes are ordered hierarchically and viewed as defining default properties of their members. [Footnote: One exception to this rule is Hudson 1985, who does introduce the notion of a hierarchy of classes to encode grammatical function information.]

Implicit in the presentation of these word classes is a notion of inheritance, which I make more precise in Chapter 3, explaining how both default properties and idiosyncracies can be accommodated with a single formal mechanism. Having introduced the mechanism for inheritance, I then illustrate its use, and some of the theoretical issues it raises, in an explanation of how lexical items subcategorize for both complements and adjuncts.

In presenting the properties of lexical entries in terms of the word classes they belong to, I have of necessity made a commitment to a particular theory of grammar, being developed by Carl Pollard and Ivan Sag under the rubric Head-driven Phrase Structure Grammar (HPSG), in order to make the presentation concrete. I provide in Chapter 4 a sketch of how the particular attributes and values used to define lexical entries interact with a small set of grammar rules and universal feature principles to provide the syntactic structures, the bases for semantic interpretation, for grammatical sentences of English.

In Chapter 5 I characterize a second kind of generalization which holds between the elements of two related classes, and introduce the familiar notion of the lexical rule to capture this kind of relationship. Based on the reasonably explicit proposal for lexical structure developed in the previous chapter, I construct a formal representation for lexical rules which should provide sufficient expressive power while incorporating some important constraints to limit that power. This choice of representation depends on both the inheritance mechanism of the previous chapter and also the consistent structuring of information in lexical entries presented there.

Chapter 6 serves as an extended set of illustrations of how the mechanisms introduced here serve in the explanation of grammatical phenomena. I develop lexical rule analyses of adjectival passives, account for familiar but difficult data involving "tough-movement" adjectives, then show how the same analysis of these adjectives also correctly predicts the grammaticality of "too/enough" constructions.

The chapter concludes with an account of one class of unbounded dependencies, again employing a lexical rule to account for what has traditionally been viewed as a phenomenon well outside the reach of lexical mechanisms.

Finally, in Chapter 7 I summarize the mechanisms introduced in the course of the presentation, and the constraints both implicit and explicit that I proposed to govern the applicability of these tools, on the one hand distinguishing the work of lexical rules from that of word classes and inheritance, and on the other hand maintaining a sharp distinction between lexical rules and phrase structure rules.

Current research leaves no doubt that morphology and syntax should be studied together, that word structure and sentence structure share important properties. I demonstrate in this work that the choice of representation developed here permits explanations for many of these shared properties, while raising new questions about the interaction of morphology and syntax.

Chapter 2

Structure of the Word Class Hierarchy

Much of the information in a fully specified entry within the lexicon is not unique to that particular entry. Viewing this information as a set of discrete properties which make up the lexical entry, related lexical items will have in common some of the properties of that entry. Lexical items can then be grouped into classes defined by those properties that are common to all members of the class. By giving a precise characterization of these word classes, one can eliminate a good deal of the redundancy found in a lexicon that consisted of fully specified entries. Put differently, one can make use of these word classes to capture generalizations about the elements of the lexicon, and to make predictions about the behavior and distribution of a lexical item on the basis of the classes it belongs to.

To avoid redundancy entirely, each property relevant to representing the elements of a lexicon should only be mentioned once in some single class, with all elements of the lexicon that have this property being members of that class. If so, a given lexical item may have to belong to many classes in order to obtain all of its properties. I argue in what follows that this multiple class membership is both necessary and desirable for representing lexical information, and I develop the notions of a word class hierarchy and the mechanisms for inheritance of properties. As will become clear in the ensuing discussion, the structure that I develop, and which I refer

to loosely as a single hierarchy, is not a simple hierarchy in the strict sense, but rather a structured collection of simple hierarchies, with a given word class often belonging to more than one hierarchy. As Chomsky argued in his discussion of lexical representation in Aspects (pp. 79-83), the need to cross-classify a given lexical item according to several distinct properties renders impossible the use of a simple branching hierarchy to represent the lexicon.

Chomsky was arguing for the use of syntactic features in describing lexical categories, and suggested relating these features in a set of hierarchical structures. The word class structure I present might well be viewed as an expansion of these two early insights, though it has grown out of more recent work on the lexicon and lexical rules, work which gives the lexicon a more central role in explanations of syntactic phenomena than Chomsky assumed in his original Aspects model.

To see more clearly the need for a rich word class structure, imagine trying to represent the information about English auxiliary verbs while maintaining the hypothesis that each entry belongs to only one word class. Then it should be enough to define a single class for auxiliaries, called AUXILIARY-VERB, which completely specified just those properties that appear in the fully specified entry of each auxiliary verb. Such a characterization misses two kinds of generalizations. First, auxiliaries have properties in common with other verbs, including the ability to be head of a sentence and to carry tense marking. This suggests that those properties defining AUXILIARY-VERB should be divided into two sets, separating properties common to all verbs from those unique to auxiliaries. Second, English auxiliary verbs are not all alike with regard to syntactic properties; they must be distinguished at least on the basis of the form of their VP complement (base form for modals[1], past participle for perfectives, present participle for copula), suggesting that an orthogonal subclassification of AUXILIARY-VERB is needed for an adequate description of the English lexicon.

A second simple illustration of the need for assigning words to more than one class involves two independent criteria for partitioning nominal forms in English,

[1]For support of the treatment of modals as auxiliary verbs, see Pullum and Wilson 1977, and Gazdar, Pullum, and Sag 1982.

and is a variant of the original argument advanced by Chomsky for the use of syntactic features. One partition of nouns distinguishes among common nouns like *book*, proper nouns like *John*, and pronouns like *they* (at least), where each of these three noun types exhibits both syntactic and semantic properties not shared by the other two types. Another, orthogonal partition of lexical nominals distinguishes on the basis of number, separating singular from plural from mass nominals. Since both a common noun like *books* and a pronoun like *they* are plural, no disjoint classification scheme will allow both the grouping together of all plurals and also the grouping of all pronouns. Since both types of grouping enjoy linguistic motivation, multiple-class membership of lexical items must be employed to capture the relevant generalizations.

In general, I take as adequate motivation for the existence of a word class the demonstration that some particular syntactic or morphological property (or cluster of properties) is shared by a number of lexical items. Again, the forcing function in the identification of word classes is my assumption that the best representation for lexical information is one in which each new piece of information, each distinct property exhibited by one or more elements of the lexicon, is introduced exactly once in the lexicon. A property shared by more than one lexical item should be introduced in a word class common to those items, or those lexical items should be related by lexical rule. Maintaining this minimalist assumption about the introduction of properties leads one to a model of the lexicon in which the information that comprises a fully specified word is rather sparsely distributed over a large number of word classes organized in a non-trivial set of hierarchies.

Structure of the chapter

In section 2.1 I present a characterization of the information that must be in a fully specified lexical entry, and motivate the use of word classes to represent that information. In section 2.2 I introduce the overall structure of the class hierarchy, then devote the rest of the chapter to a detailed (though incomplete) presentation of the classes in this hierarchy needed for English. Section 2.3 elaborates the classes

containing information about the complementation properties of lexical items, and section 2.4 describes the second main grouping of classes, which specify more traditional part-of-speech information.

2.1 Lexical entries

I follow a long tradition in taking lexical entries to contain three distinct types of information: phonological, syntactic, and semantic. Of the three, the phonological properties of lexical items receive no attemtion here, and their semantic properties little more, with the emphasis placed instead on representing their syntactic properties

For convenience of exposition, I view the syntactic properties of lexical items as being of two kinds: one a set of features, separated into those with atomic values and those with category values, and the other a set of subcategorization specifications, divided into complements (obligatory and optional), and adjuncts. As Pollard 1984 has shown, both types of information (features and subcategorization specifications) can be formally represented as attribute-value pairs, but the notation needed for this unity of representation threatens to obscure the points addressed below, so the two kinds of information are separated in the representations of lexical entries used here.

The atom-valued features that help compose a lexical entry (and specify its category) are drawn from a (small) finite set where each feature has a limited set of possible atomic values. [For a presentation of the basic theory of syntactic features adopted here, see GKPS 1985:17ff.] These atom-valued features may be simple binary features (values + or -), or they may draw their values from a larger set. The feature INVERTED, for example, is binary, indicating whether or not a verb can appear as the head of an inverted sentence. The feature VFORM, on the other hand, draws its values from a set containing among others BASE, PAST, and PAST-PARTICIPLE, to represent the morphological form of a verb.

Category-valued features, on the other hand, take as their value a set of feature-value pairs, a specification for some syntactic category. Since any non-empty set of feature-value pairs is (by definition) enough to specify a category, any such set constitutes a possible value for one of these category-valued features. Naturally, if any of the features in this set which make up a value could themselves be category-valued, the door would be open to a non-finite number of possible categories. For the present, I arbitrarily restrict the features in one of these value sets to being

atom-valued, setting aside the interesting question of whether this restriction can be motivated or rejected on principled grounds.

The other kind of syntactic information contained in a lexical entry is subcategorization information, which specifies the syntactic restrictions imposed by the lexical item on its complements and adjuncts. For expository purposes I separate these specifications into four parts: (1) a partially ordered list of names of obligatory and optional complements; (2) an unordered list of names of adjuncts; (3) descriptions of each of the complements and adjuncts mentioned in (1) and (2), specifying both syntactic and semantic properties; and (4) a set of partial ordering statements that interact with phrase structure rules, as discussed below. Each complement or adjunct entry, referred to here as a *subcat*, consists of a category specification (set of feature-value pairs) and its semantic properties, here limited to thematic role assignment. These category specifications for subcats may themselves have to employ both atom-valued and category-valued features, but crucially do not themselves make reference to subcategorization properties of their own. This last point is addressed in some detail in the next chapter.

Since reference to subcategorization properties of subcats is excluded in specifying complements or adjuncts within a lexical entry, I make use of a feature COMPLETE, quite similar to the SUBJ feature proposed by Borsley 1983, and employed in GKPS 1985:61f to distinguish *incomplete* from *complete* categories. Incomplete constituents lack one or more of their obligatory complements, including at least their final complement (usually the subject), and are marked [COMPLETE -], while complete categories are marked [COMPLETE +] to represent the property that no obligatory arguments are missing. [Complete categories correspond roughly to maximal projections in an X-bar framework.] To distinguish lexical categories from phrasal ones, I use the binary feature LEXICAL, discussed in more detail in section 2.4. With these two features COMPLETE and LEXICAL, I can follow Pollard 1984 in dispensing with the widely used (and abused) X-bar machinery, while maintaining the full range of necessary distinctions among lexical and phrasal categories.

2.1.1 Fully-specified lexical entries

To anchor this discussion, I give below a fully specified lexical entry for the past-tense form of the verb *try*, with mere placeholders for both phonology and semantics, and with the counterfactual assumption that the set of syntactic features includes only CAT, VFORM, INVERTED, LEXICAL, CASE, and COMPLETE. The names of these features should be relatively transparent for present purposes, with the possible exception of CAT, which takes as its value one of the basic non-decomposable category labels (e.g., Verb, Preposition, Determiner, Conjunction). I include a Spelling attribute since it is a property of lexical items in written English, one that all too often is not predictable from other properties of the word.[2]

(1) Fully-specified lexical entry for *tried* as in *John tried to win.*

TRY-1-PAST	
Spelling	"tried"
Phonology	/traId/
Semantics	(PAST (TRY agent:X prop:Y))
Atomic-features	(CAT Verb) (VFORM Finite) (LEXICAL +)
	(COMPLETE -) (INVERTED -)
Category-features	
Complements	Subject XComp
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative)
Subject-index	X
XComp-features	(CAT Verb) (VFORM Infinitive)
	(LEXICAL -) (COMPLETE -)
XComp-index	Y
Adjuncts	...

Using a notation that I will explain more fully below, this entry specifies that *tried* has the following syntactic properties: it is a non-inverted finite past-tense lexical verb which takes two obligatory arguments, which I have labeled *Subject* and

[2] In breaking with tradition by including spelling for lexical entries, I do not propose to give a full account of written English complete with treatment of conventions for punctuation, capitalization, and so on. I only specify that each lexical entry for English includes a conventional spelling which is only partially determined by the phonological and morphological properties of that entry. Hence lexical rules which relate these entries may also make reference to spelling, as seen in the next chapter.

XComp, the latter drawn from terminology used in Bresnan 1982. The subject is specified to be a nominative-case noun phrase, while the complement must be an infinitival verb phrase. No optional complements are available, and the possible adjuncts require a separate discussion, taken up in section 2.5, which also includes motivation for the distinction between optional complements and adjuncts.

The semantic properties of this entry are only partially represented here, to suggest the general approach I adopt, which has been developed by Pollard and Sag 1987. Briefly, the lexical semantics for *tried* is given in the Semantics attribute as a predicate expression with role assignments for its complements, where the index for each role appears in one of the subcat specifications. The control properties of this and other equi-verbs are not specified lexically, but follow from a general principle that correlates control with relative obliqueness of grammatical functions. Sag 1986, and Pollard and Sag 1987 provide a detailed account of this theory of control. For the present, what is important is that each subcat specification in a lexical entry may include both syntactic information and semantic information, where the semantic information is minimally a link to some thematic role in the main semantics of the lexical entry.

Now examine a second lexical entry, for the verb *persuaded*:

(2) Full entry for *persuaded* as in *John persuaded Sally to win.*

PERSUADE-1-PAST	
Spelling	"persuaded"
Phonology	/p@rsweId@d/
Semantics	(PAST (PERSUADE agent:X patient:Y prop:Z))
Atomic-features	(CAT Verb) (VFORM Finite) (INVERTED -) (LEXICAL +) (COMPLETE -)
Category-features	
Complements	Subject DObject XComp
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative)
Subject-index	X
DObject-features	(CAT Noun) (COMPLETE +) (CASE Accusative)
DObject-index	Y
XComp-features	(CAT Verb) (VFORM Infinitive) (LEXICAL -) (COMPLETE -)
XComp-index	Z
Adjuncts	...

The syntactic properties in these two entries have much in common; the only difference is that *persuaded* has an extra obligatory argument, a direct object which must be an accusative-case noun phrase. To eliminate the redundancy present in these two entries and in other *subject-equi* and *object-equi* verb entries, one might define classes of similar lexical entries, where the definition for each class consists of those properties shared by the members of the class. As a starting point, consider defining a class of Equi-Verbs as follows.

(3) Provisional definition for class of equi-verbs

EQUI-VERB-CLASS	
Atomic-features	(CAT Verb) (LEXICAL +) (COMPLETE -) (INVERTED -)
Complements	Subject XComp
Subject-features	(CAT Noun) (COMPLETE +)
XComp-features	(CAT Verb) (VFORM Infinitive) (LEXICAL -) (COMPLETE -)
Adjuncts	...

A comparison of this definition with the entry for *tried* will reveal that the only additional syntactic properties specified for *tried* are its morphological form, namely (VFORM Finite), and the requirement that its subject must be nominative case. Since these two properties are common to all past-tense verbs including *persuaded*, a more revealing class definition suggests itself, one which I present in (4) for illustration here, but which I will divide further in the next section of this chapter.

(4) Provisional definition for class of past-tense forms

PAST-CLASS	
Atomic-features	(VFORM Finite)
Complements	
Subject-features	(CASE Nominative)

Merging the information in the classes defined in (3) and (4) produces just those syntactic properties specified for *tried*, but not quite for *persuaded*. To represent the missing information, one might propose an "Object-Equi-Verb" class which duplicated all of the information in (3), and added a specification for the Object complement. To avoid this remaining duplication, we can define the Object-Equi-Verb class to be a subclass of the general Equi-Verb class, merging properties of the general class with its own more specialized properties. A definition of Object-Equi-Verb might then be as in (5).

(5) Provisional definition for class of object-equi verbs

OBJECT-EQUI-VERB-CLASS	
Superclass	EQUI-VERB-CLASS
Complements	DObject
DObject-features	(CAT Noun) (COMPLETE +)
	(CASE Accusative)

Now merging the information in (3-4) produces just the syntactic properties ascribed to *tried* in (1), while the properties of *persuaded* in (2) are distributed in the classes in (3-5). Given these class definitions, the entry for *tried* could be given as in (6) without loss of information, provided that the notion "merging of information" can be made precise.

(6) Non-redundant lexical entry for *tried* (preliminary)

TRY-1-PAST	
Classes	EQUI-VERB-CLASS, PAST-CLASS
Spelling	"tried"
Phonology	/traId/
Semantics	(PAST (TRY agent:X prop:Y))
Subject-index	X
XComp-index	Y

It is clear that this entry for *tried* has no idiosyncratic syntactic properties, drawing all such properties from the definitions of the classes to which it belongs. What remains in this non-redundant entry is the information about phonology, spelling, and semantics, with the semantic specifications linking the subcats with the main semantics through the index values. [3]

As will become clear in the next section, the lexical entry for a syntactically related raising verb like *failed* will be similar to (6), differing in spelling, phonology, and semantics, but having in common with *tried* all of its syntactic properties. What distinguishes *tried* from *failed* (based on distinct class membership) is a semantic property: the assignment of a thematic role by *tried* to its subject, an assignment which *failed* will not make because it is a subject-raising verb.

The entry for *persuaded* will be as simple as that for *tried*:

(7) Non-redundant lexical entry for *persuaded* (preliminary)

PERSUADE-1-PAST	
Classes	OBJECT-EQUI-VERB-CLASS, PAST-CLASS
Spelling	"persuaded"
Phonology	/p@rsweId@d/
Semantics	(PAST (PERSUADE agent:X patient:Y prop:Z))
Subject-index	X
DObject-index	Y
XComp-index	Z

[3] Note the similarity of (6) to the structure of lexical entries employed in Lieber 1981:163, a similarity which is strengthened with the introduction in the next chapter of exceptional information about lexical rules.

Again, the only idiosyncratic information about *persuaded* concerns its spelling, phonology, and semantics; all of its syntactic properties are predictable given the word classes it belongs to. Lexical entries for other object-equi verbs like *convinced* will be quite similar. [4]

The definition in (3) clearly collapses several separate clusters of properties, with some of the properties being true of English words in general, some true of verbs, and some relevant only for equi verbs, at least. For example, words are generally [LEXICAL +], all verbs are [CAT Verb] and take an obligatory subject NP, but not all verbs take a VP complement. A more principled representation of the properties in (3), then, will require the elaboration of a set of classes linked in a non-trivial set of hierarchies, with each class defined by clusters of the properties needed to characterize fully-specified lexical entries.

On this approach, the burden of representing predictable lexical information is now shifted from lexical entries to the definitions of the classes they belong to. In the remainder of this chapter I propose a partially developed hierarchy of word classes motivated by data from English, to represent this predictable information.

[4]Of course, the entry for *persuaded* in (2) is simplified for the sake of exposition; for example, the verbal complement can be either a VP or sentential, and is not really obligatory. To decide whether these properties ought to be part of the definition of the class of object-equi verbs, or idiosyncratic properties of the verb *persuade*, one needs to examine the range of verbs in this class, and extract the proper generalizations.

2.2 Word class hierarchy

To focus the following discussion of the hierarchy of word classes, I will restrict the set of syntactic features to a subset of those needed even for English. I take the identification and motivating of an adequate set of features and their values to be one of the principal ongoing tasks of the syntactician (in cooperation with the semanticist). Thus the hierarchy as presented here, while necessarily incomplete, should be viewed as representing the current state of work in progress, sufficiently detailed to provide an adequate basis for subsequent discussion of lexical entries and the lexical rules that relate them.

The division of a class of words into subclasses within the hierarchy must be based on one or more linguistically motivated distinctions among the members of that class. In developing the hierarchy below, I do not always give explicit arguments for the distinctions I draw, often because I take the relevant distinctions to be familiar. Where I omit argumentation, the lack must not obscure my position that motivation based on language-particular or universal generalizations is necessary to justify each class in the hierarchy. I have tended, for expository reasons, to draw those distinctions which are necessary for English, returning later to the important question of how to distinguish language-particular and universal information in the construction of a more general word class hierarchy.

Before launching the construction of this hierarchy for English, I note that distinctions must often be made along more than one dimension for the members of a given word class. That is, the members of the same class may be grouped differently for two separate properties; for example, members of the class of nouns can be divided into groups on the basis of number (singular, plural, or mass), but members of this class can also be grouped on the basis of whether or not they may take a determiner (proper vs. common). These two ways of sub-dividing the class of nouns are partially independent, yet each of the two holds for the full class of nouns. The need for such orthogonal distinctions means that the hierarchy of word classes does not have a simple branching tree structure where daughter nodes always represent discrete subsets of the parent class. Instead, I employ two types of links between

classes in the hierarchy: a *subset* link joins a class and one of its proper subsets, while a *perspective* link joins a class with a node that names one dimension along which that class will be sub-divided. It may prove to be the case that these two types of links need not be formally distinguished, since it seems that inheritance of information via the two types of links is the same. But for the moment I keep them distinct, introducing the notation below.

2.2.1 *"In the beginning was the Word . . . "*

The first fundamental distinction to be drawn in the lexicon is one of perspective: all words can be grouped along two separate dimensions, one involving the traditional part-of-speech distinctions, and the other involving the number and type of complements that a lexical item requires. I refer to the first distinction simply as PART-OF-SPEECH, and use the term COMPLEMENTATION for the second.

Within the PART-OF-SPEECH partitioning of the lexicon, words fall into two broad subclasses, MAJOR and MINOR, with the subclasses of MAJOR being the familiar NOUN, VERB, ADJECTIVE, and PREPOSITION, and some of the subclasses of MINOR being DETERMINER, CONJUNCTION, and COMPLEMENTIZER. I will take up each of these in some detail below. [5]

The second, orthogonal partitioning of the lexicon, for COMPLEMENTATION, groups words at the highest level into COMPLETE and INCOMPLETE categories; a member of the COMPLETE class does not require any arguments, while a member of the INCOMPLETE class specifies one or more (obligatory) arguments. The

[5] Prepositions are grouped here with the other MAJOR categories since they share with verbs, nouns, and adjectives the one property ascribed to major categories below: they take PP adjuncts. Prepositions share some additional crucial properties with verbs, nouns, and adjectives, distinct from the minor categories: prepositions subcategorize for objects, sentential complements, and subjects; they license trace introduction; and they appear as heads of predicative phrases, along with adjectives, verbs, and nouns. These properties do not make it necessary to treat PREPOSITION as a subclass of MAJOR, since said properties are not formally ascribed to the class MAJOR, but they provide additional weight in favor of the classification given here. The fact that they form a closed class might, as Croft 1986 proposes, suggest they should be treated as a minor category, but I have not here made any formal commitment to a correlation between the major/minor division and the open-class/closed-class distinction. If I had, it would also be necessary to treat auxiliary verbs as members of a minor category, a move inconsistent with the proposal made below.

structure of this part of the hierarchy is derived from Pollard 1984's hierarchy of control, though it differs in some important respects identified in the discussion to follow.

I now introduce the root class for the lexicon hierarchy, which I have labelled the WORD-CLASS.

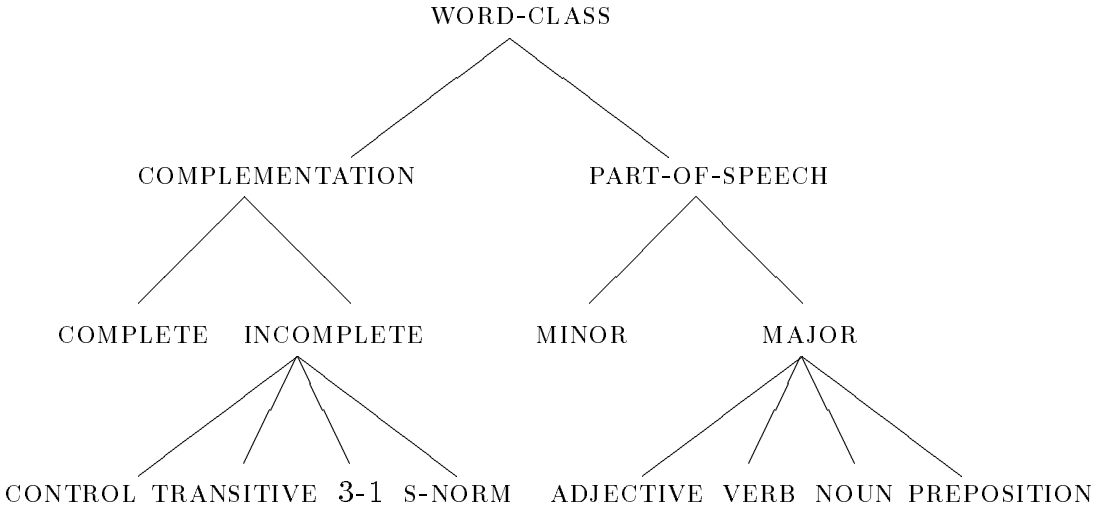
WORD-CLASS
Atomic-features (LEXICAL +)
Lexical-rules

The actual content of this class is minimal, specifying only the unsurprising claim that the value of the feature LEXICAL is by default positive for all entries in the lexicon, and noting that information about lexical rules may be specified for any subclass or actual lexical entry. Like all values of attributes specified in these class definitions, the value of the LEXICAL attribute here is a default value, one that in principle may be overridden by some subclass or actual lexical item; in Chapter 6 I introduce a lexical entry which does have a non-default value for LEXICAL. Section 2.3 presents a thorough discussion of the inheritance mechanism employed here which makes use of these defaults.

To simplify exposition, I will not introduce in this chapter any specifications for lexical rules in the definitions of word classes, reserving that topic for the next chapter. In general, what can appear in the Lexical-rules attribute for a class are the names of lexical rules that hold for members of that class, together with any idiosyncratic restrictions on those rules (see section 5.2).

I will use diagrams like those in (1) below to represent the structure of the word class hierarchy, with a dashed line indicating a *perspective* link between a class and one dimension along which that class is partitioned; an unbroken line indicates a *subset* link between the parent class and a proper subset, within a given perspective.

(1) Top level of word class hierarchy



Each of the labels in the structure in (1) can be thought of simply as a convenient name for some collection of properties that define the named class. I use the label WORD-CLASS for the class which includes all members of the lexicon, assigning only those default values which hold quite generally for elements of the lexicon, such as the property (LEXICAL +). The classes labelled COMPLEMENTATION and PART-OF-SPEECH each define the range of attributes relevent for their respective subclasses, in general leaving to those subclasses the task of assigning default values for each of those attributes.

In the next section I define the classes making up the COMPLEMENTATION part of the hierarchy, then turn in the last section to the PART-OF-SPEECH classification.

2.3 Complementation hierarchy

I begin a detailed discussion of word classes by presenting the COMPLEMENTATION part of the lexicon hierarchy. Each class description specifies either the classes that it is a proper subset of (if it is joined with a *subset* link), or the class that it partitions (if it is joined with a *perspective* link). The rest of each class description is then divided into properties shared as a default by members of the class; these include feature specifications and subcategorization information. The two types of features represent strictly syntactic information; features that appear in the *Atomic-features* slot have only atomic values, while *Category-features* take category specifications (feature bundles) as their values. Information about subcategorization includes both syntactic and semantic properties, with the semantic information here including only thematic role assignment, which I take to be assigned to each subcat by individual lexical entries.[6] Word class descriptions will only specify whether or not a given complement can be assigned a role, determined here by whether or not the relevant index attribute is available for members of that class, as illustrated by the distinction between the equi and raising classes below.

COMPLEMENTATION			
Partition-of	WORD-CLASS		
Atomic-features			
Category-features			
Complements			
LP-constraints	Head [LEXICAL +]	<	Complement

I will use several notational conventions in these class descriptions which will appear in the course of the discussion. Attribute names appear on the left, and values on the right, with each line of the description specifying an attribute (and value, if any). In the above description, I have only identified the principal types of information which make up the definition of subclasses of COMPLEMENTATION; values for these attributes will be partially specified in these subclasses. In addition,

[6] Given the substantial work now being done on thematic roles, there is hope of eventually predicting thematic role assignment on the basis of lexical semantics. I do not pursue the question of how such generalizations should be integrated with the notion of word classes developed here.

in specifying values for the Complements attribute of classes or lexical entries, I will use the familiar parenthesis convention to distinguish optional from obligatory complements.

The one attribute given a value in this class is the *LP-constraints* attribute, which contains default linear precedence constraints. While the single constraint specified here is certainly not the only one necessary for English, it is included here for illustration. It states that a lexical head in a phrase must precede all of its complements, a claim that holds true for the majority of phrase types in English, as argued by Pollard and Sag, and earlier by GKPS 1985:49. Constraints like this one interact with the phrase structure rules (introduced in chapter 4) to determine the left-to-right order of heads, complements, and adjuncts in a phrase. See Pollard and Sag (in preparation) for a detailed presentation of this interaction between LP constraints and grammar rules. I follow their approach quite closely here, though by introducing these constraints as inherited defaults, I enable subclasses or lexical items to override the defaults, a capability that I will show to be desirable.

Before embarking on an explanation of the particular LP constraint introduced in this class, I note an additional important convention which I borrow from Pollard and Sag, involving the order of elements in the Complements attribute. They argue that if the list of complements for a lexical entry is taken to be a partially ordered list which directly reflects the relative obliqueness of the complements, both the linear order of those complements in a phrase, and their semantic control properties can be predicted from such an ordered list. On this approach, subjects are defined as least oblique, then direct objects, then indirect objects, then other complements (all equally oblique), and then all adjuncts. The Complements attribute encodes this order based on obliqueness, though I delay until the next chapter an account of how this ordered list interacts with the grammar rules and the linear precedence constraints.

While obliqueness does provide a quite general account of many linear order facts, it must be taken as a default, not as an absolute, as Pollard and Sag note at least for focused constituente, which may violate the default order. Thus stressed or *heavy* phrases may appear first or last in a sequence of complements even though

that order is not congruent with properties of obliqueness. In Chapter 4 I take up some additional examples that motivate making order-by-obliqueness a default property, subject to exceptions like any other attribute of word classes and lexical entries presented here.

2.3.1 Subclasses of COMPLEMENTATION

As proposed in the previous section, there are two subclasses of COMPLEMENTATION, named COMPLETE and INCOMPLETE, corresponding to whether the *Complements* attribute is empty or not. Little needs to be said about the first of these two, the COMPLETE class.

COMPLETE	
Superclasses	COMPLEMENTATION
Atomic-features	(COMPLETE +)

The only default property of all elements of the COMPLETE class is here claimed to be that they do not require any obligatory complements, given the intended interpretation for the COMPLETE feature. Consistent with this, no value (obligatory or otherwise) is specified for the *Complements* attribute in this class description, nor was any given in the parent class, COMPLEMENTATION.

Lexical items which are members of the COMPLETE class include proper nouns, pronouns, and pro-forms like *so*, which can appear in place of finite S complements, as illustrated in (1).

(1) a. John said Sally would sing.
b. John said so.
c. John told me Sally would sing.
d. John told me so.

The other subclass of COMPLEMENTATION is the INCOMPLETE class, which claims as members all those lexical items which require one or more obligatory complements.

INCOMPLETE	
Superclasses	COMPLEMENTATION
Atomic-features	(COMPLETE –)
Complements	Subject
Subject-features	(CAT Noun) (COMPLETE +)
Subject-index	

Distinct from the COMPLETE class, the INCOMPLETE definition specifies that members of this class require (at least) one obligatory argument, a property that seems to be redundantly specified in this definition. But the two statements referring to 'COMPLETE' and 'Subject' respectively, serve quite different functions. The atomic feature specification (COMPLETE –) carries only the information that predicates of this class require at least one complement. This feature value interacts with the phrase structure rules introduced in chapter 4 to ensure the proper constituent structure for English. On the other hand, the mechanism for specifying which complements a head takes and what their properties are makes use of the list of complements, including but not restricted to subjects. This latter mechanism uses the labels Subject, DObject, and the like to point to the specifications for their syntax and semantics which also appear as part of the class description or lexical entry. In the INCOMPLETE class, the least oblique complement (labelled Subject) is introduced, and partial specifications are given for its syntactic properties, also making provision for assignment of a thematic role index. Additional properties may be specified in subclasses of this class, or in subclasses of MAJOR, or in actual lexical entries. [7]

Lexical items which are members of the class INCOMPLETE include intransitive verbs and common nouns, as well as items which are members of subclasses of INCOMPLETE. Put differently, the subclasses of INCOMPLETE do not exhaustively partition the class INCOMPLETE, since some lexical items are directly members of the class itself, rather than members of some subclass. If motivation is found

[7] Bresnan 1982:301 mentions the notion of default syntactic properties of subcats, and gives as possible defaults for Subject, DObject, and XComp values equivalent to those presented here, modulo notation.

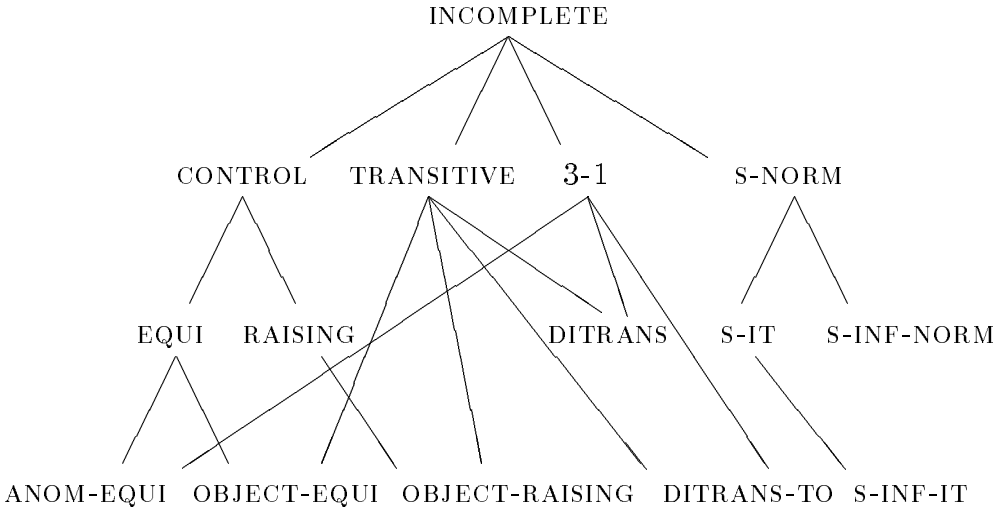
for a separate subclass of INCOMPLETE called INTRANSITIVE, then the resulting three subclasses of INCOMPLETE would exhaustively partition it; but classes in general may have both subclasses and actual instances, so the postulation of a class INTRANSITIVE would require demonstration of properties shared by, say, intransitive verbs and common nouns which are not shared by, say, transitive verbs.

2.3.2 Subclasses of INCOMPLETE

The subclasses of INCOMPLETE are CONTROL, TRANSITIVE, 3-1, and S-NORM. Members of CONTROL define one additional oblique argument, a controlled complement, and fall into two subclasses, for equi and raising predicates. Members of the TRANSITIVE subclass have, in addition to the subject, at least one argument, the direct object, which is a possible controller, more oblique than the subject, but less oblique than all other complements. These predicates include simple transitives, ditransitives, and object-control predicates (both object-equi and object-raising). Members of the 3-1 subclass of INCOMPLETE introduce a complement more oblique than the direct object, one which is neither controlled nor a possible controller, here given the familar label of indirect object. Members of the S-NORM class take an oblique verbal complement which is either sentential or given arbitrary control.

The subclasses of INCOMPLETE can be arranged in an extension of the diagram of the previous section as follows:

(2) Subhierarchy of WORD-CLASS rooted in INCOMPLETE



To begin with, the definition of the CONTROL class specifies that its members have a controlled complement, by default obligatory (in addition to the subject which they have by virtue of being members of a subclass of INCOMPLETE). The properties of this argument, the XComp, are partially specified here, and will be further specified (or overridden) in subclasses of this class, or in actual lexical entries. Included in this definition is the attribute XComp-index, whose value will be assigned by each lexical item belonging to this class.

CONTROL	
Superclasses	INCOMPLETE
Complements	XComp
XComp-features	(CAT Verb) (COMPLETE –) (LEXICAL –)
XComp-index	

The CONTROL class is divided into two familiar subclasses, EQUI and RAISING, with the distinction treated here as primarily a semantic one involving the way in which thematic roles are assigned to the controller of the XComp. What the CONTROL classes determine is whether or not a Subject-index or DObject-index attribute is permitted as part of the definition for members of that class.
The RAISING class does not permit a Subject-index assignment; this is indicated

with an obvious notation, which should be interpreted to mean that the attribute itself is not inherited, so a lexical item directly a member of this class could not assign a value to Subject-index, since it does not inherit the attribute. (I will use this notion of blocked-inheritance again in the chapter on lexical rules.) The EQUI class does permit the assignment of an index for the thematic role of the subject, since the attribute is inherited from the INCOMPLETE class, and not blocked. Likewise, the OBJECT-RAISING subclass of RAISING does permit a Subject-index attribute, but not one for DObject-index, while the OBJECT-EQUI class does, as seen below.

RAISING	
Superclasses	CONTROL
Complements	
Subject-index	DO-NOT-INHERIT

EQUI	
Superclasses	CONTROL
Complements	
Subject-features	(NFORM Norm)

Lexical entries which are directly members of the RAISING or EQUI class will have as their complements a subject and an XComp, with the subject predicted to be the controller, since it is the next (indeed only) less oblique complement assigned to members of this class. These subject-control predicates will be either subject-equi (like *try*) or subject-raising (like *seem*). EQUI predicates also have a syntactic constraint which prevents the expletives *there* and *it* from appearing as subject; the OBJECT-EQUI class includes a like constraint on the direct object. [8][9]

Object-control verbs are only indirectly members of RAISING or EQUI, by virtue

[8] This syntactic constraint might be better represented as a semantic constraint on co-indexing, but this would require a more detailed semantic account of expletive pronouns (and noun phrases in general) than I am prepared to give here. So I retain the syntactic account, in part to show how the mechanism could work. This is a good example of where the need for particular syntactic features is affected by the semantic assumptions made.

[9] One unexpected but desirable minor consequence of this formalism is the prediction that there will be no class of verbs which have a *there* or *it* subject but a normal direct object along with an XComp, since the restriction on normal objects is imposed on OBJECT-EQUI, which inherits from EQUI, which blocks *there* and *it* subjects.

of being directly members of the class OBJECT-RAISING or the class OBJECT-EQUI, which I introduce after defining the other two immediate subclasses of INCOMPLETE.

TRANSITIVE, the second subclass of INCOMPLETE, includes all lexical items with a direct object: a complement more oblique than the subject but not semantically controlled, and whose default syntax is an accusative NP.

TRANSITIVE	
Superclasses	INCOMPLETE
Complements	DObject
DObject-features	(CAT Noun) (COMPLETE +) (CASE Accusative)
DObject-index	

The content of this class is similar to that of CONTROL, though with a direct object specified instead of the XComp. Again, the properties of this direct object are partially specified here, and will be further specified (or overridden) in subclasses of this class or in actual lexical entries. Lexical entries in this class will then have at least a subject and a direct object as complements; those entries which are directly members of TRANSITIVE will have only a subject and a direct object. As shown in the diagram in (2), each subclass of TRANSITIVE is also a subclass of another COMPLEMENTATION class, so lexical entries which are members of some subclass of TRANSITIVE will inherit an additional argument, as described below.

Lexical entries inheriting from the TRANSITIVE class which also inherit from the EQUI or RAISING classes will have the controlled complement be controlled by the direct object, rather than the subject, since the direct object is the next complement after the XComp, in descending order of obliqueness. Again, this object-control does not have to be stipulated here, since the general theory of semantic control is driven by properties of obliqueness, where these are (in part) independently motivated by linear order facts. The classes for these object-control predicates are given below.

The third subclass of INCOMPLETE I have labelled 3-1, drawing the term from work in relational grammar; this class specifies an oblique complement I will call the indirect object (IObject), whose default syntax is a prepositional phrase introduced

by the preposition *to*. I will assume that as a default, indirect objects are no more or less oblique than controlled complements, which accounts for the relative freedom of order between indirect objects and other oblique complements, when other ordering constraints do not interfere.

3-1	
Superclasses	INCOMPLETE
Complements	IObject
IObject-features	(CAT Preposition) (COMPLETE –) (PFORM To)
IObject-index	

Lexical entries which are members of this 3-1 class (and not also members of TRANSITIVE) can be distinguished from members of the TRANSITIVE class in at least two ways relevant to the present discussion: if such entries also have a controlled complement, it is the subject, not the object, which is the controller; and such entries have no corresponding passive form where the subject of the passive is assigned the role that the active form assigns the indirect object. As the diagram in (2) shows, ditransitive lexical items are members of both 3-1 and TRANSITIVE, inheriting the direct object from TRANSITIVE and the indirect object from 3-1. Such items, by virtue of being transitive, will have a corresponding passive form, but where as usual the thematic role of the passive's subject is the same as that of the active's direct object, not that of its indirect object.

As seen in the diagram in (2), there are two subclasses of TRANSITIVE for the two types of ditransitive verbs in English, exemplified by the two forms of *give* as in (3), which I have for convenience labelled GIVE-1 and GIVE-2.

(3) a. Mary gave John a book. GIVE-1
b. Mary gave a book to John. GIVE-2

I assign verbs like GIVE-1 to the class DITRANS, and verbs like GIVE-2 to the class DITRANS-TO, relating members of the two classes with a lexical rule of Dative Shift which I discuss in the next chapter. Members of each of these classes inherit their three complements (Subject, DObject, and IObject) from the classes TRANSITIVE, 3-1, and (indirectly) INCOMPLETE. But where the DITRANS-TO class simply inherits default information about the IObject from the 3-1 class, the

DITRANS class assigns it the non-default syntax of an NP rather than a TO-PP. A fuller discussion of these two classes, and their interactions with lexical rules and the grammar, must wait until I introduce the lexical rules for passive and dative shift verbs, in the next chapter.

DITRANS	
Superclasses	TRANSITIVE, 3-1
Complements	
IObject-features	(CAT Noun) (COMPLETE +) (CASE Accusative)

DITRANS-TO	
Superclasses	TRANSITIVE, 3-1

The remaining three subclasses in the COMPLEMENTATION hierarchy are also subclasses of CONTROL. Like the ditransitive predicates, members of these classes OBJECT-RAISING, OBJECT-EQUI, and ANOMALOUS-EQUI also have three complements, but the most oblique of these is in each case an XComp. Members of OBJECT-RAISING and OBJECT-EQUI each have a subject (assigned in INCOMPLETE), a direct object (assigned in TRANSITIVE), and a controlled complement (assigned in CONTROL).

OBJECT-RAISING	
Superclasses	RAISING, TRANSITIVE
Complements	
DObject-index	DO-NOT-INHERIT

Note here that members of the OBJECT-RAISING class do inherit a Subject-index, not from the RAISING class (where it is blocked), but from the INTRANSITIVE class via TRANSITIVE, of which OBJECT-RAISING is also a subclass. This means that object-raising predicates can assign a thematic role to their subjects, but not to their direct objects, given that the DObject-index attribute is blocked for members of this class. No such hindrance is placed on any of the complement role assignments for OBJECT-EQUI predicates, so all three complements (subject, direct object, and VP complement) will be assigned thematic roles by each member of the OBJECT-EQUI class.

OBJECT-EQUI	
Superclasses	EQUI, TRANSITIVE
Complements	
DObject-features	(NFORM Norm)

As with the earlier EQUI and RAISING classes, the (NFORM Norm) specification is in complementary distribution with the blocking of an index assignment for the given subcat (here the DObject, not the Subject). This complementarity is not captured in the present analysis, but might be predicted in a semantically-driven treatment of expletive pronouns, as acknowledged earlier.

Similar to the OBJECT-EQUI class is the ANOMALOUS-EQUI class, which is also a subclass of EQUI, but has as its second superclass 3-1 rather than TRANSITIVE. [10]

ANOMALOUS-EQUI	
Superclasses	EQUI, 3-1
IObject-features	(CAT Noun) (COMPLETE +)
	(CASE Accusative)

Members of this class, like those of OBJECT-EQUI, have three complements, but instead of a direct object, they have an indirect object which was assigned in the 3-1 class. I have here suggested that an accusative NP is the default for anomalous-equi indirect objects, but the relative rarity of members of this class makes defending this choice difficult. Examples of this class include entries for verbs like *promise* *make*, and *strike*, illustrated in (4).

(4) a. John promised Mary to sing a song.
b. John made Mary a good father.
c. John struck Mary as a good singer.

Of course, the lexical entries for *make* and *strike* will have to include an idiosyncratic specification of the syntax for its XComp, to override the default (infinitival VP) specification defined in the CONTROL class which *make* belongs to; in the

[10] I am grateful to Diana Roberts for helping me in the proper placement of this class within the COMPLEMENTATION hierarchy.

case of *make*, its XComp must be an NP of some sort, while that of *strike* must be a predicative prepositional phrase introduced by *as*.

The fourth and final subclass of INCOMPLETE is one I have labelled S-NORM, since lexical items belonging to this class take both a verbal complement whose default syntax is sentential, and a subject whose syntax is the inherited default, a non-expletive NP (NFORM Norm). I define this class as follows.

S-NORM	
Superclasses	INCOMPLETE
Complements	XComp
XComp-features	(CAT Verb) (COMPLETE +) (LEXICAL –) (VFORM Finite)
XComp-index	

An example of a lexical item belonging directly to this class is the verb *hope*, as in *John hoped (that) Mary would win.* Another entry in this class is one of the entries for *persuade*, which is also a member of the TRANSITIVE class, inheriting a direct object in addition to the subject and sentential complement obtained through the S-NORM class. This verb appears as the head of the sentence *Mary persuaded John that he snored.*

As seen in the diagram in (2) above, there are two immediate subclasses of this class, one overriding the default information about the Subject subcat, the other overriding the defaults for the XComp. The first of these, S-INF-NORM, retains the default subject specification, but assigns non-default syntactic properties to the sentential complement, making it infinitival, and allowing it to be either a VP or a sentence introduced by the complementizer *for*.

S-INF-NORM	
Superclasses	S-NORM
Complements	
XComp-features	(COMPLETE + –) (VFORM Infinitive) (COMP For)

An example of a lexical entry that is directly a member of this class is the verb *prefer* as in *I prefer (for John) to win.* Here, if the verbal complement is an S, it is not controlled, while if the complement is a VP, it is controlled by the subject.

The other subclass of S-NORM leaves the default specification for the verbal complement, but requires that the subject be the expletive *it*, and blocks inheritance of the Subject-Index attribute, similar to the approach used in the definition of the classes for raising predicates given above.

S-IT	
Superclasses	S-NORM
Complements	
Subject-features	(NFORM It)
Subject-index	DO-NOT-INHERIT

This class includes predicates like *seem* and *obvious*, as in *It seems that John snores* and *It is obvious that John snores.*

Finally, this class has a subclass of its own, which inherits the non-default expletive *it* subject, and also overrides the default specification for the verbal complement, making it infinitival. I do not define this S-INF-IT class as a subclass of both S-INF and S-IT since that would give rise to conflicting information about the subject, defeating the efforts of the S-IT class to specify non-default properties for the subject.

S-INF-IT	
Superclasses	S-IT
Complements	
XComp-features	(COMPLETE + –) (VFORM Infinitive) (COMP For)

Examples of lexical entries in this class include *possible*, as in *It is possible (for Mary) to distress John.*; and the verb *please*, which is also a member of the TRANSITIVE class, inheriting in addition to the Subject and XComp a direct object, which controls the XComp if it is a VP. This can be seen in the sentence *It would please John to win the race.* (More precisely, *please* will be an immediate member of a class, say one named TRANS-INF, which is a subclass of both TRANSITIVE and S-INF-IT, in the same way that OBJECT-EQUI is a subclass of both TRANSITIVE and EQUI.)

I do not offer here any further subclassification of COMPLEMENTATION, but turn instead to a discussion of the other major classification of the full lexicon, a hierarchy I have termed PART-OF-SPEECH.

2.4 Part-of-speech hierarchy

The PART-OF-SPEECH side of the picture is a classification of the lexicon independent of COMPLEMENTATION. The class definition for PART-OF-SPEECH itself simply identifies the types of information which may be provided by subclasses of this class, but does not contain any default value assignments for any of the attributes. The range of possible attributes along this dimension of the lexicon is much like that in the COMPLEMENTATION part of the hierarchy, with the additional possibility of selecting particular adjuncts for some class. I delay until section 3.2 a detailed discussion of adjuncts, where I take up the distinction between adjuncts and optional complements, and I also motivate the decision to have heads select for particular adjuncts, rather than having adjuncts select for their heads.[11]

PART-OF-SPEECH	
Partition-of	WORD-CLASS
Atomic-features	
Category-features	
Complements	
Adjuncts	

2.4.1 Subclasses of PART-OF-SPEECH

As seen in the initial sketch of the hierarchy given in section 2.1, the two subclasses of PART-OF-SPEECH are MAJOR and MINOR, with MINOR subclasses including at least DETERMINER, CONJUNCTION, COMPLEMENTIZER, and the like, while the MAJOR classes include just VERB, NOUN, ADJECTIVE, and PREPOSITION.

MAJOR	
Superclasses	PART-OF-SPEECH
Adjuncts	PP-Adjunct
PP-Adjunct-features	(CAT Preposition) (LEXICAL –) (PFORM Neutral)

[11] In principle, it seems to be possible to maintain that only classes select for particular adjuncts, while lexical items do not; the machinery I present here, however, does not embody that restriction.

The only information specified as the default for all MAJOR lexical items is that all members of this class allow prepositional phrase adjuncts with some default properties; this is clearly true for verbs, nouns, and adjectives, and arguably true for prepositions as well. [Cf. Jackendoff 1973:350ff.] Other possible adjuncts seem to be restricted to subclasses, as indicated below. The (PFORM Neutral) feature-value pair fits into a general scheme which permits subcategorization for particular prepositional phrases, for verbs like *rely* as in *rely on John*, and for passive verbs, which permit an agentive *by* phrase. Such prepositions are marked with a distinctive value for the PFORM attribute which is usually the same as the spelling of the preposition. Ordinary prepositions, then, such as *in* or *under*, have the unmarked (Neutral) value for PFORM, which means phrases they head will not match the requirements for particular complements, but will match those for ordinary adjuncts. Moreover, special prepositions like the passive *by* will not appear freely as adjuncts to non-passive predicates.[12]

Of the four MAJOR subclasses, VERB and NOUN have a rather rich set of subclasses, while ADJECTIVE and PREPOSITION remain relatively unanalyzed here. I present first the subclassification of VERB, then that of NOUN.

2.4.2 Subclasses of VERB

The properties of verbs group along two dimensions, requiring two separate partitionings of VERB, one involving the auxiliary/main verb distinction, and the other involving the form of the verb, (e.g., base, past, participial, etc.). I will call the first partition VERB-TYPE, and the second VERB-FORM. The VERB class itself specifies only the obvious value for CAT.[13]

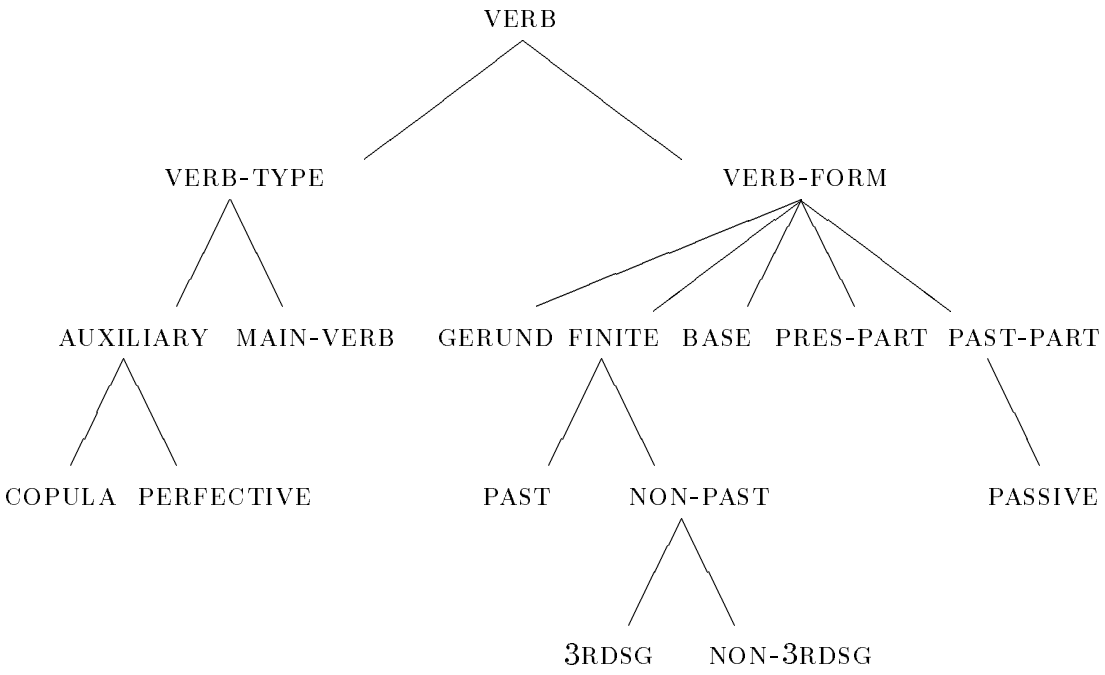
VERB	
Superclasses	MAJOR
Atomic-features	(CAT Verb)

[12] The PFORM feature is introduced in GKPS 1985:23.

[13] As will become clear below, I do not follow Chomsky's (1970) approach to capturing certain cross-categorial generalizations through use of the binary features N and V, instead capturing such generalizations by means of lexical rules. Accordingly, the values of the attribute CAT include the non-decomposable values Verb, Noun, Adjective, and Preposition.

An extension of the graph in section 2.1 which represents the subclasses of VERB appears in (1):

(1) Subclasses of VERB



2.4.3 Verb Types

The distinction between auxiliary and main verbs rests on several observed differences in distribution, the most obvious being that auxiliary verbs (usually) can appear as the heads of inverted sentences, while main verbs never can. This difference, illustrated in (2), is encoded using the feature INVERTED, which appears with a negative value in the MAIN-VERB subclass of VERB-TYPE below, and is unmarked in the AUXILIARY subclass.[14]

(2) a. Has Ferdinand bought an island?
b. *Bought Ferdinand an island?

[14]Following GKPS, I assume that the lexical entry for the first person singular *aren't* as in *Aren't I clever?* will have the idiosyncratic specification (INVERTED +), since this auxiliary can only appear in inverted sentences.

VERB-TYPE	
Partition-of	VERB

AUXILIARY	
Superclasses	VERB-TYPE, RAISING
Complements	
XComp-features	(VFORM Base)

The AUXILIARY class, with fixed membership, includes auxiliary verbs of three types: modals, the forms of perfective *have*, and the forms of *be*. This class definition assigns the base form as the default syntactic form of the VP complement for auxiliaries, a default which is overridden both by the perfective and by the copula, as well as by the individual lexical item for *ought*. Since all auxiliary verbs are arguably subject-raising predicates, the definition of this class includes RAISING as one of the two superclasses. Members of this class are unmarked for the feature INVERTED, signifying that as a default they may appear as heads of either inverted or non-inverted sentences.

COPULA	
Superclasses	AUXILIARY
Complements	
XComp-features	(PREDICATIVE +)

I will not go into detail here about the subclasses of COPULA, though the idiosyncracies of members of this class can provide a thorough exercise of the inheritance mechanisms being introduced. I only note the introduction here of the binary feature PREDICATIVE, used to distinguish predicative from non-predicative phrases, a distinction needed both here and for the description of possible nominal modifiers usually referred to as reduced relatives.

The only members of the PERFECTIVE class are the various forms of the auxiliary *have*, which require that the XComp be headed by a past-participle verb.

PERFECTIVE	
Complements	
XComp-features	(VFORM Past-Part)

Except for the above closed classes of auxiliary verbs, all other verbs in English are main verbs, and therefore cannot appear as heads of inverted sentences, but can appear as heads of non-inverted sentences. Another difference between the two classes is the default form of the VP complement, if there is one; with few exceptions, this complement is infinitival for main verbs, as specified in the following class definition.

MAIN-VERB	
Superclasses	VERB
Atomic-features	(INVERTED –)
Complements	
XComp-features	(VFORM Infinitive)

I do not provide here any further subclassification of main verbs, since the purpose of this class is to distinguish main from auxiliary verbs. It may well prove to be the case that borderline verbs like *better* will render the main/auxiliary split inadequate; it will suffice for my purposes here.

2.4.4 Verb Forms

The paradigm for an ordinary verb like *laugh* contains several syntactically distinct forms, as suggested by the names of the subclasses under VERB-FORM in the graph in (1) above. Not all of these forms have a corresponding morphological distinction; for example, the base form and the present plural (non-third-person-singular) form of all verbs except *be* are the same in English:

(4) a. Does John laugh at his own jokes?
b. People laugh at anything.

Similarly, the passive and the past participle forms of all verbs with passives have identical morphology, yet they must be kept distinct because of their very different subcategorization and semantic properties.

Each of the subclasses of VERB-FORM will introduce a distinct value for the feature VFORM, and some will specify additional properties, both syntactic and semantic. Each subclass also specifies whether it is predicative or non-predicative;

see the discussion of the COPULA class above for introduction of this feature. These subclasses play an important role in expressing regularities in the inflectional paradigms of verbs in English, though the explanation of their interdependency with lexical rules must wait until the following chapter.[15]

VERB-FORM	
Partition-of	VERB

BASE	
Superclasses	VERB-FORM
Atomic-features	(VFORM Base) (PREDICATIVE –)

It is to this BASE subclass of VERB-FORM that the unmarked lexical entries for verbs belong.

FINITE	
Superclasses	VERB-FORM
Atomic-features	(VFORM Finite) (PREDICATIVE –)
Complements	
Subject-features	(CASE Nominative)

Finite verbs specify that their subjects must be nominative rather than accusative (though this is relevant only for pronouns). Of the two subclasses of FINITE, the class for past tense verbs does not need further subclassification, but the class for non-past verbs must be divided into third-singular and non-third-singular subclasses. I do not add classes for the other combinations of person and number, since the only members of such classes would be irregular forms of the copula, and these have to be entered separately in the lexicon anyway, so the classes would serve no purpose in characterizing English.

[15] I have not included an INFINITIVE subclass of VERB-FORM, since it would be a class with exactly one member, the verb *to* of *to work* (following Pullum 1982). Since the sole member of this class participates in no lexical rules and cannot hope to recruit additional class members, the extravagance of a separate form class seems unjustified. The word *to*, then, will be entered in the lexicon without having membership in a subclass of VERB-FORM, so will itself specify the features (VFORM Infinitive) and (PREDICATIVE –). Since it is a member of the AUXILIARY class, *to* will inherit the property that the features of its complement must include (VFORM Base).

PAST	
Superclasses	FINITE

NON-PAST	
Superclasses	FINITE

These two subclasses of FINITE should include the relevant semantic distinction between past and non-past verbs, but I leave the interesting question of how to represent that distinction for further work.

THIRD-SINGULAR	
Superclasses	NON-PAST
Complements	
Subject-features	(AGREEMENT Third-Singular)

NON-THIRD-SINGULAR	
Superclasses	NON-PAST
Complements	
Subject-features	(AGREEMENT –Third-Singular)

I introduce here a simple notational convention for prohibited feature values, in giving the value of the feature AGREEMENT. Since I assume that all attributes in a class description have a finite list of possible values, the negation of a value is simply shorthand for the disjunctive specification of all the other values. So in this case the value of *Subject-features* could be equivalently expressed as the disjunctive (AGREEMENT 1st-Sg 2nd-Sg Plural)). This disjunctive specification is in turn to be interpreted as requiring that the subject of a verb which is a member of this NON-THIRD-SINGULAR class will have as its value for AGREEMENT one of these three values (1st-Sg 2nd-Sg Plural). How this requirement is enforced remains a topic reserved for the section on feature unification, in section 4.1. The disjunction of all values of AGREEMENT except Third-Singular is needed in the above definition, since each pronoun must be marked with just one of the possible AGREEMENT values, in order to co-occur with the proper form of the verb *be*, yet all pronouns except the third-person-singular ones can occur as subjects of non-third-singular verbs.

PRES-PARTICIPLE	
Superclasses	VERB-FORM
Atomic-features	(VFORM Pres-Participle) (PREDICATIVE +)

Members of the PRESENT-PARTICIPLE class head verb phrases which appear as complements to the verb *be*, and as modifiers for common nouns. They do not include verbal gerunds, as argued in Gazdar, Pullum, and Sag 1982, based on contrasts like that between **Kim is having left* and *Having read that book, he returned it.*

VERBAL-GERUND	
Superclasses	VERB-FORM
Atomic-features	(VFORM Gerund) (PREDICATIVE +)

Members of the VERBAL-GERUND class must also be distinguished from nominal gerunds, as argued by Wasow and Roeper 1972, since the latter are simply well-behaved deverbal nouns, belonging to the COMMON-NOUN class. Phrases headed by verbal gerunds still have the internal structure of VPs, and hence those heads must be classed as verbs, not as nouns; the phrase structure rules will license the appearance of verbal gerunds as noun phrases.

PAST-PARTICIPLE	
Superclasses	VERB-FORM
Atomic-features	(VFORM Past-Participle) (PREDICATIVE –)

To the PAST-PARTICIPLE class belong verbs which head VPs that appear as complements to the perfective *have*. In addition, this class includes the subclass of passive participles, which have the same morphological form as past participles, but distinct syntactic and subcategorization properties.

PASSIVE	
Superclasses	PAST-PARTICIPLE
Atomic-features	(VFORM Passive) (PREDICATIVE +)
Complements	(PP-By)
PP-By-features	(CAT Preposition) (COMPLETE –) (PFORM By)
PP-By-index	

The class to which all passive forms of verbs belong is here taken to be one of the possible verb forms, with the work of relating members of this class to their corresponding active forms handled by lexical rule, as discussed in the next chapter. What this class definition provides, in addition to the distinguishing syntactic feature specifications, is the information that as a default, passive verbs allow an optional PP complement introduced by the preposition *by*, with the thematic role of the PP-By determined by the passive lexical rule.[16]

2.4.5 Subclasses of NOUN

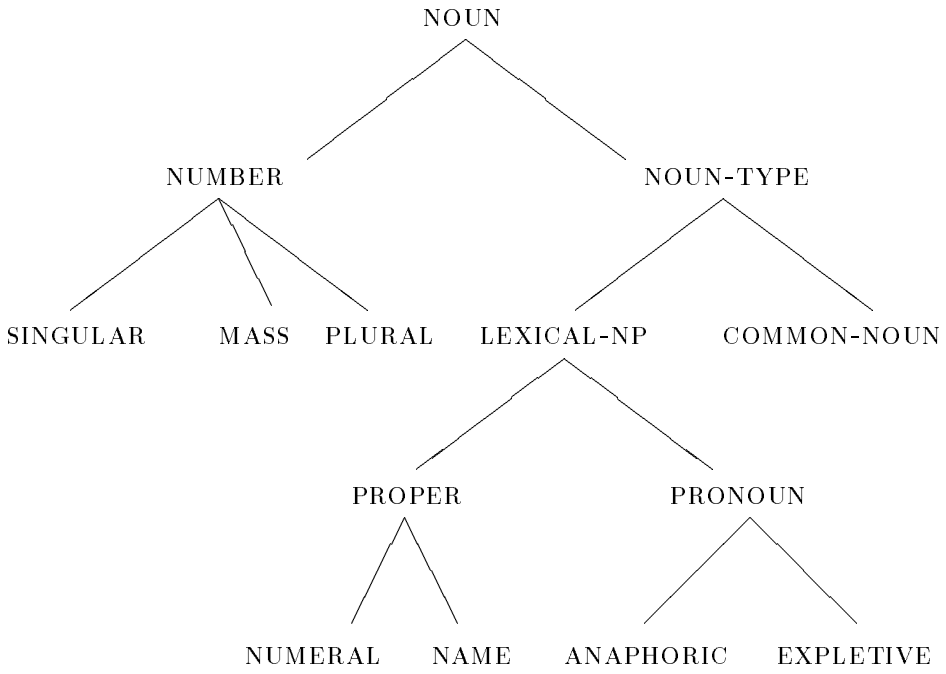
Like verbs, nouns also divide into two essentially independent partitions, one involving number and the other involving noun type. Most of the properties of nouns will be found in one or the other of these two subclassifications; the only properties given here in the definition of NOUN are the obvious category feature, and the default value for the feature NFORM, which has as other possible values *There* and *It*, to distinguish expletive pronouns.

NOUN	
Superclasses	MAJOR
Atomic-features	(CAT Noun) (NFORM Norm)

An extension of the hierarchy graph which presents the subclasses of NOUN appears in (5):

[16] It is worth noting here that the informal parenthesis notation used to indicate the optionality of the PP-By should more properly be represented as another attribute of each subcat, which I will assume is a Status attribute alongside the Features and Index attributes already illustrated for subcats. The possible values of this attribute would be at least Obligatory and Optional, but the attribute might also be used to distinguish several kinds of optionality. Following an approach being developed by John Nerbonne, and by Pollard and Sag 1987, one might distinguish the class of verbs like *kick* from that of verbs like *eat* by marking the optional direct object as having a distinct status for each of the two, in order to capture the different entailment properties of the two types (where *eating* entails that something gets eaten, while *kicking* does not entail that anything gets kicked). I do not pursue this idea here, though it holds promise.

(5) Subclasses of NOUN



2.4.6 Noun types

I take up first the partitioning of nouns along the dimension of noun type. English distinguishes two broad types of lexical nouns, often labelled *common* and *proper*, where the distinction is in part a semantic one, with common nouns naming classes of entities in the world, and proper nouns naming individuals. This distinction is also characterizable in terms of complementation, with common nouns permitting or requiring one obligatory argument, and proper nouns permitting none. If this were the only difference, a separate noun-type classification would be unnecessary, since the COMPLEMENTATION classification would suffice, in the COMPLETE/INCOMPLETE distinction. However, common nouns allow a range of adjuncts not permitted with proper nouns, including adjectives and restrictive relative clauses; the converse also seems to be true, with proper nouns allowing at least one type of adjunct not possible for common nouns, namely titles, as in *professor*

Smith or *serial number 54321*. Selection for adjuncts is not covered by making reference to the COMPLEMENTATION hierarchy; hence the need for a NOUN-TYPE partitioning of the class NOUN.

Further distinctions need to be made in noun types, between proper nouns and pronouns, in order to capture the familiar differences of distribution in linear order for verb-particle constructions, illustrated in (6).

(6) a. John looked the answer up.
b. John looked up the answer.
c. John looked it up.
d. *John looked up it.

If this difference in acceptability is to be given a syntactic account, then some feature-value distinction is needed which permits the rules which admit the sentences in (6) to be able to distinguish proper nouns from pronouns in imposing ordering constraints. A second motivation for making a syntactic distinction between proper nouns and pronouns comes from contrasts like the following, where (7b-c) are unacceptable even if put in some context where the referent for the pronoun is given explicitly.

(7) a. John read the Chomsky book.
b. *John read the he book.
c. *John read the him book.

To make the necessary distinctions among types of nouns, I introduce the feature NTYPE, with values *Common*, *Proper*, and *Pronoun*. The classes for proper nouns and pronouns are then grouped into a class of lexical NPs. In addition to being a subclass of NOUN-TYPE, the LEXICAL-NP class is a subclass of COMPLETE, which ensures that all members of this class will have an empty *Complements* list, a defining property of noun phrases.[17]

NOUN-TYPE	
Partition-of	NOUN

LEXICAL-NP	
Superclasses	NOUN-TYPE, COMPLETE

[17]For the sake of exposition, I did not indicate this second link for LEXICAL-NP in the diagram in (5); further discussion of multiple-class membership is found in section 3.1.

The class of lexical NPs has two subclasses here: PROPER and PRONOUN, as described below. Proper nouns are further divided into two subclasses, one for ordinary names, and one for numerals like *two*; one distinction between these two classes is that only members of the NUMERAL class are related to morphologically identical forms that appear as numerical determiners, as in *one book is missing* but not *John book is missing*. If the relation between *two* the name of a number and *two* the determiner is to be captured via lexical rule, then numbers must be taken as a distinct class so the lexical rule that captures that relation can be restricted to apply to the right subset of lexical NPs (namely names of numbers).

PROPER	
Superclasses	LEXICAL-NP
Atomic-features	(NTYPE Proper)
Adjuncts	Title
Title-features	(CAT Noun) (COMPLETE –)

NUMERAL	
Superclasses	PROPER

This class clearly contains lexical entries for the numerals *one*, *two*, and so on, but it is less clear whether more complex names of numerals like *two hundred fifty-three* are lexical items in this class, or instead phrasal constituents constructed by phrase structure rule. Issues relevant for deciding this case are treated in the discussion of the relative domains of lexical rules vs. phrase structure rules, in Chapter 4.

NAME	
Superclasses	PROPER

In the NAME class are all proper nouns used to name individuals. This is perhaps the most open of all the word classes, since new names for new and existing entities are created constantly.

The other subclass of lexical NPs, for pronouns, is divided into at least two subclasses: those which are anaphoric and those which are expletive, including *it* and *there*.

PRONOUN	
Superclasses	LEXICAL-NP
Atomic-features	(NTYPE Pronoun)

ANAPHORIC	
Superclasses	PRONOUN

Anaphoric pronouns submit to a further subdivision, distinguishing reflexive from non-reflexive forms.

REFLEXIVE	
Superclasses	ANAPHORIC

NON-REFLEXIVE	
Superclasses	ANAPHORIC

The other subclass of pronouns, for expletive pronouns, includes at least the lexical entries for *it* and *there*, where each of these lexical items specifies a unique value for the feature NFORM, in order to exclude sentences like the one in (8).

(8) *There kicked the ball.

EXPLETIVE	
Superclasses	PRONOUN

In addition to PROPER, the other basic NOUN-TYPE subclass is COMMON-NOUN, which has the expected value for the feature NTYPE, and also specifies several adjuncts possible as modifiers for common nouns. Note that no specification about agreement of relative clauses with their head is made here; those specifications appear in the various subclasses of NUMBER, and will be merged with the features given in the COMMON-NOUN class, in a manner detailed in the discussion of inheritance, in section 3.1.

COMMON-NOUN	
Superclasses	NOUN-TYPE
Atomic-features	(NTYPE Common)
Complements	Determiner/Subject
Subject-features	(CAT Noun) (COMPLETE +) (CASE Genitive)
Determiner-features	(CAT Determiner)
Adjuncts	Adjective Compound-Noun Full-Rel Thatless-Rel Reduced-Rel
Adjective-features	(CAT Adjective) (LEXICAL +) (PREDICATIVE –)
Cmpnd-Noun-features	(CAT Noun) (NTYPE –Pronoun)
Full-Rel-features	(CAT Verb) (COMPLETE +) (VFORM Finite) (REL [])
Thatless-Rel-features	(CAT Verb) (COMPLETE +) (VFORM Finite) (SLASH ((CAT Noun) (COMPLETE +)))
Reduced-Rel-features	(CAT Verb Adjective Preposition) (COMPLETE –) (LEXICAL –) (PREDICATIVE +)

The first curiosity in this definition is the disjunctive specification in the Complements attribute, which I annotate with a slash, in order not to confuse disjunction with optionality. Common nouns may occur with either a determiner or a possessive NP subject; this complementarity is made explicit by the use of the disjunctive specification, which is completely analogous to the ordinary disjunction of feature values as described above.

The specification for features of the reduced relative clause adjunct also makes use of the disjunction of possible feature values. Again, the specification (CAT Verb Adjective Preposition) is to be interpreted as a requirement that the category which fits this adjunct description must have as the value of its CAT feature one of the three values Verb, Adjective, or Preposition. I say more about the use of this disjunctive notation in the section on feature unification, in chapter 4.

In the specification for Full-Rel-features appears the attribute-value pair (REL []), which is to be interpreted as a specification for any non-null set of features: this

adjunct must contain a relative pronoun of any sort.

There may well need to be further subclasses of COMMON-NOUN, possibly distinguishing between root nouns and deverbal nouns, and perhaps even carving up deverbal nouns into agentives and other nominalizations. As usual, the motivation for these further subdivisions must arise out of demonstrable differences in syntactic distribution or behavior; I leave the further subclassification of common nouns unspecified here.

2.4.7 Noun Number classes

The partitioning of the NOUN class by number requires three classes to make the traditional singular/plural/mass distinction in English, as illustrated in (9-11).

(9) a. A book is missing.
b. *Several book are missing.
c. *Book is missing.
d. *Book are missing.

(10) a. *A children is missing.
b. Several children are missing.
c. *Children is missing.
d. Children are missing.

(11) a. *A money is missing.
b. *Several money are missing.
c. Money is missing.
d. *Money are missing.

The agreement picture is complicated slightly by the remnants of person agreement (first, second, third) which are found in the pronominal system of English, and which are further reflected in the idiosyncratic inflectional paradigm for the verb *be*. To account for these manifestations of person while still acknowledging the fact that for all the other English verbs only a binary distinction is needed for agreement, I make use of the single feature AGREEMENT for both number and person; the usual values are *Third-Singular*, *Plural*, and *Mass*; others may be introduced as needed.

NUMBER	
Partition-of	NOUN

SINGULAR	
Superclasses	NUMBER
Atomic-features	(AGREEMENT Third-Singular)
Complements	
Determiner-features	(AGREEMENT Third-Singular)
Adjuncts	
Full-Rel-features	(REL ((AGREEMENT Third-Singular)))
Thatless-Rel-features	(SLASH ((AGREEMENT Third-Singular)))

PLURAL	
Superclasses	NUMBER
Atomic-features	(AGREEMENT Plural)
Complements	
Determiner-features	(AGREEMENT Plural)
Adjuncts	
Full-Rel-features	(REL ((AGREEMENT Plural)))
Thatless-Rel-features	(SLASH ((AGREEMENT Plural)))

MASS	
Superclasses	NUMBER
Atomic-features	(AGREEMENT Mass)
Complements	
Determiner-features	(AGREEMENT Mass)
Adjuncts	
Full-Rel-features	(REL ((AGREEMENT Mass)))
Thatless-Rel-features	(SLASH ((AGREEMENT Mass)))

Since the notation used here can be misleading, let me remind the reader that each of these subclasses of NUMBER specifies default properties that relative clause adjuncts will have if they are selected for, but none of these subclasses select for relative clauses. So members of the COMMON-NOUN class which are also members of the SINGULAR class will select for relative clauses, and specify that they must show third-person-singular agreement (given the way the feature principles work, as discussed in GKPS 1985:75ff). In contrast, members of the PROPER noun

class which are also members of the SINGULAR class do not select for (restrictive) relative clauses, so the default specifications in SINGULAR about relative clauses are ignored.[18]

2.4.8 The ADJECTIVE class

I will not offer a detailed discussion of the ADJECTIVE subclass of MAJOR here, since the structure of the VERB and NOUN classes will provide an adequate basis for discussion of multiple-class membership and inheritance in the sections to follow, and will also suffice for the presentation of lexical rules in the next chapter. I merely introduce the basic class, and draw the familiar predicative/attributive distinction, then show how this will interact with the LEXICAL attribute.

ADJECTIVE	
Superclasses	MAJOR
Atomic-features	(CAT Adjective)

Adjectives can appear in at least three positions in a sentence: as pre-nominal modifiers, as post-nominal modifiers, and as complements of verbs like *be* and *become*. Adjectives like *mere* can only appear in pre-nominal position, while those like *afraid* cannot appear pre-nominally, but can appear as head of a post-nominal adjective phrase, or as a complement to verbs like *be*. Many adjectives can appear in any of the three contexts, provided that they are lexical in pre-nominal position, and phrasal in post-nominal position.

I will assume that verbs like *be* and *become* subcategorize for a predicative complement (one marked (PREDICATIVE +)), and that common nouns introduce two distinct adjuncts, as seen above; the *Adjective* adjunct is (PREDICATIVE –) and (LEXICAL +), while the post-nominal adjective phrase is subsumed in the *Reduced-relative* adjunct, which is predicative and non-lexical. Then the lexical entry for *mere*

[18]The reader will have noticed that the same agreement specification appears four time in each of the three subclasses of NUMBER for nouns, suggesting that a generalization is being missed. Though I do not present here a specific proposal for eliminating the redundancy, it seems that the spirit of the Control Agreement Principle of GKPS 1985:83ff could be preserved in terms appropriate to this word class hierarchy. An alternative proposal developed by Pollard and Sag treats this kind of agreement as semantic, not syntactic.

will either belong to an ATTRIBUTIVE subclass of ADJECTIVE, or simply be lexically marked as (PREDICATIVE –), if there is no independent justification for such an ATTRIBUTIVE class. Likewise, the entry for *afraid* will be (PREDICATIVE +), and that for an adjective like *eager* will be unmarked for the PREDICATIVE feature. This produces the necessary set of distinctions to account for the judgments in (12-14), provided that the phrase structure rules and linear precedence constraints ensure that lexical adjuncts occur before the head, while non-lexical adjuncts occur after the head. See Chapter 4 for presentation of the PS rules and LP constraints.

(12) a. A mere lieutenant should not interrupt.
b. *That lieutenant is mere.

(13) a. *An afraid man should not train lions.
b. John is afraid.
c. John is afraid of lions.
d. *The man afraid refused to enter the cage.
e. The man afraid of lions refused to enter the cage.

(14) a. A fearful man should not train lions.
b. John is fearful.
c. John is fearful of lions.
d. *The man fearful refused to enter the cage.
e. The man fearful of lions refused to enter the cage.

2.4.9 The PREPOSITION class

The last subclass of MAJOR is for prepositions, which I will not classify in detail.

PREPOSITION	
Superclasses	MAJOR
Atomic-features	(CAT Preposition)

The only subclass of PREPOSITION that I describe here is the class of lexical adverbs, which I take to be of the same category as prepositional phrases, given the overlap in distribution among adverbs and PPs. Since a lexical preposition is distinguished from a prepositional phrase by the value assigned to the feature LEXICAL, I define the ADVERB class in what may appear to be a counterintuitive

fashion, making its members (LEXICAL -), in order to capture the syntactic similarity between adverbs and prepositional phrases. I say nothing else about them here.

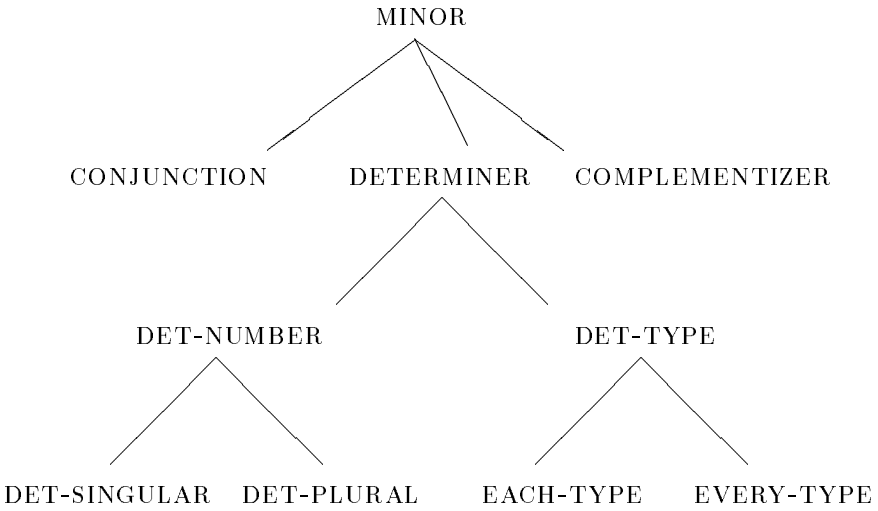
ADVERB	
Superclasses	PREPOSITION
Atomic-features	(LEXICAL –)

2.4.10 Subclasses of MINOR

Aside from the four MAJOR subclasses, there are a number of classes of lexical items grouped together as subclasses of the class MINOR, which is the other immediate subclass of PART-OF-SPEECH. The existence of the MINOR class itself is not given any independent justification here, but serves as a convenient node for grouping these non-major classes. I include here only three of these MINOR subclasses, for determiners, conjunctions, and complementizers, and I will sketch some further detail only for determiners. The extension of the hierarchy sketch for the MINOR class hierarchy is given in (15).

MINOR	
Superclasses	PART-OF-SPEECH

(15) Subclasses of MINOR



Determiners in English can show number agreement, and also divide along an orthogonal dimension that involves the partitive construction, so at least two partitions of this class are needed.

DETERMINER	
Superclasses	MINOR
Atomic-features	(CAT Determiner)

The number distinction among determiners is only a binary distinction, since singular determiners agree with both singular and mass (third-person) nouns. Review the definitions of SINGULAR, MASS, and PLURAL subclasses of the NUMBER subclass of NOUN, to see the corresponding specifications for determiners of one or the other of the two classes defined here.

DET-NUMBER	
Partition-of	DETERMINER

DET-SINGULAR	
Superclasses	DET-NUMBER
Atomic-features	(AGREEMENT Third-Singular Mass)

DET-PLURAL	
Superclasses	DET-NUMBER
Atomic-features	(AGREEMENT Plural)

A second perspective in which to classify determiners involves their distribution with respect to the partitive construction, among others. I illustrate the contrast in (16-17).

(16) a. Each runner finished.
b. Each one of the runners finished.
c. Each of the runners finished.

(17) a. Every runner finished.
b. Every one of the runners finished.
c. *Every of the runners finished.

As with several of the distinctions I have drawn in presenting this hierarchy, it may be that the difference need not be encoded syntactically, if an adequate

semantic account can be provided, but determiners clearly pattern either like *each* or in the more restricted fashion of *every*. To express this difference, I introduce a second subclass of DETERMINER called DET-TYPE, and name its two subclasses after their exemplar members. In lieu of a semantic account of the distinction, I employ the feature DTYPE to distinguish members of the two subclasses.[19]

DET-TYPE	
Partition-of	DETERMINER

EACH-TYPE	
Superclasses	DET-TYPE
Atomic-features	(DTYPE Each)

EVERY-TYPE	
Superclasses	DET-TYPE
Atomic-features	(DTYPE Every)

Two other subclasses of MINOR given here are those for conjunctions and for complementizers; I do not provide detailed descriptions of either, but will refer to these classes in subsequent discussions, so include them for completeness.

CONJUNCTION	
Superclasses	MINOR
Atomic-features	(CAT Conjunction)

COMPLEMENTIZER	
Superclasses	MINOR
Atomic-features	(CAT Complementizer)

[19] But see Ladusaw 1982 for a promising semantic treatment of this distinction.

Chapter 3

Inheritance and Subcategorization

In order for the hierarchy of word classes that I introduced in the previous chapter to be useful, the mechanism for passing information from superclass to subclass, or from class to member, needs to be made explicit. This flow of information in a class hierarchy is commonly termed "inheritance", since members of a class exhibit the properties defined for their class and its parent classes; these class members are said to "inherit" the attribute values defined for classes nearer to the root of the hierarchy.

Much of the predictable information about a lexical entry concerns its subcategorization properties, which can be stated using this same inheritance mechanism, as illustrated in the previous chapter. This use of inheritance for describing both complements and adjuncts depends on several assumptions about the form and content of subcat specifications, assumptions which are reflected in the class definitions already given, and which merit discussion here.

Structure of the chapter

Section 3.1 provides a detailed introduction to the inheritance mechanism I employ, identifying the representation issues that arise in a hierarchy of the sort presented

here. In section 3.2, I introduce a basic constraint on subcat specifications in lexical entries, then give motivation for some of the distinctions made in the types of subcats that an entry may inherit.

3.1 Inheritance

The notion of inheritance that I present here, while simplified, is closely related to the kind of inheritance developed for general-purpose knowledge representation. Indeed, the multiple-parent structure of the lexicon presented above, the use of default values, and the distinctions among different types of attributes are all examples of ideas borrowed from this field of research. While I do not make further explicit reference to parallels between the structure of the lexicon used here and the structure of more standard knowledge bases that have been developed, there is much held in common, suggesting that as difficulties arise in the use of these tools for representation, solutions may be found in the work in general knowledge representation.[1]

As illustrated in the previous chapter, a word class or lexical entry may inherit from one or more parent classes. In the simple cases where a class or entry belongs to just one superclass, the rule for how values of an attribute are assigned in the word class hierarchy is quite straightforward. The two alternatives to be considered depend on whether a given attribute permits only one value, or multiple values. Individual features such as VFORM are single-valued, while attributes like *Complements* permit more than one value, as seen in the previous section. (I did not make this distinction explicit in the preceding section, attribute by attribute, since for most there should be no danger of confusion.)

(1) **Inheritance of Values**

> The value assigned to a particular word class (or member) W for a given attribute is determined as follows:
>
> a. For a single-valued attribute, the assigned value is either introduced directly in W, or is the one introduced in the most specific class to which W belongs. If there is no value introduced anywhere in the linked classes between W and the root WORD-CLASS, inclusively, no value is assigned to W for that attribute.

[1]For an introduction to this field of research via the description of one particular scheme for knowledge representation, see Rosenberg 1983 and references cited there.

b. For a multiple-valued attribute, the assigned values are the members of the set consisting of all distinct values introduced for that attribute in W and in any of the classes linking W with the root WORD-CLASS, inclusively.

For example, if W specifies a value V1 for some multiple-valued attribute, and if W's parent class specifies values V2 and V3, and if no other classes up the hierarchy specify any value, then W's inherited values for that attribute consist of V1, V2, and V3.

In cases where a class or member belongs to more than one superclass, the picture might be more complicated, since each of two immediate superclasses might specify a different value for the same single-valued attribute. One way to address the potential conflict would be to define another rule of inheritance to take account of multiple parents, a rule which for each attribute assigns priority to some one of the parent classes. An alternative approach would be to constrain the hierarchy in such a way that each single-valued attribute of some given class or member was assigned a value by at most one of the immediate superclasses (or its parents, so conflicting values could not occur.

As a working hypothesis I adopt the latter strategy, for three reasons: first, because it imposes a stronger constraint on the hierarchy, and is thus more readily falsifiable; second, because the hierarchy presented here is consistent with this hypothesis; and third, because it is not clear that there is any principled basis for assigning a ranking among the superclasses of each class or lexical entry. What I do not provide here is a precise formulation of the desired constraint on value assignment for single-valued attributes in a class hierarchy that permits multiple immediate superclasses for a class. I will assume that for present purposes the intuition is clear enough.[2]

[2]If this strong restriction on word class definitions proves to be untenable, a convention like the following would have to be adopted instead, to govern attribute-value conflicts with inheritance.

(i) Multiple-parent inheritance

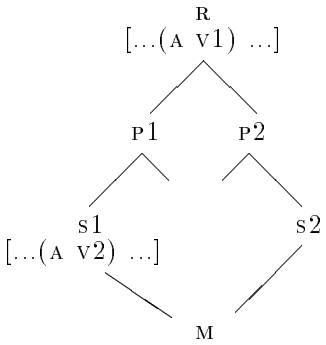
a. For a single-valued attribute, give priority to the parent class mentioned first in the list of superclasses specified in a class or member. Given this convention for multiple parents (which applies transitively up the hierarchy), the assignment of a

value for the relevant attribute for the class or member W is then made according to the principle in (1a) above.

b. For a multiple-valued attribute, the convention is essentially the same as in (1b), though the walk up the hierarchy must be performed for each superclass specified at each step. Again, all values are collected for the relevant attribute, and all are assigned as the values of that attribute for the class or member W.

One interesting variant of this potential conflict in inheritance of values can arise because of the distinction I have drawn in the two kinds of links which can join one class with another. Inheritance proceeds as described in (1) along what I have termed *subset* links (signified by dashed lines in the diagrams of the preceding section), but something more may have to be said about what happens across *perspective* links (the dotted lines in the diagrams). No example of this type of potential conflict arises given the hierarchy presented in the preceding section, but there is again nothing that excludes such a conflict. I present the problem abstractly with a simple scenario, which may be more easily grasped with the aid of the following diagram:

(ii) Hypothetical hierarchy



Imagine that a default value V1 is specified for some single-valued attribute A in the root class R of a hierarchy, and that R is partitioned two ways, with the partitions labelled P1 and P2. Now if some subclass S1 of P1 overrides the default V1 for A with a new value V2, then trouble rears its head, because a member M of S1 is also going to be a member of some subclass of P2, since P1 and P2 are each (exhaustive) partitions of the root R. Hence M will inherit the more specific default value V2 from within the P1 hierarchy, but since M also inherits from the P2 hierarchy, in which no overriding of values for A is found, it looks like M will also inherit the original value V1.

Since the conflict just described is due to my use of these perspective links, and may not involve the more common subset links, a more principled resolution suggests itself than the general but arbitrary one proposed in (ia): unlike subset links which support inheritance for all attributes, it may be that perspective links should be disjoint in the set of attributes for which they support inheritance. The intuition here should be fairly clear: for a given way of partitioning a class, only some proper subset of the attributes might be relevant, and if there are two independent ways of partitioning that class, the second partition might well be expected to refer to a subset of attributes which is disjoint from the set of those needed in the first partition; otherwise, the two partitions would not be independent.

Lacking evidence from the hierarchy at hand to guide exploration of these proposals for resolution (or elimination) of conflicts, I leave the questions open, but note that L. Karttunen has developed a strategy for dealing with conflicts arising in multiple inheritance of this kind.

3.1.1 An illustration of inheritance

To illustrate the use of the conventions in (1), I return to the lexical entry for *tried*, which I introduced in section 2.1. I suggested there that the entry for *tried* should not include any specifications that could be predicted from knowledge of the word classes to which *tried* belongs. For illustrative purposes, I suggested that *tried* might be assigned to two classes labelled the EQUI-VERB-CLASS and the PAST-CLASS. Now that we have the more carefully constructed hierarchy of classes developed in section 2.2, we can return to the question of which classes *tried* is a member of. A short inspection of the classes presented should reveal that the three most specific classes are MAIN-VERB, PAST, and EQUI. So the minimal lexical entry for *tried* will be as given in (2).

(2) Minimal lexical entry for *tried*

TRY-1-PAST	
Superclasses	MAIN-VERB, PAST, EQUI
Spelling	"tried"
Phonology	/traid/
Semantics	(PAST (TRY agent:X prop:Y))
Complements	
Subject-index	X
XComp-index	Y

To show how the inheritance of attribute values works, I repeat the fully specified entry for *tried*, but now labelling each attribute value with the class in which that value was specified.

(3) Annotated, fully-specified lexical entry for *tried*

TRY-1-PAST		Source class
Superclasses	MAIN-VERB	(Local)
	PAST	(Local)
	EQUI	(Local)
Spelling	"tried"	(Local)
Phonology	/traid/	(Local)
Semantics	(PAST (TRY agent:X prop:Y))	(Local)
Atomic-features	(CAT Verb)	VERB
	(VFORM Finite)	FINITE
	(INVERTED –)	MAIN-VERB
	(LEXICAL +)	WORD-CLASS
	(COMPLETE –)	INCOMPLETE
	(PREDICATIVE –)	FINITE
Category-features		
Complements	Subject	INCOMPLETE
	XComp	CONTROL
Subject-features	(CAT Noun)	INCOMPLETE
	(COMPLETE +)	INCOMPLETE
	(CASE Nominative)	FINITE
Subject-index	X	(Local)
XComp-features	(CAT Verb)	CONTROL
	(VFORM Infinitive)	MAIN-VERB
	(LEXICAL –)	CONTROL
	(COMPLETE –)	CONTROL
XComp-index	Y	(Local)
Adjuncts	PP-Adjunct	MAJOR
PP-Adjunct-feats	(CAT Preposition)	MAJOR
	(COMPLETE –)	MAJOR
	(PFORM Neutral)	MAJOR
LP-constraints	...	COMPLEMENTATION

Though this example is relatively simple, values were contributed by nine distinct word classes in the hierarchy to compose the entry for *tried*. Illustrated here are inherited values for both types of attributes, single-valued (like INVERTED) and multiple-valued (like Complements). Note carefully in this regard that the attribute

Atomic-features is multiple-valued, with its possible values themselves attribute-value pairs, where the latter attributes are all single-valued. The inheritance conventions behave the same for each attribute, recursively within the structure of a lexical entry or word class.

3.1.2 Overriding of default values

What the annotated entry for *tried* does not illustrate is the inheritance mechanism which allows the overriding of a default value with another, more narrowly applicable value. For this purpose I provide as a second example the entry for the auxiliary verb *has*, as in *John has arrived.*

(4) Annotated, fully-specified lexical entry for *has*

HAVE-PRES-3RDSG		Source class
Superclasses	PERFECTIVE	(Local)
	THIRD-SINGULAR	(Local)
	RAISING	(Local)
Spelling	"has"	(Local)
Phonology	/hAz/	(Local)
Semantics	...	(Local)
Atomic-features	(CAT Verb)	VERB
	(VFORM Finite)	FINITE
	(INVERTED +)	AUXILIARY
	(LEXICAL +)	WORD-CLASS
	(COMPLETE –)	INCOMPLETE
	(PREDICATIVE –)	FINITE
Category-features		
Complements	Subject	INCOMPLETE
	XComp	CONTROL
Subject-features	(CAT Noun)	INCOMPLETE
	(COMPLETE +)	INCOMPLETE
	(CASE Nominative)	FINITE
	(AGREEMENT 3rdSg)	THIRD-SINGULAR
XComp-features	(CAT Verb)	CONTROL
	(VFORM Past-Part)	PERFECTIVE
	(LEXICAL –)	CONTROL
	(COMPLETE –)	CONTROL
Adjuncts	PP-Adjunct	MAJOR
PP-Adjunct-feats	(CAT Preposition)	MAJOR
	(COMPLETE –)	MAJOR
	(PFORM Neutral)	MAJOR
LP-constraints	...	COMPLEMENTATION

In this entry the value of interest is the one for the feature VFORM in the subcat specification for the XComp, marked to be PAST-PART, which is not the default value (BASE) specified in the AUXILIARY class. Since *has* is a member of the PERFECTIVE class, and this class is a subclass of AUXILIARY, the rule for inheritance in (1) above requires that the more specific value for VFORM in

PERFECTIVE take precedence over the more general value in the AUXILIARY class.[3]

The illustration just given involved a subclass overriding a default specified in a higher class; also possible are cases where individual lexical entries override a class default. For example, the lexical entry for the modal verb *ought* must stipulate that its VP complement is infinitival, unlike the default base form for complements of other modals, as seen in (5). The minimal entry for *ought* is given in (6).

(5) a. John ought to read a book.
b. *John ought read a book.
c. John should read a book.
d. *John should to read a book.

(6) Minimal lexical entry for *ought*

OUGHT-1	
Superclasses	AUXILIARY, FINITE
Spelling	"ought"
Phonology	/at/
Semantics	(OUGHT agent:X proposition:Y)
Complements	
Subject-index	X
XComp-features	(VFORM Infinitive)
XComp-index	Y

Additional examples of this mechanism for overriding defaults will surface in succeeding chapters on lexical rules, where I make use of the inheritance convention to provide an account of *blocking* phenomena in both inflectional and derivational morphology.

[3]This precedence convention has a long tradition, dating back to Panini, and familiar in more recent work as the Elsewhere Condition of Kiparsky 1968, or the Proper Inclusion Precedence principle of Koutsoudas, Sanders, and Noll 1974.

3.2 Subcategorization

An important use of the inheritance mechanism just described is in the representation of subcategorization properties of lexical entries. There is a substantial amount of information that goes into a complete subcat entry, yet the vast majority of that information is predictable. In the present section I make explicit some of the properties of subcats that I have assumed above, to provide a more complete account of the form and content of lexical entries.

3.2.1 A constraint on subcategorization

In section 2.2 I stipulated that descriptions of complements and adjuncts within a lexical entry cannot themselves include specifications about subcategorization. Instead, I introduced the binary feature COMPLETE to signify whether or not a given category was required to have an empty list of complements. On this approach, a lexical entry can only specify values for any of the atomic-valued or category-valued features of a complement or adjunct category C (attributes like VFORM or SLASH), where COMPLETE and LEXICAL are the only such features that reflect properties of the subcat attributes of that category C. The claim implicit in this approach can be formulated as in (1). [4]

> (1) In specifying the properties of a complement or adjunct, the only piece of information a lexical entry ever needs to supply about the subcategorization properties of that category is whether or not the category itself still requires one obligatory complement.

This restriction embodies a strong prediction about the possible subcategorizations of lexical entries, excluding any entry which would make reference to the internal structure of its complements or adjuncts. For example, no verb may subcategorize for a verb phrase whose head verb itself subcategorizes only for a verb phrase, or only for a noun phrase. The most a verb can require of its complement is

[4] See also the independently developed, equivalent Locality Constraint of Pollard and Sag.

that it be a verb phrase instead of a full sentence; it cannot impose any requirements about the internal structure of that complement VP or S. Likewise, no preposition can require that its complement noun phrase contain a relative clause, or that it not contain a relative clause. If these seem silly, it is because they are, even though I have only suggested examples that drop down a single level in imposing restrictions. Without a principle like that in (1) constraining the power of lexical representation, one might expect to find a verb which requires that its sentential complement contain a verb phrase which itself contains a sentential complement which contains a noun phrase which contains a relative clause headed by a ditransitive verb. There

are no verbs like this.[5]

[5]One broad class of apparent counterexamples to the principle in (1) involves predicates that select for prepositional phrases headed by a specific preposition, as with the verb *rely* which must be followed by a PP headed by *of*. Here, as mentioned earlier, I follow GKPS in making use of a head feature PFORM whose value identifies a particular member of one of the closed classes including prepositions, and I take advantage of prepositions being the heads of their phrases, to effect (via the Head Feature Convention) the presence of this PFORM feature on the phrasal node dominating the PP, so it can be referred to in specifying subcategorization. Selection by a verb for a particular preposition, then, is quite analogous to selection by a verb for a verb phrase whose head has a particular morphological form such as finite or base or present participle. In both cases it is information about the morphology of the head of the complement phrase which is being specified by the lexical entry, and it is just this information which is guaranteed to be available at the phrasal level given the presence of the Head Feature Convention.

A more awkward apparent counterexample to the strong claim in (1) is posed by words like *kind* or *sort* or *type*, as illustrated in (i).

(i) a. What kind of movie makes you cry?
b. What sort of horror movie makes you cry?
b. *What kind of any movie makes you cry?
c. *What sort of the movie makes you cry?
d. *What type of Fellini's movies makes you cry?

It would appear that these nouns subcategorize for a prepositional phrase whose complement must be a nominal phrase (common noun with or without modifiers, but still missing its final complement) but cannot be a full noun phrase. However, the data is at best confusing. For example, it does seem possible for the complement of *of* to be a noun phrase with determiner *a(n)*, as in (ii).

(ii) a. What sort of a fool do you take him for?
b. What kind of a dog are you looking for?

Also, if the object of *of* is a plural noun, the main verb may have to show plural agreement instead of agreeing with what is presumably the head noun *kind*:

(iii) a. What kind of movies make you cry?
b. *What kind of movies makes you cry?

If *kind* itself is plural, then it seems that the noun must also be plural or mass, though still without determiner or genitive subject:

(iv) a. What kinds of movies make you cry?
b. *What kinds of movie make you cry?
c. *What kinds of these movies make you cry?
d. What kinds of wine do you prefer?

What these observations suggest is that it may be more accurate to treat *kind of* as part of a complex determiner which agrees in number with its head noun (what looked like the complement of *of*). On this approach, *kind* would have to agree with both singular and plural nouns, if (iiia) is really good (though it seems awkward), and the examples in (ii) would remain unaccounted for. Phonological evidence seems to support this analysis, given the examples in (v) where the *of* has been reduced (using conventional orthography such as it is).

(v) a. What kinda book is that?
b. *What kinda a book is that?
c. That type o' thing should not be allowed.
d. This sorta picture is what I'm looking for.

What we do find are lexical entries that require a verb phrase complement rather than a sentential complement, or vice versa, or a noun phrase rather than a noun. These are just the distinctions expressible using the binary features LEXICAL and COMPLETE. What these features embody is a precise hypothesis concerning the amount of information about a category that must be available within a lexical entry for subcategorization. No other mechanism is provided in the lexical entry structure to specify additional constraints on complements or adjuncts.

This constraint on possible subcategorization finds a strong parallel in the limitations on possible control relations that exist. It is well-known that in both raising and equi constructions, it is always the grammatical subject of the complement phrase which is controlled, and never any more oblique element within that complement phrase. So while examples like those in (2) are common across languages, no cases are attested where an argument of a higher clause predicate has to control a non-subject argument in a lower clause, as illustrated in (3).

(2) a. John tried to interview Mary.
(where John controls the grammatical subject of *interview*)
b. John persuaded Sally to interview Mary.
(where Sally controls the grammatical subject of *interview*)

(3) a. *John tried (for) Mary to interview.
(meaning *John tried to be interviewed by Mary*
where John controls the grammatical object of *interview*)

b. *John persuaded Sally (of?) Mary to interview.
(meaning *John persuaded Sally to be interviewed by Mary*
where Sally controls the grammatical object of *interview*)

If control is a relation which is constrained in part by the subcategorization properties of lexical entries, then the fact that only subjects are controlled might follow directly as a consequence of the principle in (1). I merely raise the possibility here, since to work it out in detail would be beyond the scope of the present discussion, requiring the elucidation of a theory of control, not the main thrust of this work.

While it is clear that the proper analysis of these constructions will bear on the principle of subcategorization I proposed in (1), more work needs to be done to determine that analysis. In the meantime, I leave the principle stated as in (1), lacking clear evidence to the contrary.

See Sag 1986 for such an elucidation.

In short, by incorporating the principle in (1) within the basic structure of the lexical entry, I constrain the power of subcategorization in a way which permits sufficient complexity without opening the door to unmotivated complexity in the specification of complements and adjuncts. Moreover, this implicit constraint on lexical entries will play an important role in restricting the power of lexical rules which relate lexical entries, without requiring any stipulation to this effect about the form or functioning of lexical rules themselves, as I show in the next chapter.[6]

3.2.2 Optional complements vs. adjuncts

In the introduction to the structure of lexical entries in section 2.1, I separated subcat information about obligatory and optional complements from information about adjuncts. I summarize here the motivations for making a sharp distinction between optional complements and adjuncts (which are also optional), a distinction needed both in lexical entries and in phrase structure rules.

The single most important difference between adjuncts and optional complements is that optionals may be assigned a thematic role in the lexical entry while adjuncts may not. To make this general claim explicit will require that more be

[6]This hypothesis about subcategorization dates back at least to Chomsky 1965:96, and is discussed in Kajita 1968, where a potential counterexample is considered, namely the verb *serve*. See the contrast in (i), then note the sentence in (ii), which suggests that the phenomenon is probably semantic in nature, not syntactic - a solution developed in Pollard and Sag 1987.

(i) a. This ice will serve to chill drinks. (from Kajita 1968)
b. *This ice will serve to melt.

(ii) a. This article should serve to antagonize, just like every other article by that author.

Another class of possible counterexamples concerns the analysis of the finite VP in sentences like those in (iii), where number agreement is required between the filler and the finite-VP complement.

(iii) a. Which book did John think has become a bestseller?
b. *Which book did John think have become a bestseller?
c. That's the author who you said was never going to be famous.
d. *That's the author who you said were never going to be famous.

I take up these constructions in chapter seven, and show there that the necessary number agreement is handled by general mechanisms while remaining consistent with the principle set forth at the beginning of this section. Contra Pollard 1985b (WCCFL), who does not subscribe to said principle.

said about the semantic framework I adopt, but the notion itself should be clear. Consider the sentence in (4):

(4) Mary was loudly applauded by Sally.

An optional complement like *by Sally* has a particular thematic role assigned by the passive verb *applauded*, namely that role which is assigned to the subject of the corresponding active verb *applaud*. On the other hand, there is no thematic role assigned to an adjunct like the adverb *loudly*, which contributes information about manner, regardless of the predicate.

A second, related distinction is that individual lexical items can and do impose idiosyncratic restrictions on optional complements (including but not restricted to thematic role assignment), but no lexical item imposes idiosyncratic constraints on any of its adjuncts. All adjunct specifications are supplied in class definitions, and hold without exception for the members of that class. For example, the entry for the noun *preference* contains an idiosyncratic specification for an optional prepositional phrase headed by *for*, as in (5a), with the role assigned to the complement of *for* being the same as that assigned to the direct object of the corresponding verb *prefer*, as in (5c). Not even the ubiquitous preposition *of* will do in this case, as shown in (5b).

(5) a. Your preference for chocolates is no secret.
b. *Your preference of chocolates is no secret.
c. It is no secret that you prefer chocolates.

This kind of idiosyncratic selection by deverbal nouns for particular prepositions is commonplace, but no common noun idiosyncratically blocks modification by a relative clause, which I take to be an adjunct. In fact, all common nouns without exception submit to modification by each of the syntactic variants of relative clauses, as illustrated in (6). It is this lack of exceptionality which serves to distinguish adjuncts from optional complements.

(6) a. The book that John purchased was expensive.
b. The book John purchased was expensive.
c. The book purchased by John was expensive.

3.2.3 Selection by heads for adjuncts

Having drawn the distinction between optional complements and adjuncts, one might now question the assumption that heads select for adjuncts in the same way that they select for complements, both obligatory and optional. Since in many cases the adjunct might be semantically the functor, with the head the semantic argument, one might expect the syntactic dependency to mirror the semantic one. While this might seem appealing at first glance, I will argue that to give adjuncts the task of selecting for their heads would require additional expressive power and lead to loss of generalization.

Before presenting arguments to this effect, I suggest three properties which could be expected to contribute to the notion of *head*, and identify the one of these three which I take to be definitional. First is the syntactic dependency in a local phrase between the mother and one of the daughters, where that one daughter determines the values for any features not otherwise specified on the mother. This dependency, often expressed as the Head Feature Convention, allows that distinguished daughter to propagate syntactic information outside of the local phrase, by contributing the default properties of its mother. Second is the dependency between two daughters, where one daughter determines some properties of the other, as when a transitive verb requires that its first complement be a noun phrase. Third is the semantic dependency already mentioned, where (assuming semantic operations that expect functors and arguments) one distinguished daughter is the semantic functor, and the others are semantic arguments. While it is often the case that the head of a phrase is the distinguished daughter in all three of these dependencies, I take only the first of these to be definitional for heads, and in later sections I give illustrations of the latter two dependencies in which a non-head is the distinguished daughter.

Since the ability to impose restrictions on a sister in a local phrase is not reserved for heads alone, it is consistent within this framework to imagine that adjuncts might select for their heads. Yet there are some significant formal obstacles which seem surmountable only at the cost of some otherwise unmotivated extensions to the framework I have sketched. For the sake of concreteness, consider relative clauses, which are clearly adjuncts for common nouns. In order to have a relative clause

select for a common noun, that selectional information would have to reside on the relative clause node, and the source of the information would have had to be some lexical entry contained within the relative clause, presumably either that of the head (the inflected verb), or that of the relative pronoun. But it couldn't very well be the responsibility of the relative pronoun, since some relative clauses don't contain a relative pronoun (cf *The man I met*). So it would have to be part of the entry for the head verb, but then the selectional information would have to be conditional, only being relevant in case the verb happened to end up as the head of a relative clause. More precisely, the information would have to be stored in such a way that if a verb phrase like *singing a song* appeared as the complement of a verb like *continue* as in *continued singing a song*, then the selectional information that *singing* (and hence *singing a song*) might impose on the head it was sister to would not be used in this case, but if that VP *singing a song* appeared as an adjunct to a common noun, as in *the man singing a song*, then the selectional information contained in the entry for *singing* would be imposed on the head. Worse yet, the information would have to be represented in such a way that a VP like *sings a song* would not appear as adjunct to a noun, while still allowing *who sings a song* to be such an adjunct, as well as *John sings*, as in *the song John sings*. Building these conditions and constraints into the lexical entries for verbs would require a significant increase in the expressive power permitted for lexical representation, including some mechanisms for context-sensitivity that would have to interact with the parser in some fashion.[7]

A second class of arguments against the notion of having adjuncts select syntactically for their heads involves loss of generalization which would result from such an approach. A straightforward illustration involves those adjuncts to common nouns which are often misleadingly termed *reduced relatives*, as in (7).

[7]There is of course a third option available: the addition of a separate phrase structure rule that just combines nouns with that-less relative clauses. Such a rule would stand out from the others proposed here by virtue of its highly specific nature, and therefore seems undesirable, but I have not provided any formal basis for excluding such a rule from the grammar.

(7) a. The woman singing a song is famous.
b. The woman assigned to this committee is famous.
c. The woman taller than you is famous.
d. The woman in the corner is famous.

Those phrases which can modify *woman* in (7) are similar to those which can appear as complements of the copula *be*, as illustrated in (8). The post-nominal modifiers differ from copular complements in at least two ways: nominal phrases cannot appear as post-nominal modifiers, and gerundive forms as in *anyone knowing the answer* cannot appear as complements to the copula. But both sites permit a wide range of predicative phrases headed by adjectives, prepositions, and verbs.

(8) a. The woman is singing a song.
b. The woman is assigned to this committee.
c. The woman is taller than you.
d. The woman is in the corner.
e. The woman is a doctor.

To account for the constructions in (8), one can simply say that the verb *be* subcategorizes for a predicative phrase, having distinguished between predicative and non-predicative verb forms and adjectives. Now if heads select for adjuncts, a quite similar generalization can be expressed for the examples in (7), since the definition of the common noun class can include as one of the possible adjuncts predicative phrases, with the syntactic specification of this type of adjunct similar to that for *be*, though with a restriction excluding nominal phrases. If, on the other hand, one were to have adjuncts select for their heads, then each of the predicative verb forms, predicative adjectives, and prepositions would have to contain independent stipulations that they (or phrases that they head) could appear as adjuncts to common nouns. The generalization that any (non-nominal) predicative phrase can serve as adjunct to any common noun (syntactically) would be lost.

For these reasons, I will continue to have heads select for adjuncts rather than the converse, consistent with the word class hierarchy sketched at the beginning of this chapter. The resulting lack of congruence between syntactic dependency and semantic dependency (at least for some adjuncts) does not seem to present any real difficulty.

Chapter 4

Phrase structure rules

To make sense of the structure for lexical entries proposed here, one needs to know how these entries interact with the syntactic rules used to combine them into phrases. Given the wealth of information contained in a fully specified lexical entry, it should be no surprise that relatively little information needs to be specified in the syntactic rules that admit these entries. What information does appear there must be combined with the information in lexical entries, so the conventions which govern that merging of information need to be made explicit.

The phrase structure rules and feature conventions that I present in this chapter will serve as an adequate basis for the illustrations and analyses that I offer in succeeding chapters, though they are rather informally presented. The reader interested in a more detailed presentation of this kind of grammar is encouraged to consult Pollard and Sag 1987, who also suggest which properties of such a grammar are language-particular, and which can be ascribed to universal grammar.

Structure of the chapter

Section 4.1 introduces the connection between lexical entries and phrase structure rules, then presents the principal feature conventions and phrase structure rules needed for the fragment of English studied here. In section 4.2 I illustrate in some detail the way in which these mechanisms are used in producing (or admitting) the

structure for some simple sentences. Then in section 4.3 I introduce one additional phrase structure rule, to complete the basic account of unbounded dependencies that I adopt, an account which makes a slight improvement on the Lexical Head Constraint introduced in Flickinger 1983 and employed by GKPS 1985.

4.1 Rules and feature conventions

Consider the phrase structure rule in (1) below, which combines a lexical head with zero or more complements immediately following it, to form a phrasal constituent. Making use of one of the linear order constraints supplied by the lexical head (inherited from the COMPLEMENTATION class), the linear order of the head with respect to its complements is predicted. And given our assumption that the Complements attribute of the lexical head is ordered to reflect the relative obliqueness of the complements, we can take advantage of the observation by Sag 1987, and Pollard and Sag 1987, that this order corresponds directly to the (default) linear order for those complements in a phrase. Thus the PS rule does not need to stipulate either the order of the head with respect to its complements, or the relative order of those complements.

(1) Lexical Complements PS Rule

$$\begin{array}{lll} \text{X} & \longrightarrow & \text{Head, Complement*} \\ & & \\ \text{where} & & \text{X = (COMPLETE –) (LEXICAL –)} \\ & & \text{H = (COMPLETE –) (LEXICAL +)} \end{array}$$

This rule is to be interpreted as follows: the phrase X is composed of one or more constituents, consisting of a single head daughter together with zero or more complements (using the Kleene star notation). The complements may be either obligatory or optional, with the order of the complements reflecting relative obliqueness.

Since the Complements attribute imposes only a partial order on its values, two complements introduced for a given lexical head may be freely ordered with respect to each other. This is desirable given examples like those in (2), where PP-By and VP complements of the passive verb *persuaded* are unordered with respect to each other, giving both grammatical examples.

(2) a. Mary was persuaded by her press secretary to run for president.
b. Mary was persuaded to run for president by her press secretary.

Restrictions are placed by the rule on both the mother category X and the head daughter H, but none are imposed by the rule on the complement daughter.

The restrictions on the mother and head daughter make reference to the binary features LEXICAL and COMPLETE, which have the following interpretations for each of their two possible values. These interpretations are based on the assumption (familiar from GKPS and others) that English attaches all complements except the last obligatory one as sisters of the lexical head.

LEXICAL	+	No restriction is placed on the complements of the constituent.
LEXICAL	–	The list of complements may have at most one obligatory element remaining.
COMPLETE	–	The list of complements must have at least one obligatory element remaining.
COMPLETE	+	The list of complements must be empty.

The intuitions behind the definitions for the feature LEXICAL are straightforward: a constituent which is (LEXICAL +) retains its full list of Complements as defined in the lexicon, and may therefore have any available combination of obligatory and optional complements. Constituents which are marked (LEXICAL –) are in general constructed by phrase structure rule, so given the above assumption about attachment of complements in English, a non-lexical constituent must have either one remaining obligatory, or none.

The interpretation of values for the feature COMPLETE given above simply follows from the definition of that feature given in the previous section. A constituent which is (COMPLETE –) cannot yet have found its final obligatory complement (by definition), so must have at least that one complement remaining. And a constituent which is (COMPLETE +) must (again by definition) have found its last obligatory complement, leaving empty the list of complements.

Given these definitions, one can see that the rule in (1) requires that the head be lexical and be missing at least one complement; (1) also requires that the resulting constituent be still missing just the final complement, having picked up all other complements (obligatory or optional).

As mentioned, the properties that each complement C must satisfy are determined by the subcategorization properties found on the head constituent, together with any constraints imposed by the relevant PS rule. So in the case where the head

is a finite auxiliary verb like *did* as in the sentence *John did try to work* the subcategorization requirement in that verb's lexical entry is for the rightmost complement to be a verb phrase which is morphologically base in its form. To help see this, I provide a partially redundant entry for *does* in (3).

(3) Partially redundant lexical entry for *does*

DO-1-PAST	
Superclasses	AUXILIARY, PAST
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL +) ...
Complements	Subject XComp
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative)
XComp-features	(CAT Verb) (COMPLETE –) (LEXICAL –) VFORM Base)
...	

4.1.1 Three conventions

Given the schematic nature of phrase structure rules like that in (1), some additional mechanisms are needed to effect the merging of information contributed by the lexicon with information contributed by syntactic rules. I present these mechanisms in the form of three conventions, though the second could be collapsed with the first, given certain reasonable assumptions about the grammar. Since my primary concern here is to show how lexical entries as I have presented them here interact with syntactic rules, I keep these first two conventions separate for the sake of exposition. See Pollard and Sag 1987 for a more formal presentation of the grammar which does collapse these first two conventions.

The basic principle which constrains the *merging* of information both for inheritance and for participation in phrase structures is a restricted form of unification which I will term Feature Unification, since its principal use is to govern the merging of two sets of syntactic feature-value pairs. [1]

(4) **Feature Unification**

[1] This notion. introduced by Martin Kay and Ron Kaplan, plays a central role in much current work in formal linguistics; a good introduction is provided by Shieber 1986. Cf. Pollard and Sag 1987 for more detail about this mechanism as used in the HPSG framework.

> Given two sets A and B of feature-value pairs, the result of unifying A and B will be the set C of feature-value pairs, where for each pair (F V) consisting of a feature F and (possibly singleton) set of values V, the assignment for V is determined as follows:[2]
>
> (a) If only one of A or B contains an entry for F, then the entry (F V) in C will be identical with that entry.
>
> (b) If both A and B contain an entry for F, where A assigns V1 to F, and B assigns V2 to F, the value assigned to F for C is the result of intersecting the sets V1 and V2, unless the intersection is empty, in which case unification fails.

The principle is perhaps most easily grasped by considering an example or two. Let A and B be defined first as in (5).

(5) A = [(CAT Determiner) (AGREEMENT Third-Singular Mass) (LEXICAL +)]
B = [(CAT Determiner) (AGREEMENT Mass)]

According to the definition in (4), the unification C of A and B will be that in (6), with the justification for each feature's value following.

(6) C = [(CAT Determiner) (AGREEMENT Mass) (LEXICAL +)]

CAT:	Both A and B contain a specification, so clause (4b) makes the value for C the result of intersecting the singleton set (Determiner) with itself, giving the same value.
AGREEMENT:	Again, both A and B contain a specification, so C's value is the intersection of the set (Third-Singular Mass) with the set (Mass), resulting in the singleton set (Mass).
LEXICAL:	Only A contains a specification, so by clause (4a) the value for C is the same as for A.

[2] This principle must be interpreted as applying recursively, since some features such as SLASH are category-valued. Again, see Shieber 1986 for a careful treatment of feature unification.

An example illustrating a failure of unification is given in (7).

(7) A = [(CAT Determiner) (AGREEMENT Third-Singular Mass)]
B = [(CAT Determiner) (AGREEMENT Plural)]

The unification C of A and B fails in this case, since both A and B specify values for the feature AGREEMENT, where the intersection of the two sets (Third-Singular Mass) and (Plural) is empty.

It is this convention defined in (4) which governs all merging of syntactic feature information in this framework, including both the collecting of information for inheritance and also the merging of information involved in matching lexical entries with category specifications in syntactic rules. But given the form of grammar rules like that in (1) and lexical entries like that in (3), a second convention is required to make the subcategorization specifications in (3) useful in concert with the syntactic information in (1). I term this Subcat Unification, and note that it makes essential use of Feature Unification.

Subcat Unification

> The restrictions that a head imposes on each one of the complements and adjuncts it subcategorizes for must be unified with any restrictions imposed on that complement or adjunct by the relevant phrase structure rule which admits it.

This convention can be thought of most simply as a constraint which must be satisfied by each relevant local subtree in a phrase structure tree for each grammatical sentence of the language.

The third convention will be the most familiar, accounting for the relationship between features of a head daughter and its immediately dominating node in a local subtree. The formulation of this principle assumes some distinguished subset of the set of features, referred to here as the *head features*, which I take here to include all atomic-valued features and the category-valued feature SLASH. The intuition is the same one formalized in GKPS 1985, simplified for purposes of exposition.[3]

[3]This formulation differs from that of Pollard and Sag 1987 in at least two important respects:

(8) **Head Feature Convention**

In a local subtree dominated by a node X and containing a head daughter H, where the subtree is admitted by a phrase structure rule PSR,

(a) All head features which are not assigned a value by PSR for X have the same values for X as they do for H.

(b) The subcategorization specifications for X are the same as those for H, augmented with links for any subcats already associated with complements (which are omitted in node descriptions here).

---

first, I have not attempted to generalize the Head Feature Convention to treat subcat information in the same way as the atomic and category-valued features; and second, Pollard and Sag make the strong assumption that head features all have identical values on mother and head daughter, allowing no override. Given my use of the features COMPLETE and LEXICAL to encode properties of subcategorization, with their values propagated in part by means of the Head Feature Convention, I have not explored the incorporation of this strong (and thus in principle desirable) constraint within the framework presented here.

4.1.2 Some phrase structure rules

Before illustrating the interaction of the lexicon and the grammar, I present several phrase structure rules which are basic to the grammar of English, though additional rules will of course be needed. The first of these rules is a repetition of rule (1) given above; this rule and the Inverted-S rule use the Kleene * notation to signify zero or more constituents. In addition to the features COMPLETE and LEXICAL, these rules refer to the INVERTED feature, which is only relevant for auxiliary verbs. Remember that the right-hand side members do not have an order stipulated in the rule; rather, that order is determined by independent LP constraints, which are introduced in word classes or lexical entries.[4]

(PS1) Lexical Complements

X $\longrightarrow$ Head, Complement*

where X = (COMPLETE –) (LEXICAL –)
H = (COMPLETE –) (LEXICAL +)

(PS2) Adjuncts

X $\longrightarrow$ Adjunct, Head

(PS3) Final Complement

X $\longrightarrow$ Complement, Head

where X = (COMPLETE +)
H = (COMPLETE –) (LEXICAL –)

(PS4) Inverted S

X $\longrightarrow$ Head, Complement*

where X = (COMPLETE +)
H = (LEXICAL +) (INVERTED +)

[4]Among the familiar constructions of English not covered by these rules are coordination, sentence-initial PP modifiers, and imperatives. Constructions that may or may not be analyzable with these rules, suitably modified, include negation, possessive NPs, and comparatives.

(PS5) Filler-Gap

X $\longrightarrow$ Filler, Head

where H = (CAT Verb) (COMPLETE +) (SLASH α)
F = α

(PS6) S-bar

X $\longrightarrow$ Complementizer, Head

where X = (COMP α)
H = (CAT Verb) (COMPLETE +) (COMP None)
C = (COMP α)

One more phrase structure rule will be added to this collection in the next section.[5]

4.1.3 Linear order constraints

When introducing the notion of linear order constraints in chapter 2, I was concerned primarily with the issue of how to properly introduce such constraints as properties of word classes, to be inherited or overridden by lexical entries. Here I consider more carefully the kinds of constraints necessary, and illustrate where the default constraints are overridden.

For convenience, I repeat here the constraint taken from Pollard and Sag which I included in the definition of the COMPLEMENTATION class in chapter 2, as a value of the LP-constraints attribute.

(9) a. Head [LEXICAL +] < Complement

This constraint, together with the obliquely ordered Complements list for lexical entries, will account for the order of elements in most applications of rules PS1 and PS4, including verb phrases like *gave a book to John* and inverted sentences like *Is*

[5] Rules almost identical to PS1, PS3, and PS4 were proposed by Pollard 1985:255, and are now incorporated along with a rule much like PS5 in Pollard and Sag 1987, though they make use of a feature convention to do the work undertaken here by the feature COMPLETE.

John afraid of lions?. See Sag 1987, and Pollard and Sag 1987 for a discussion of the predictive power of (9a) and the obliqueness convention. However, additional constraints are needed to account for the desired order of elements in the other rules given above. For the sake of the discussion to follow, I add some additional LP constraints; some I take from Pollard and Sag 1987, and others I introduce to provide the desired effect with the rules given.

Pollard and Sag propose a constraint to account for the relative order of the daughters in the Filler-Gap rule, a constraint that they suggest can be generalized to account for the order of daughters in the Final-complement rule as well. I introduce the two constraints separately, since my aim here is not to produce an elegant account of linear precedence; see Sag 1986 and Pollard and Sag for such work. To these, I add a third LP constraint, to express the relative order of a sentence and its complementizer. These constraints, which should be introduced in the appropriate classes within the lexicon hierarchy, provide the desired orders for the right-hand side members in rules PS3, PS5, and PS6, respectively.

b. Complement < Head [LEXICAL –]
c. Filler < Head
d. Complementizer < Head

There are two distinct uses that I will make of the Head-Adjunct rule in discussion to follow: one use is to combine a lexical adjunct with a (lexical) head, as in *very tall* and *old man*, where in each case the adjunct precedes the head. The other use is for phrasal adjuncts, which always follow the head, as in *man I met* and *worked on Monday*. The two LP rules needed (barring an elegant collapsing of some sort) are given in (e-f).

e. Adjunct [LEXICAL +] < Head
f. Head < Adjunct [LEXICAL –]

4.1.4 Exceptions to default linear order

In presenting constraints on linear order as inheritable properties of lexical entries, I suggested that treating such constraints as defaults was desirable, since exceptions exist for at least some of these constraints. I illustrate here two such exceptions, one

affecting the lexical head LP constraint introduced in the definition of the COMPLEMENTATION class, and the other overriding the default order of elements in the Complements attribute, which is based on relative obliqueness of the complements.

The first of these constraints, repeated in (9a) above, states that lexical heads precede their complement sisters in a phrase. This correctly predicts that verbs, prepositions, nouns, and adjectives all precede their objects, and also predicts that auxiliary verbs in inverted sentences precede their subjects, as illustrated in (10).

(10) a. saw John
b. in Seattle
c. preference for chocolate
d. eager to win
e. does John sing

These examples make it clear that (10) holds true for the overwhelming majority of lexically headed phrases in English. However, as Thomas Wasow (p.c.) notes, English can boast of at least a few postpositions, illustrated in (11-14), where the postpositional head, as its name indicates, follows its complement.[6] Wasow notes further that these postpositions cannot be assimilated to adverbs and measure phrases as in *Mary left two weeks later*, since the postpositions must have a preceding NP, while adverbs like *later* do not.

(11) a. Mary left two weeks ago.
b. *Mary left ago two weeks.

(12) a. That town is ten miles away.
b. *That town is away ten miles.

(13) a. Mary will graduate five years hence.
b. *Mary will graduate hence five years.

(14) a. His injury notwithstanding, John will win.
b. Notwithstanding his injury, John will win.

[6] I am grateful to T. Wasow for sharing his hoard of postpositions with me; that hoard contained all of the following examples except *away*. Additional postpositions include *across*, *around*, and perhaps *off*, as in *That town is ten miles off*.

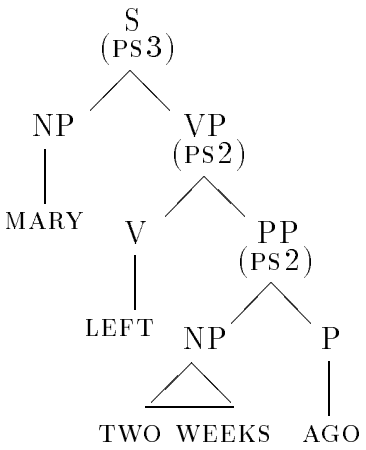
What the lexical entry for *ago* must do is override the default ordering constraint given above in (9a), replacing it with the reverse constraint. To express this idiosyncratic property of *ago*, I employ the rather clumsy but transparent notation illustrated in (15), where I give the lexical entry for *ago*, ignoring phonology, spelling, and semantics.

(15) Non-redundant entry for *ago*

AGO	
Superclasses	PREPOSITION
Semantics	...
Linear-order	Complement < Head [LEXICAL +]
	Block: Head [LEXICAL +] < Complement

This entry explicitly blocks the inherited default ordering constraint, replacing it with the reverse of the default. It is then the interaction of this locally specified LP constraint with the phrase structure rule PS1 that will provide the desired structure for *two weeks ago*, as illustrated in (16).

(16) Mary left two weeks ago.



If the word *notwithstanding* can serve as either preposition or postposition, its entry in (17) will differ from that for *ago* in that it will block the default constraint, but introduce no replacement, leaving free the relative order of lexical head and complement, to produce both examples in (14) above.

(17) Non-redundant entry for *notwithstanding*

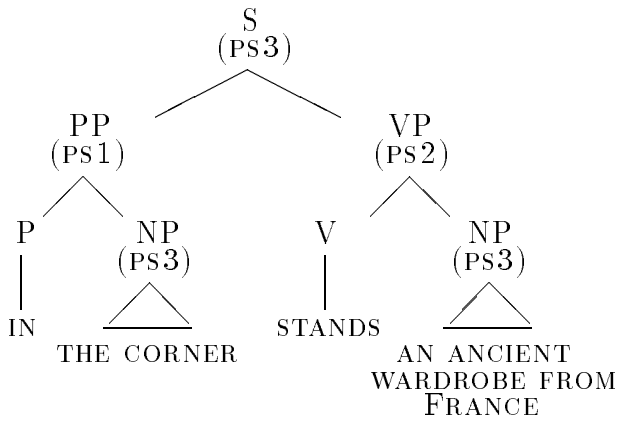
NOTWITHSTANDING				
Superclasses	PREPOSITION			
Semantics	...			
Linear-order	Block:	Head [LEXICAL +]	<	Complement

The second class of exceptions involves verbs which seem to override the default order of elements in the Complements attribute, as illustrated in (18).

(18) a. In the corner stands an ancient wardrobe from France.
b. On my desk sit five memos from the same dean.

In these sentences, the noun phrase that shows agreement with the verb, and which must therefore be the subject, follows the verb, while the locative phrase, more oblique than the subject, precedes the verb. Since these sentences cannot be instances of topicalization (which always leaves the subject preceding the main verb), I will assume that the structure of (18a) must be that in (19), annotated with the obvious PS rules.

(19) In the corner stands an ancient wardrobe from France.



For this structure to be admitted using the PS rules indicated, one lexical entry for *stands* must include a non-default order for its complements, since PS3 picks up just the final complement, and PS1 picks up all of the other complements (in this case just one). This unusual lexical entry will then be related by lexical rule

to the more familiar entry for *stands*, illustrated in (20a), which I assume will make no mention of the Complements attribute, since the order of its values will be the default as given in (21), with the least oblique element first.

(20) a. An ancient wardrobe from France stands in the corner.
b. Five memos from the same dean sit on my desk.

(21) Complements: Subject XComp

In contrast, the entry for the irregular *stands* of (19) will include (in its non-redundant form) the following stipulation for the Complements attribute:

(22) Complements: XComp Subject

Given the straightforward interaction of PS rules with this Complements attribute, the structure in (19) follows directly from the stipulation in (22).

This analysis predicts the number agreement between verb and post-verb complement seen in (18), since that complement is still the subject, a position reinforced by the fact that the sentence initial PP in (18) does not happily participate in raising constructions, as it ought to if it were the subject. Contrast (23a) with (23b), in which the PP is arguably the subject, so does appear in raising constructions.[7]

(23) a. *In the corner seems to stand an ancient wardrobe.
b. In the bathtub seems to be his favorite place to write poetry.

Examples like *stand* as in (18a) provide support for the notion that the elements of the Complements list have a default order which can be overridden, maintaining the view that position in the sentence is, in principle, independent of the agreement properties, thematic role assignment, and control properties which cluster for particular grammatical functions. Here again the notion of default properties which can be overridden is crucial in enabling the expression of relevant generalizations while allowing for exceptions.

[7]This argument was made by Adrian Akmajian in an unpublished LSA paper (1973). I am grateful to T. Wasow for rehearsing the argument, and for providing example (23b).

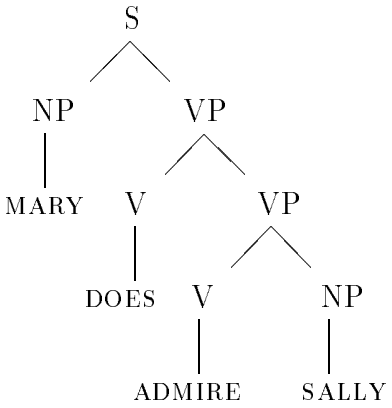
4.2 Two examples

To illustrate the effect of the three conventions introduced in the previous section, together with the grammar rules just given, I work through two examples, one grammatical and one not. For the sake of exposition in these examples, I will mention in the Complements attribute of phrasal nodes only those complements which have not yet been associated, though in fact they are still present, linked to constituents in the phrase structure.

The first example is given in (1), with its constituent structure given in (2), where the node labels are convenient abbreviations which I will flesh out shortly.

(1) Mary does admire Sally.

(2)



Starting with the verb phrase *admire Sally*, consider the lexical entries given in (3-4).

(3) Partially-specified lexical entry for *Sally*

SALLY-1	
Superclasses	PROPER SINGULAR
Atomic-features	(CAT Noun) (COMPLETE +) (LEXICAL +) (NTYPE Proper) (NFORM Norm) ...
Complements	
...	

(4) Partially-specified lexical entry for *admire*

ADMIRE-1	
Superclasses	TRANSITIVE MAIN-VERB BASE
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL +) (VFORM Base)
Complements	Subject DObject
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative)
DObject-features	(CAT Noun) (COMPLETE +)
...	

Given this entry for *admire*, the phrase structure rule given as PS1 will suffice to combine the verb with its first complement, which is the direct object, specified to be a noun phrase. For convenience, I repeat the definition of this rule here.

(PS1) Lexical Complements PS Rule

X ⟶ Head, Complement*

where X = (COMPLETE –) (LEXICAL –)
H = (COMPLETE –) (LEXICAL +)

One procedural description of how this rule can be used to admit the phrase *admire Sally* is the following.

(5) Application of phrase structure rule PS1

a. Unify the features of the entry for *admire* with those of the head H. This makes H have the following features:

(CAT Verb) (COMPLETE –) (LEXICAL +) (VFORM Base)

b. For each of the subcats in the Complements attribute of this head except the leftmost one (the subject), do the following two-part unification, preserving the order of these subcats on the list in the order of the corresponding daughters in the phrase. (Since *admire* only has one such subcat, for the direct object, the Complement* in PS1 will be realized

by only one complement, and the ordering constraint will be trivially satisfied.)

(i) Unify the features of the complement subcat in the entry for *admire* with those of the complement C in the rule (a trivial unification), Given the lexical entry in (4), the features on C in the rule will now be
(CAT Noun) (COMPLETE +).

(ii) Unify the features of the entry for *Sally* with those of C in the rule, which are from the unification in (b.i), to give
(CAT Noun) (COMPLETE +) (LEXICAL +)
(NTYPE Proper) (NFORM Norm)

c. Since both the head H and the complement C in the rule are now filled, and exactly one subcat remains in the Complements attribute for the head, apply the Head Feature Convention to propagate the relevant information from the head H to the dominating node X. This results in the node description given in (6) for the verb phrase for *admire Sally*.

(6) Description of phrasal node dominating *admire Sally*

admire Sally	
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Base)
Complements	Subject
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative)

A second application of rule PS1 will combine *does* with *admire Sally* to admit the verb phrase *does admire Sally*, using steps quite analogous to those just taken in admitting *admire Sally*. The entry for *does* is partially specified in (7).

(7) Partially-specified lexical entry for *does*

DO-1-PRES-3RDSG	
Superclasses	AUXILIARY, THIRD-SINGULAR
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL +) (VFORM Finite)
Complements	Subject XComp
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative) (AGREEMENT Third-Singular)
XComp-features	(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Base)
...	

Using this entry together with the specification in (6) for the node dominating *admire Sally*, the second application of rule PS1 is as in (8).

(8) Second application of phrase structure rule PS1

a. Unify the features of the entry for *does* with those of the head H. This makes H have (at least) the following features:

(CAT Verb) (COMPLETE –) (LEXICAL +) (VFORM Finite)

b. For each complement subcat of *does* that is not the leftmost one in the Complements attribute, in this case only the XComp, perform the following two-step unification:

(i) Unify the features of the complement subcat in the entry for *does* with those of the complement C in the rule (a trivial unification), Given the lexical entry in (7), the features on C in the rule will now be

(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Base)

(ii) Unify the features of the node for *admire Sally* with those of C in the rule, which are from the unification in (b.i), to give

(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Base)

c. Since both the head H and the complement C in the rule are now filled, apply the Head Feature Convention to propagate the relevant information from the head H to the dominating node X. This results in the node description given in (9) for the verb phrase dominating *does admire Sally*, where I leave off of the Complements list the already-associated DObject complement.

(9) Description of phrasal node dominating *does admire Sally*

does admire Sally	
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Finite)
Complements	Subject
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative)

To combine the subject *Mary* with the verb phrase *does admire Sally*, we use the Final Complement phrase structure rule given above as PS3, repeated here.

(PS3) Final Complement PS Rule

X $\longrightarrow$ Complement, Head

where X = (COMPLETE +)
H = (COMPLETE –) (LEXICAL –)

This rule states quite simply that a phrase X which is complete (COMPLETE +) can be composed of two constituents: a head H which has exactly one obligatory complement remaining to be found, and the complement C itself. Like the Lexical Complements rule (PS1), rule PS3 stipulates nothing about the complement C; all constraints on that constituent C will come from the subcategorization information provided by the constituent that is the head of this phrase. The ordering of the complement before the head is ensured by the linear precedence rules discussed above. This rule will serve to combine verb phrases with their subject noun phrases to make sentences, and also serves to combine nominals with their determiners to

make noun phrases. In (10) I walk again through the steps involved in applying the rule.

(10) Application of phrase structure rule PS3

a. Unify the features of the entry for *does admire Sally* with those of the head H. This makes H have (at least) the following features:

(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Finite)

b. Unify the features of the one remaining subcat in the Complements attribute for the head *does admire Sally* with those of the complement C in the rule (a trivial unification), so C will now have the following features:

(CAT Noun) (COMPLETE +) (CASE Nominative)

c. Unify the features of the entry for *Mary* (which will be just like those for *Sally* in (3)) with those of C in the rule, which are from the unification in (b), to give

(CAT Noun) (COMPLETE +) (CASE Nominative) (LEXICAL +)
(NTYPE Proper) (NFORM Norm)

d. Since both the head H and the complement C in the rule are now filled, apply the Head Feature Convention to propagate the relevant information from the head H to the dominating node X. In this case, there is no subcategorization information to be transferred, so only the features from the verb phrase are passed up. This results, finally, in the description given in (11) for the node dominating the full sentence *Mary does admire Sally*.

(11) Description of sentence node dominating *Mary does admire Sally*

Mary does admire Sally	
Atomic-features	(CAT Verb) (COMPLETE +) (LEXICAL –) (VFORM Finite)
Complements	

Having worked carefully through an example of how a grammatical sentence is admitted by the rules and lexical entries provided, I now sketch in less detail an example of how an ungrammatical sentence fails to be admitted, given those same rules and entries. Consider the string *Mary does to admire Sally*, where the form of the embedded verb phrase is infinitival rather than base. In (12) I give a partially specified entry for the infinitival *to*, and in (13) a partial description of the node dominating *to admire Sally*, admitted by two successive applications of rule PS1, with the second application taking *to* as the head and *admire Sally* as the complement.

(12) Partially-specified lexical entry for infinitival *to*

TO-INFINITIVAL	
Superclasses	AUXILIARY
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL +) (VFORM Infinitive)
Complements	Subject XComp
Subject-features	(CAT Noun) (COMPLETE +)
XComp-features	(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Base)

(13) Description of phrasal node dominating *to admire Sally*

to admire Sally	
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Infinitive)
Complements	Subject
Subject-features	(CAT Noun) (COMPLETE +)

Now when we try to combine this verb phrase with the auxiliary verb *does*, again using rule PS1 with *does* as head and *to admire Sally* as complement, we run afoul

of the Subcat Unification convention. To aid in seeing this, I repeat the entry for *does* as (14).

(14) Partially-specified lexical entry for *does*

DO-1-PRES-3RDSG	
Superclasses	AUXILIARY, THIRD-SINGULAR
Atomic-features	(CAT Verb) (COMPLETE –) (LEXICAL +) (VFORM Finite)
Complements	Subject XComp
Subject-features	(CAT Noun) (COMPLETE +) (CASE Nominative) (AGREEMENT Third-Singular)
XComp-features	(CAT Verb) (COMPLETE –) (LEXICAL –) (VFORM Base)
...	

The specification in this entry for the XComp is for a verb phrase which is morphologically base, not infinitival. So when we try to unify the set of features which is the value of XComp-features in (14) with the set of features which is the value of Atomic-features in (13) in order to satisfy Subcat Unification, feature unification fails. Each set of features has a single value for the feature VFORM, but the two values are not the same, causing failure of unification. This failure means that rule PS1 cannot be used to combine *does* with *to admire Sally*. The only other relevant rule presented here is PS3, but the head in that rule must be (LEXICAL –), which prevents *does* from filling the head position. In order to admit *does to admire Sally* as a verb phrase, the lexicon for English would have to include an entry for *does* which subcategorized for infinitival verb phrases. No such lexical entry exists, so the verb phrase (and hence the string *Mary does to admire Sally*) cannot be admitted, and is ungrammatical.

Syntactic ill-formedness in this framework, in fact, will always be grounded in a failure of feature unification at some level, in trying to match a phrase or lexical entry either with some daughter specification in a phrase structure rule, or with a subcat specification in a lexical entry. The determination of grammaticality will depend both on the definitions of the word classes and lexical entries, and on the definitions of the phrase structure rules in the grammar.

4.3 Trace introduction

Before leaving this presentation of the grammar, I give a sketch of how traces are introduced, to account for sentences involving unbounded dependencies, which will also make use of the Filler-Gap rule introduced above as PS5. In GKPS 1985 a metarule was used to supply, for each lexically-headed PS rule in the base grammar, a corresponding set of derived rules each of which eliminated one of the sisters of the head, encoding the description of that sister as the value of the SLASH feature on the mother.

With most of the representation of complements now being done in the lexicon, the number of necessary phrase structure rules is much smaller than it was for GKPS. Indeed, there is now only one PS rule that would need to be input to a Slash Termination Metarule like that in GKPS, which leads to a rather obvious, but simple solution for trace introduction in the framework I have sketched: I simply add one more rule to the grammar.

(PS7) Traced Complement PS Rule

X	$\longrightarrow$	Head,	Traced-Complement, Complement*
where		X =	(COMPLETE –) (LEXICAL –) (SLASH α)
		H =	(COMPLETE –) (LEXICAL +)
		T =	α

This rule is simply a modification of the Lexical Complements rule (given as PS1 above) where one of the complements of the lexical head is left unfilled, and the description of that complement category is passed to the left-hand side of the rule as the value of its SLASH attribute. It is a natural counterpart for the Filler-Gap rule (PS5), which closes off the other end of the syntactic dependency represented by the SLASH feature on nodes in a phrase structure tree. The Traced Complement rule introduces a gap in a phrase, by introducing on a node a SLASH value that was not passed up by the Foot Feature Principle from one of the daughters. The Filler-Gap rule, in turn, provides the filler for that gap, terminating the unbounded

dependency conveyed by the Foot Feature Principle.

The Traced-Complement element in the rule differs from other complements in that it is not unified with a lexical item or phrase in a tree, so the only specifications for this element come directly from its head sister. It is this set of features which must be duplicated as the value of the SLASH attribute on the left-hand side. The *alpha-matching* device here is formally the same as that used in the Filler-Gap rule, intended to ensure identity of category description in the places marked α.

This mechanism for introducing traces enjoys all of the benefits of the lexically constrained metarule of GKPS 1985, as argued in Flickinger 1983, with one nice additional property. The metarule approach, while accounting for many of the constraints on extraction, failed to explain the lack of subject extraction in inverted sentences, since the auxiliary verb's subject is sister to a lexical head and hence should be extractable. See Pollard 1985 for a discussion of this weakness. On the account given here, the only place a trace can be introduced is in a construction whose left-hand side is still incomplete; inverted sentences do not meet this criterion, and cannot be introduced by PS7. so there could not be a subject trace in an inverted sentence.

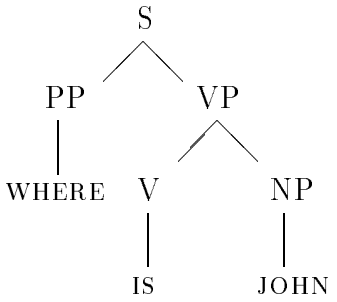
For completeness, I note that some constructions involving the copula which would be treated in GKPS as containing a trace will be handled more simply on the approach I sketch here. Since by hypothesis the only mechanism in English for introducing traces is by means of PS7 above, examples like those in (1) must not contain a trace.

(1) a. How tall is John?
b. Where is John?
c. How many unicorns are there in the garden?

Given an earlier assumption that lexical entries may override the default complement order which is based on obliqueness, the examples in (1) can be admitted by providing entries for the copulas which have the reverse order of the first two subcats in the Complements attribute, compared with the entries needed for simple declarative constructiosn. These two sets of entries for the forms of the copula are related by a lexical rule which I do not present here, sketching instead the PS tree

that results for example (1b), to illustrate the effect of the lexical rule.

(2) Where is John?



Since the set of WH-questions illustrated in (1) is limited to those with copular heads, and since there is no need for an unbounded dependency mechanism here, the lexical rule approach for these constructions seems markedly preferable to the obvious alternative of adding yet another phrase structure rule similar to PS7 which would introduce a trace for the non-subject complement of the copula.[8]

In most of the discussion in this work, the generalizations that would be captured via metarule in GKPS are expressed in this framework in terms of inheritance or lexical rule, but in this one case, a single phrase structure rule best captures the intent of the relevant metarule, preserving the notion that constraints on trace introduction are properties not of the lexicon, but of phrase structure. That is, lexical entries do not specify which of their complements can be extracted, and which cannot; the fact that the subject of the verb in English cannot be extracted is due to the fact that English is an SVO language, where the grammar causes the subject in a declarative sentence to be combined with its verb separately from the verb's other complements.

This account, like that of Flickinger 1983 and GKPS 1985, predicts that a language whose grammar allows a lexical predicate to pick up all of its complements (including the subject) with a single PS rule should also allow subject extraction.

[8]If one distinguishes the lexical entry for the predicative copula of *John is tall* from the identity copula of *John is the king of France*, then this analysis of copular constructions has the added virtue of allowing one to eliminate a spurious second parse for *Who is John*, by not allowing the lexical rule used for (1) to apply to the identity copula.

This prediction, which holds true for VSO languages like Irish and Welsh, is based on differences in the grammars of SVO vs. VSO languages, not differences in lexical representation.[9] Of course, not all languages will have a grammar rule like PS7 introducing traces, but for languages like Irish and English which do exhibit unbounded dependencies, the licensing of a trace at one end of the dependency is a function of the grammar, not of the lexicon.[10]

[9]For analyses of Welsh and Irish within a phrase structure grammar framework, see Harlow 1983 and Sells 1983.

[10]GKPS 1985 introduce two distinct metarules for "slash termination"; one of the two (STM1, p. 142ff) is essentially the Trace Introduction Metarule proposed in Sag 1982, and is the one whose work is now to be done by PS7 proposed here. The other metarule for slash termination (STM2, p. 160ff) accounts for a set of extraction phenomena which are lexically constrained, and which can and should be accounted for by lexical rule, as proposed in Pollard 1985. In a later section I will take up Pollard's proposal in some detail.

Chapter 5

Lexical Rules

The word class hierarchy and associated notion of inheritable information that were presented in the previous chapters provide powerful tools for describing one kind of shared structure in the lexicon, eliminating redundancy that is category-wide in nature. These tools are helpful, for example, in distinguishing that which is common to all verbs from that which is unique to some subclass of verbs, but still shared by all (or most) members of that subclass. In order to capture a second type of generalization in the lexicon, I develop in this chapter an account of lexical rules, familiar in much other current work on the lexicon, but here allied with the notions of word class and inheritance to provide a more flexible framework, one which will extend to lexical explanations of phenomena once believed to be outside the scope of the lexicon.

The key idea on which this chapter is based is the notion that a lexical rule represents a systematic relationship holding between two word classes, or more precisely, between the members of one class and the members of another class. As I illustrate below, each word class can include, as part of its definition, information about relevant lexical rules that relate members of that class to members of other word classes. This information can also be part of an actual lexical entry, since an entry inherits and can override any of the properties of the word classes it belongs to. In associating a lexical rule with a class or lexical entry, any one of three types of information may be specified: 1) the rule is simply applicable; 2) the rule applies,

but with morphological or semantic idiosyncracies stipulated by the entry or class; or 3) the rule is not applicable.[1]

In the course of illustrating both this general notion and the particular machinery I employ to express the notion more precisely, I indirectly address the issue of what kinds of relationships among words merit capture via lexical rule, whether morphological, syntactic, or semantic. After sketching the extent of applicability, I also suggest a partial classification of lexical rules along several dimensions, some familiar from previous work (e.g., major/minor, inflectional/derivational, same-arity/changed-arity). Central to this account is the view that while lexical rules can be grouped along these dimensions, they are instances of one formal mechanism, whether they express generalizations about inflectional morphology or derivational morphology, or about phenomena traditionally viewed as syntactic, not lexical. In each case, a rule simply expresses a cluster of regularities holding between respective members of two sets of lexical entries, drawing on the definitions of word classes given in Chapter 2.

Structure of the chapter

I begin in section 5.1 with a general characterization of a lexical rule in terms that are consistent with the representation of lexical entries developed in the previous chapter, and I establish the connection between lexical rules and word classes. Section 5.2 addresses the three types of exceptional information that a class or entry may specify for a given lexical rule, and provides an account of *blocking* phenomena using these mechanisms. In section 5.3 I take up the question of the kinds of relationships properly expressed with lexical rules, distinguishing the work of lexical rules from that of inheritance and from that of phrase structure rules.

[1]This three-way distinction should not be confused with G. Lakoff's (1970) non-isomorphic set of distinctions labelled *positive, negative,* and *absolute* exceptions to transformations; cf Green's (1974:75) criticism of Lakoff's divisions.

5.1 On the Content and Form of Lexical Rules

I take a lexical rule to be the expression of a systematic (but not necessarily exceptionless) relationship holding between two sets of lexical entries, where each of the two sets is defined in terms of word classes as presented in Chapter 2. In the simplest case, a lexical rule relates the members of one word class to the members of a second word class, but more generally a set of lexical entries affected by a lexical rule may be described as the intersection of two or more word classes, as will become evident below. Given a word belonging to the first set, then, a lexical rule predicts the existence of a corresponding word belonging to the second set, with the differences and similarities between the two words captured both in the formulation of the rule, and in the definitions of the classes of each word.

Abstractly, a lexical rule will be of the form given in (1), where each of the word classes is of the sort presented in the previous chapter. In the absence of evidence to the contrary, I assume that lexical rules are bi-directional, allowing either entry of a related pair to predict the existence and properties of the other entry.

(1) General form of a lexical rule

Set 1					Set 2	
Classes:	A1		M1		Classes:	A2
Spelling:	B1		M2		Spelling:	B2
Phonology:	C1	$\Longleftarrow$	M3	$\Longrightarrow$	Phonology:	C2
Syntax:	D1		M4		Syntax:	D2
Semantics:	E1		M5		Semantics:	E2

The content of a rule can be viewed as having two parts: (1) identifying the sets of lexical entries being related, by specifying the class or classes which define each of the two sets; and (2) defining the mappings M1 - M5, which relate the spelling of Set 1 members to the spelling of Set 2 members, the phonology of Set 1 members to that of Set 2 members, and likewise for each of the other kinds of information. Of course, each of these mappings M1 - M5 may be the identity relation in a given rule, though not all of them in any one rule.

In expressing the relationships between members of two sets of lexical entries, I make crucial use of the distinction between idiosyncratically specified information (which appears in a non-redundant, minimally specified lexical entry) and inherited information. I adopt here the restrictive hypothesis that lexical rules hold for minimally specified lexical entries, without having access to inherited, predictable information. Adopting this hypothesis imposes a constraint on the form and function of lexical rules which is strong, but not too strong, allowing a simpler formulation of rules by keeping to a minimum the amount of information to be managed. Only two kinds of information are relevant for a lexical rule: the word classes that each of the two related entries belong to, and any idiosyncratic properties specified by either lexical entry.

By applying only to minimally-specified forms, a lexical rule can also guarantee that the two forms will share any idiosyncratic properties not relevant to the rule, by enforcing identity as the default relation for any attributes not mentioned in the rule. This correct prediction would be much harder to make if lexical rules applied to fully-specified entries, and had to account for all of the predictable differences between the two related forms, differences which are the result of the two entries belonging to different classes.

The applicability of a lexical rule is constrained in two additional ways, one external to the rule, the other internal. First, the name of a lexical rule must apppear as the value of the Lexical-rules attribute for two word classes in the hierarchy of the previous chapter (two, because a lexical rule is always a two-place relation). It is for members of these classes that the rule holds (though there may be exceptions within either class). Thus for any given lexical entry, the set of potentially applicable lexical rules is inherited in the same way that other shared properties are, so a lexical entry can specify idiosyncratic information about any of the applicable rules, as will become clear in the next section. Second, the rule may itself specify further constraints on applicability, effectively reducing its applicability to only a proper subset of a class that introduced the rule. In general these constraints are expressed in terms of word classes, as illustrated with several of the rules introduced below.

To simplify the formulation of lexical rules, I employ two conventions. First,

where the relationship between corresponding properties of two related entries is identity for a given lexical rule, that rule will make no mention of those properties. Second, I will identify only implicitly the two sets of lexical entries related by a rule, by mentioning the relevant parent classes in the first statement of the rule, the one stating how the list of parent classes for the one entry is related to the list of parent classes for the other. Both of these conventions are rehearsed in the example below.

I introduce the notation for lexical rules with an example of a relatively simple lexical rule, the inflectional rule relating the base forms of verbs with their past tense forms. Making use of the word classes sketched in the previous chapter, I can express the relationship between, say, *walk* and *walked* as in (2), leaving out specifics of phonology and only hinting at semantics. For convenience, in cases where a class that introduces the rule is not mentioned directly in the rule's statement about classes, I follow the rule's name here with the name of that word class.

(2) LR-PAST lexical rule

LR-PAST		
LE2-Classes – PAST	=	LE1-Classes – BASE
LE2-Spelling	=	(AFFIX-ED LE1-Spelling)
LE2-Phonology	=	...
LE2-Semantics	=	(PAST LE1-Semantics)

This lexical rule, like any other, expresses a relation holding between two sets of lexical entries, the first set represented by a canonical lexical entry LE1, and the second set by LE2. The rule's applicability is governed by the relevant classes that LE1 and LE2 each belong to, with these classes named in the statement within the rule that relates the one entry's list of parent classes with the other entry's class list. Having specified the range of applicability, each rule then states the particular dependencies holding between properties of LE1 and corresponding properties of LE2.

The LR-PAST rule expresses a relation holding between members of the BASE word class and members of the PAST word class. It is thus a rule which may apply to every verb in English which has a base form or a past form, regardless of

the number of arguments the verb requires, and regardless of whether the verb is an auxiliary or not. (I take up the issue of exceptions to applicability in section 3.2.) To some extent, the decision to relate the past form of a verb directly to the base form (rather than, say, to the past-participle or even to some archi-entry) is arbitrary, though it has interesting consequences. One benefit of this decision is that it simplifies expression of some of the dependencies (discussed below), and allows a more symmetrical expression of the various inflectional rules, since each expresses a relation between the base form and some one other form of a verb. A possible objection is that on this approach there is no direct mapping between a present tense form like *walks* and the past tense form *walked*; rather, the relation between the two is expressed indirectly, in terms of the two separate relations that link, on the one hand, *walks* and *walk*, and on the other hand *walked* and *walk*. Since no adverse effects seem to result from this indirection, and since it allows some regularization in expression of rules, I will assume that in a paradigm like that for verbal inflection in English, some one member of the paradigm will serve as the anchor for the other members (analogous to the use of "hub cities"[2] by airlines in scheduling inter-city flights).[3]

The first statement in this rule, as in every lexical rule, links the set of classes that LE1 is directly a member of with the set that LE2 is directly a member of. With the informal notation I have adopted here, the rule says the classes that LE1 belongs to are the same as those that LE2 belongs to, except that where LE1 is a member of the BASE class, LE2 is a member of the PAST class. For example, if we take LE1 to be the entry for the base verb *walk*, then the LE1-Classes will be the set (BASE INTRANSITIVE MAIN-VERB), LE2 will be the entry for the past tense verb *walked*, and this lexical rule states that the classes *walked* belongs to (which I

[2]I am grateful to T. Wasow for the analogy, though he may choose to disclaim the idea.

[3]An alternative approach would permit lexical rules to relate a class of actual lexical entries with a class of archi-entries. To allow this kind of relation would weaken the notion of a lexical rule, since one side of the relation would no longer be anchored in actual lexical entries whose properties can be as easily verified. I have therefore adopted the hypothesis that the relevant generalizations can be expressed without depending on such archi-forms. One ready example that might drive one to the use of archi-forms is the verbal paradigm for a Romance language like Spanish, where the stem constant in each of the inflected forms does not itself appear as an independent form. I do not pursue this issue here.

abbreviate LE2-Classes) must be the set (PAST INTRANSITIVE MAIN-VERB). This may seem obvious, but in some lexical rules the mapping between the classes of LE1 and those of LE2 is not so simple, as will be seen.

The second statement in (2) links the spelling of LE1 with that of LE2, stating that the default relationship is one I have simply abbreviated AFFIX-ED, which would properly be fleshed out as the relevant (non-trivial) spelling rule that handles affixation of the suffix -ed, mentioning both rule-specific and general English spelling conventions. This default spelling can be overridden by particular lexical entries, as discussed in section 3.2. Not all lexical rules have a related change in spelling for the two forms; if this rule had said nothing about spelling, it would mean that as a default the two entries LE1 and LE2 would have the same spelling.

A relationship analogous to the AFFIX-ED one for spelling will have to be specified for the phonological properties of the two lexical items. More generally, each lexical rule which is not morphologically transparent will specify some default morphological rule which relates the phonological forms of the two entries. Each of these morphological rules will generally have two parts, one specifying the relevant affix (or affixes), and the other specifying the operation used to combine affix and stem. In English, the operation is almost always simple concatenation of affix and stem, with the only variable being whether the affix is a prefix or a suffix. I leave until Chapter 7 a fuller discussion of these morphological rules, and how inflectional and derivational affixes are represented in the lexicon.

Similar in form to the statement in rule (2) about spelling is the one relating the lexical semantics of LE1 and LE2. The rule says the default relationship is one I have abbreviated PAST-TENSE, standing in for the (again non-trivial) relevant semantic rule that connects, for example, the meaning of *walk* with the meaning of *walked*.

Equally important are the things this lexical rule does not say, given the convention that identity is assumed for any properties not mentioned. Since the lexical rule says nothing, for example, about the syntactic features of either entry, any idiosyncratic feature values specified in one must be identically specified in the other. Note here that this does not mean each of the syntactic features in the fully specified

entry for *walk*, for example, will have exactly the same values as the features in the fully specified entry for *walked*. Since *walk* inherits values for some of its features from BASE, while *walked* inherits some values from PAST instead, the resulting sets of values for at least some features of the two entries will be distinct (e.g., in the value for the feature VFORM). But these important differences are handled by the inheritance mechanisms already introduced in the previous chapter, so the lexical rule does not need to mention them.

Notice that if lexical rules were insensitive to the distinction between idiosyncratic and predictable properties of lexical entries, the statement of even a simple rule like LR-PAST would be much more difficult. If the lexical rule for past tense verbs had to cope with fully-specified entries that blurred this distinction, it would be difficult to express in the rule just which properties of the one entry had to match in the related entry. For example, the verb *like* idiosyncratically requires a verbal complement which is either infinitival or gerundive, while the verb *enjoy* does not allow the infinitival form, allowing only the participial form for its complement, as illustrated in (3-4).

(3) a. John likes to make noise.
b. John likes making noise.

(4) a. *John enjoys to make noise.
b. John enjoys making noise.

Since all of the inflected forms of *like* allow the same choice of two permissible forms for the verbal complement, while all of the inflected forms of *enjoy* insist on the participial complement, the lexical rules like LR-PAST or the similar one for present-third-singular forms must preserve these idiosyncracies. Yet a fully-specified entry for the base form *enjoy* stipulates not just the form of the complement, which would have to be identical in the present third-singular entry *enjoys*; the fully-specified base entry for *enjoy* also specifies that its subject be unmarked for number, an indifference which crucially must not be shared by the entry for *enjoys*. Short of tagging each attribute value in a fully-specified entry as local or inherited, it is not clear how the lexical rule for present-third-singular forms could be constrained to ensure identity of the verbal complement's VFORM value while ignoring differences

in the subject's AGREEMENT value for these two entries for *enjoy*. In sharp contrast, this difference in idiosyncratic vs. inherited information can be exploited by lexical rules without stipulation when they are constrained to apply only to minimally-specified entries.

5.1.1 Lexical rule for passive

In order to see the interaction between inheritance and lexical rules more clearly, let us consider another example of a lexical rule, the passive rule, presented initially as (3).

(3) LR-PASSIVE lexical rule

LR-PASSIVE		(TRANSITIVE)
LE2-Classes – PASSIVE	=	(CHANGE-ARITY LE1-Classes) – PAST-PARTICIPLE
LE2-Subject	=	LE1-Object
LE2-PP-By	=	LE1-Subject

This rule, which is introduced in both the TRANSITIVE class and the PASSIVE class, holds for the set of lexical entries belonging to the TRANSITIVE class, but is further restricted within the rule to those which are also members of the PAST-PARTICIPLE class, relating them to corresponding members of the PASSIVE class. Unlike the inflectional rules such as the past-tense rule in (2), the passive rule makes reference not to the BASE subclass of verbs, but rather to the PAST-PARTICIPLE subclass, since passives have all of the properties specified in the PAST-PARTICIPLE class. For convenience, I repeat the definition of the PASSIVE word class as (4).

(4) PASSIVE word class

PASSIVE	
Superclasses	PAST-PARTICIPLE
Atomic-features	(VFORM Passive) (PREDICATIVE +)
Complements	PP-By
PP-By-features	(CAT Preposition) (COMPLETE –) (PFORM By)

Most of the work of this rule (as with most lexical rules) is handled in the statement of the dependency between the classes that LE1 is a direct member of, and the classes that LE2 is a direct member of. In this rule the relationship is somewhat more complicated than it was for the rule in (2). In addition to exchanging PASSIVE for PAST-PARTICIPLE in the set of parent classes, analogous to the exchange of PAST for BASE in (2), the passive rule specifies a second, non-trivial exchange which I have abbreviated CHANGE-ARITY. Given the hierarchy as presented in Chapter 2, this exchange connects the COMPLEMENTATION subclass that LE1 mentions directly, and the one that LE2 mentions directly. For convenience I term these C-1 and C-2, respectively, and express the CHANGE-ARITY relation informally in (5).

(5) **CHANGE-ARITY Relation**

C-2 is the immediate superclass of C-1 distinct from TRANSITIVE.

Thus, if LE1 is the entry for *devour*, which belongs directly to TRANSITIVE, the corresponding passive LE2 for *devoured* belongs directly to INCOMPLETE. If LE1 is the entry for *persuade*, a direct member of OBJECT-EQUI, then the passive LE2 for *persuaded* belongs directly to EQUI. In each case, the passive is not a member of the TRANSITIVE class, so it will not have a direct object. In more detail, the entry for the passive *persuaded* will be a direct member of the EQUI class, which inherits from CONTROL and from INCOMPLETE, so the Complements attribute of *persuaded* will include a controlled complement (from CONTROL) and a subject (from INCOMPLETE). The structure of the hierarchy permits a straightforward expression of the relation holding between the complement requirements of an active verb and those of its passive counterpart.

The other two statements in the passive rule in (3) identify the active form's object properties with those of the passive form's subject, and likewise for the active's subject and the passive's oblique argument, here labelled PP-By. Each such statement ensures that the thematic role and any idiosyncratic syntactic properties assigned by the active to one subcat are assigned by the passive to its corresponding subcat. In English there do not appear to be any transitive verbs which assign

idiosyncratic syntactic properties to their object, so the sole effect of these two statements in (3) is to capture the thematic role relationships between active and passive.

5.1.2 Passives and idiosyncratic case in Icelandic

The passive construction in Icelandic provides another good illustration of the way in which inheritance interacts to good effect with the lexical rule mechanisms proposed here. From the work of Andrews 1976, 1982, Thrainsson 1979, Maling 1980, and Zaenen 1980, it is clear that many verbs in Icelandic assign idiosyncratic case to subjects as well as objects. What is of particular interest is that if an active transitive or ditransitive verb assigns an idiosyncratic or *quirky* case to its direct object, the passive form of that verb assigns the same quirky case to its subject. Three examples given in (7-9) are taken from Zaenen, Maling, and Thrainsson 1985, who provide a good summary of the relevant arguments for treating these non-nominative arguments as subjects.

(7) a. Ég hjálpaδi honum.
I helped him(D)
I helped him.

b. Honum var hjálpaδ.
him(D) was helped
He was helped.

(8) a. Ég mun sakna hennar.
I will miss her(G)
I will miss her.

b. Hennar var saknaδ
her(G) was missed
She was missed.

(9) a. Ég skilaδi henni peningunum.
I returned her(D) the-money(D)
I returned the money to her.

b. Henni var skilaδ peningunum
her(D) was returned the-money(D)
She was returned the money.

This assignment of quirky case to subjects of passives is just what is expected for Icelandic, assuming that its rule for relating actives and passives is essentially the same as that in (3) above (modulo the naming of the oblique subcat). The lexical entry for the active *hjálpaδ* specifies dative case for its direct object subcat, so that same case will be preserved in the corresponding passive entry's subject, lexically overriding the default (nominative) case ordinarily assigned to subjects.

In fact, it would be difficult to have Icelandic passives behave otherwise with respect to quirky case assignment, given the treatment of lexical rules presented here. Since quirky case is by hypothesis assigned by a particular lexical item, that idiosyncratic information will be preserved in other lexical entries related to the first by lexical rule, so idiosyncratic case introduced for a subcat in one lexical entry will necessarily be introduced in lexically related entries. This constraint on possible lexical relationships is supported by the Icelandic data.

The approach taken here is thus consistent with the position adopted by Zaenen, Maling, and Thrainsson 1985, who argue (p. 466) that "idiosyncratic case is assigned to thematic roles and not to grammatical functions". More precisely, idiosyncratic case is assigned to a particular subcat by a lexical item, which also assigns a thematic role to that subcat. However, the terminology they use to distinguish kinds of case assignment is not helpful, since they treat "lexical" and "idiosyncratic" as interchangeable, and distinct From "default" case assignment. It should be clear that within the framework developed here, both idiosyncratic and default case assignment occur within the lexicon, with the relevant distinction made in terms of word classes vs. individual lexical entries.[4]

5.1.3 Lexical rule for To-datives

By way of illustration I offer one more lexical rule here, one which raises several issues that I take up in more detail in the course of this chapter. This rule is intended

[4]I have only touched on the wealth of relevant data on Icelandic passives summarized by Zaenen, Maling, and Thrainsson, who concentrate on the case assignment properties of several classes of ditransitive verbs. A full treatment of the data they present, and an account of how their proposed association principles could be expressed in terms of the mechanisms described here, would be instructive, but will have to await further work.

to capture the familiar relation between pairs of entries for ditransitive verbs like the two varieties of *give* shown in (6), where one entry has two NP complements (in addition to the subject), while the other has an NP and a PP introduced by the preposition *to*.[5]

(6) a. Sally gave John a book.
b. Sally gave a book to John.

To relate such pairs of entries, I propose the To-Dative lexical rule given in (10).

(10) LR-TO-DATIVE lexical rule

LR-TO-DATIVE		
LE2-Classes – DITRANS-TO	=	LE1-Classes – DITRANS
LE2-DObject	=	LE1-IObject
LE2-IObject	=	LE1-DObject

This rule differs from the passive rule in some interesting ways. First, it is well known to be much more subject to exceptions than is the passive rule, with some verbs (like *donate*) insisting on a PP indirect object, and others (like *spare*) insisting on an NP. This raises the question of whether to make the rule hold as a default for the two classes of ditransitives, and lexically specify exceptions for those members which do not submit to the rule, or to not make the rule a default, instead specifying lexically just those verbs for which the rule holds. The presence of both variants for newly-formed dative verbs, illustrated in (11), argues in favor of making the rule a default for the two ditransitive classes.

(11) a. John faxed me the documents yesterday.
b. John faxed the documents to me yesterday.

A second way in which the dative rule differs from the two previous rules is that the dative rule never has a morphological effect. Many lexical rules have this

[5] For earlier treatments of this relationship, see Green 1974, Oehrle 1976, Dowty 1978, Bresnan 1978, 1982:25, Baker 1979, Wasow 1980:294-5. I do not attempt to provide here an account of "for-Datives" as illustrated in (i), even though the double-object forms of verbs like that in (ia) are indeed members of the DITRANS class; they will have to be related to forms like that in (ib) by a lexical rule which copes with the complexities of the (optional) benefactive FOR-PP, a rule which I assume will be distinct from the To-Dative rule given here.

(i) a. Sally bought John a book.
b. Sally bought a book for John.

property, but these must then be motivated on syntactic or semantic grounds with a care not so essential when there is morphological evidence for two distinct but related lexical forms.

This rule for relating ditransitive verbs does have in common with the passive rule a systematic change in the mapping between thematic roles and grammatical functions. Given that both entries for *give* have a direct and an indirect object, with the indirect more oblique than the direct object, and given the linear precedence constraint which orders less oblique elements before more oblique ones, the exchange effected by the lexical rule is essential. If the entry for the DITRANS-TO *give* in (12a) assigns the PATIENT role to its direct object *a book* and the RECIPIENT role to its indirect object *to Mary*, then the entry for the corresponding DITRANS *give* of (12b) must assign the RECIPIENT role to its direct object *Mary*, and the PATIENT role to its indirect object *a book*.[6]

(12) a. John gave a book to Mary.
b. John gave Mary a book.

In sum, this rule captures the two alternations observed in a pair of related ditransitive verbs; the first of these is the syntactic form of the indirect object (NP vs. PP), which follows from the differing default syntactic assignments made by the DITRANS and DITRANS-TO classes for the indirect object. The second alternation must be dealt with explicitly in the rule, since it involves a change in the assignment of thematic roles, which I take to be lexically idiosyncratic. Note here that this rule, like the passive rule, does not violate our assumption that lexical rules

[6]There are examples which strongly suggest that this dative rule's applicability is subject to semantic constraints which I have not attempted to include in the rule's definition. Consider the pair in (ia,b) (cf. Oehrle 1975 for these and similar examples).

(i) a. John gave the teacher a headache.
b. *John gave a headache to the teacher.
c. John gave an apple to the teacher.

One way to exclude (ib) might be to say that *give* has (at least) two distinct meanings, and that the dative rule is sensitive to that distinction, however it be made precise. The two different senses here are subtle, but the one involves a transfer of some sort, as in (ic), while the other sense involves causality, but not transfer. At the very least, the entailments are different, since (1c) entails that John had an apple, but (1a) does not entail that John had a headache. This data raises important issues involving both the proper formulation of lexical rules, and the permissible constraints on rules; but I do not pursue them here.

apply only to minimally-specified entries. Indeed, it is precisely because thematic role assignments are lexically idiosyncratic, and therefore present in the minimal entry, that they need to be re-assigned by the lexical rule.

The To-Dative rule interacts in some important ways with the Passive rule, providing a convenient forum for further explication of the connections between word classes and lexical rules, but before taking up this discussion in section 5.3, I present in 5.2 the mechanisms for handling exceptions of several types involving lexical rules.

5.2 Exceptional Properties

A lexical rule expresses a relationship holding between members of two sets of lexical items, with the basic constraints on set membership being provided by the word classes introducing the lexical rule. However, few if any lexical rules are completely exceptionless in applying to members of the relevant word classes. The inflectional rules may have the fewest exceptions (ignoring morphology for the moment), but even the past-tense rule does not hold for all base verbs; consider the verb *beware* illustrated in (1), which does have a base form but does not have a past form (or indeed any other inflected form).[7]

(1) a. John was told to beware of the lions.
b. Beware of the lions!
c. *John bewared of the lions.

The passive rule also holds for the vast majority of transitive verbs, but also has some exceptions, such as *resemble*, *suit*, and the *have* of possession. The rule for To-Datives, on the other hand, admits of a large number of exceptions among ditransitive verbs. In those cases where a rule holds quite generally for the class, I introduce it as a default property of that class, entering the name of the rule as a value of the Lexical-rules attribute that I included in the definition of the top-level word class in the previous chapter. The mention of a lexical rule in a class definition means that unless a lexical item belonging to that class says something to the contrary, the generalization expressed in the rule holds for that item.

I first illustrate the notation for linking rules and classes, ignoring exceptions, and show how this information about lexical rules interacts with the inheritance mechanism to provide as part of a lexical entry the specification of those lexical rules which hold for that entry. I then present three kinds of exceptional information that classes or (more often) individual entries can provide about lexical rules, and provide a notation for expressing such information.

[7]Note that the lack of past-tense forms for *to* and the modals is correlated with the absence of a base form for these same verbs. What must be stipulated as exceptional about the modals is that the inflectional rules relating present-tense forms with base forms do not hold.

5.2.1 Linking of lexical rules and word classes

The class definition for the PAST class is given below as it was defined in section 2.2, but now with the additional information about lexical rules included in the definition.

PAST	
Superclasses	FINITE
Lexical-rules	LR-PAST

The definition for the BASE class will also mention LR-PAST, along with the names of the other inflectional rules, given my assumption that the base form serves as the "hub" for the inflectional paradigm. I include in the following definition only a couple of additional rule names, by way of illustration, without attempting to specify and defend a complete list of these inflectional rules.

BASE			
Superclasses	VERB-FORM		
Atomic-features	(VFORM Base) (PREDICATIVE –)		
Lexical-rules	LR-PAST	LR-PRES-3RD-SG	LR-PAST-PART

Since lexical entries can inherit information from more than one class, any given entry may specify participation in more rules than appear in any one class definition. For example, a simple transitive base verb like *devour* will include in its set of lexical rules not only all of the inflectional rules mentioned in the BASE class, but also at least the LR-PASSIVE rule which will appear as part of the definition of the TRANSITIVE class. Since each rule is defined as an independent regularity holding for two sets of lexical entries, without any notion of sequential application, no difficulties arise about one rule "feeding" or "bleeding" another.[8] Rather, each rule appearing in a lexical entry simply predicts the existence of a related lexical entry,

[8]This terminology was introduced by Kiparsky 1968 for phonological rules, and adopted in part by Wasow 1977 for describing the applicability of lexical rules and transformations. Given sequential application of rules, for a rule A to feed a rule B means that rule A applies before rule B and transforms its input in such a way that its output is of the right form so rule B can apply. For rule A to bleed rule B means that rule A again applies before rule B, but transforms its input so that rule B cannot apply to the output of rule A even though rule B could have applied to the input of A, if it could have gotten there first.

and also predicts those properties of the second entry which depend on properties of the first entry.

5.2.2 Exceptions

For each lexical rule that is mentioned in a lexical entry, any of three types of idiosyncratic information can be specified to affect the way in which the rule applies to that entry - information which is not predictable from the classes that entry belongs to. First, the entry might simply specify an additional rule which applies idiosyncratically to that entry, though not generally for any of the classes the entry is a member of. Second, the entry may block the application of a rule which otherwise holds for some class the given entry is a member of. And third, the entry may specify unpredictable phonological, orthographic, or semantic properties of the other entry related to it by a given lexical rule.

Positive exceptions

The first type of exception, where a rule applies to some entry but not to most members of its class, is rare if it exists at all. A convincing example would involve finding a pair of lexical items which clearly did not belong to the classes for which a lexical rule held, but nevertheless exhibited the relation described by the rule. I do not know of any convincing example of this kind.

A more vexing illustration of positive exceptions might involve a rule with very limited applicability, one which held only for a few pairs of lexical items, but not even for most members of the classes these items belonged to. One seemingly plausible candidate to illustrate this kind of positive exception involves pairs of entries like those for the verb *open*, illustrated in (2).

(2) a. The door opened.
b. Sally opened the door.

Pairs like this can be related by a causative lexical rule, as proposed in Jackendoff 1975 (cf. Wasow 1977:331-3). I formulate this rule in (3), and note that it has much in common with the passive rule given in the previous section, though it is restricted

to immediate members of the TRANSITIVE class, as indicated here by the mention of the TRANSITIVE class itself in the first statement of the rule.

(3) LR-CAUSATIVE lexical rule

LR-CAUSATIVE		
LE2-Classes – TRANSITIVE	=	LE1-Classes – INCOMPLETE
LE2-DObject	=	LE1-Subject
LE2-Semantics	=	(CAUSE-TO-BE agent: LE2-Subject-index situation: LE1-Semantics)

What this rule says is that the lexical entry for an intransitive verb like *open* as in (2a) is identical to the entry for the transitive *open* of (2b) except in three respects: first, there is the obvious substitution of the class TRANSITIVE for the class INCOMPLETE; second, the thematic role assigned to the subject of the intransitive is the same as that assigned to the direct object of the transitive; and third, the semantics of the transitive introduces some notion of causality.[9]

The examples in (4-7) taken from Wasow 1977:333 show that there are a number of apparently unpredictable exceptions to the causative rule, at least in one direction: many transitive verbs which seem to have the right semantic properties fail to have intransitive counterparts.

(4) a. John dropped the rope.
b. John lowered the rope.

c. The rope dropped.
d. *The rope lowered.

[9]In formulating a rule like this, one glimpses just over the horizon the hoary head of Generative Semantics, for one might well ask whether the rule in (3) should also be used to relate the entries for, say, *kill* and *die*, even though they do not share a common morphological base. Before dismissing the notion too quickly, note that we surely wish to use the same inflectional rule for past-tense verbs to relate not only *walk* and *walked* but also *go* and *went*; so it doesn't seem that a common morphological base is a prerequisite. On the other hand, to relate *kill* and *die* with the rule in (3) has the feel of opening Pandora's box; what criterion can be used to exclude this use while allowing for suppletion? Jackendoff 1975:651 imposes the requirement that "items linked by a lexical rule show morphological relatedness," but does not propose to treat inflection with lexical rules. Perhaps the requirement of morphological relatedness, however formulated, should hold for some classes of rules but not others, serving as one criterion for identifying such classes (e.g., not needed for inflection, but needed for derivation).

(5) a. John shattered the light bulb.
b. John demolished the light bulb.

c. The light bulb shattered.
d. *The light bulb demolished.

(6) a. John darkened his hair.
b. John tinted his hair.

c. His hair darkened.
d. *His hair tinted.

(7) a. We moved the boxes.
b. We transported the boxes.

c. The boxes moved.
d. *The boxes transported.

However, the mere existence of exceptions is not reason enough to treat a rule as completely idiosyncratic; first, it may indicate that the rule needs to be further constrained, in this case semantically. Second, it might prove to be the case that the exceptions, while numerous, are still in the minority, in which case it would seem preferable to introduce the rule as a default for transitives. In the worst case, one might have to determine whether the exceptions were indeed the majority, to decide whether or not to introduce the rule in the class. Such tallying seems an unwieldy way to decide issues of representation; I adopt instead the approach that if a regularity like that illustrated in (2) holds for several lexical items, the rule should be introduced in the class common to those lexical items, and blocked on the other members of that class, ignoring the issue of democracy.

If no clear examples of positive exceptions can be found, the framework presented here could be strengthened by excluding this possibility entirely. I leave the issue unresolved here, and turn to the other two kinds of exceptions, both of which are clearly needed.

Negative exceptions

Since the To-Dative rule relates members of the DITRANS class with corresponding members of the DITRANS-TO class, and is here assumed to be a default rule

for members of both classes, the name of the rule will appear in the Lexical-rules attribute for each of the two classes. It will therefore also appear as part of the lexical entry for each member of either class, unless the lexical entry says something to override the default. Clearly, the entry for the verb *donate* needs to say something special in order to exclude sentences like (8b).

(8) a. Sally donated a book to the library.
b. *Sally donated the library a book.

To mark the fact that *donate* does not have a double-NP alternate form as the rule would lead one to expect, I annotate the mention of the lexical rule with the tag *Not-Applicable*, as in the following partially specified entry for the base form *donate*.

DONATE-1-BASE	
Superclasses	MAIN-VERB, BASE, DITRANS-TO
Spelling	"donate"
Phonology	...
Semantics	(DONATE agent:X patient:Y goal:Z))
Complements	
Subject-index	X
DObject-index	Y
IObject-index	Z
Lexical-rules	(LR-TO-DATIVE Not-Applicable)

This entry predicts the existence of a passive form *donated* as well as the various inflectional forms such as the past-tense *donated*, but states that there is no double-object form like that in (8b) above.

A second example of a lexical item refusing to participate in a relevant lexical rule is provided by the verb *get*, which does not have a corresponding passive form, unlike similar verbs such as *persuade*.

(9) a. Mary will get Bill to dance.
b. *Bill was gotten to dance.
c. Mary will persuade Bill to dance.
d. Bill was persuaded to dance.

Since the verb *get* is otherwise like its fellow object-equi verbs, all of which have passive forms, the lack of a passive for *get* appears to be an idiosyncratic gap. This

exception in the applicability of the passive lexical rule (introduced in section 5.1) can be represented here as a part of the lexical entry for the object-equi *get*, given in its minimally specified form in (10).

(10) Minimal lexical entry for *get*

GET-1-BASE	
Superclasses	MAIN-VERB, BASE, OBJECT-EQUI
Spelling	"get"
Phonology	...
Semantics	(GET agent:X patient:Y goal:Z))
Complements	
Subject-index	X
DObject-index	Y
XComp-index	Z
Lexical-rules	(LR-PASSIVE Not-Applicable)

Exceptional morphology/semantics

The third type of exceptional information a lexical entry may specify about a given lexical rule involves cases where the rule holds as a default for this entry as for others of its class, but where some aspect of the related entry is unpredictable and must be stipulated. Clear examples occur with the various inflectional rules, where the usual affix marking the inflection is altered or replaced entirely, as in the past-tense form *sang* for the verb *sing*. The default -ED suffixation rule must be blocked, with the phonology and spelling of the past-tense form stipulated instead in the base entry. At least three properties of lexical entries can be so stipulated: phonology, spelling, and semantics. It may also be that cases exist where the syntactic features of one form cannot be completely predicted from the other for some rule and some particular form, but I don't know of such a case. So I present the notation in such a way that it could easily be extended to cope with syntactic idiosyncracy as well, should such cases arise. The notation is again straightforward: a lexical entry may annotate the name of a particular rule with one of the tags *Irregular-Phonology*, *Irregular-Spelling*, or *Irregular-Semantics*, supplying in addition the unpredictable information which replaces the default.

An example of non-default semantics might be the case of the verb *walk* with the causative rule, where the meaning of the transitive *walk* in (11a) is less transparently tied to the meaning of the intransitive *walk* of (11b) than for the corresponding pair *march* in (12). If the idiosyncratic semantic property of the transitive *walk* can be formalized, that information would be supplied with the *Irregular-Semantics* tag for the LR-CAUSATIVE value in the lexical entries for *walk*.

(11) a. John walked the dog.
b. John walked.

(12) a. The sergeant marched his troops.
b. The troops marched.

As an illustration of the other two types of irregularity, note that *sing/sang* would require both phonology and spelling tags on LR-PAST in the lexical entries; an example where only the phonology tag would be needed is on the LR-PRES-3RD-SING rule for the verb *say*, which has the expected spelling *says* for its present-tense form, but an idiosyncratic pronunciation (cf. *do/does* compared with regular *go/goes*). In contrast, the past-tense form for a verb like *pay* has a completely regular phonology, but an idiosyncratic spelling *paid* (cf regular *pray/prayed*, *play/played*, *bay/bayed*). I illustrate the notation for such irregular properties by giving the partially-specified entry for the base form *pay* in (13).[10]

[10]I ignore in this entry the many subtleties of subcategorization for the verb *pay*, since my intent here is to illustrate its exceptional behavior with respect to the inflectional rule for past. I return to the interesting subcategorization properties of *pay* in the discussion of adjectival passives below. As given, this entry would account for (ia) below, but not any of (i.b-g); as usual, I assume that additional entries for *pay*, related by the appropriate lexical rules, would provide for these examples, similar to the way other familiar transaction verbs such as *buy* or *trade* are handled.

(i) a. John paid the grocer.
b. John paid for the bananas.
c. John paid twenty dollars.
d. John paid the grocer for the bananas.
e. John paid the grocer twenty dollars.
f. John paid twenty dollars for the bananas.
g. John paid the grocer twenty dollars for the bananas.

I am also not concerned here with the proper thematic role to be assigned to the direct object of *pay*; I use familiar labels such as *agent* and *patient* for thematic roles, though I have no theoretical commitment to some small, universal set of thematic role names. The issue involved here is not affected by just what the particular roles are, as long as *pay* does assign roles.

(13) Partially-specified lexical entry for *pay*

PAY-1-BASE	
Superclasses	MAIN-VERB, PAST, TRANSITIVE
Spelling	"pay"
Phonology	...
Semantics	(PAY agent:X patient:Y)
Complements	
DObject-index	Y
Subject-index	X
Lexical-rules	(LR-PAST Irregular-Spelling: "paid")

The lexical entry for *paid* will need a similar tag on the LR-PAST rule specifying that its corresponding base form has the spelling *pay*, on the assumption that lexical rules are non-directional, allowing the prediction of either entry from the other in a relation.[11]

Summary

This mechanism for exceptions to lexical rules gives a lexical entry a way of overriding default properties of lexical rules which complements the ability of an entry to override default properties of the classes an entry belongs to. What a lexical rule provides is a description of the default relationship between corresponding pairs of the two sets of related entries, but a particular pair may, as we have seen, be idiosyncratic in some respect while still submitting to the general correspondence. I have suggested three particular outlets for individual expression by a pair of lexical entries related by a lexical rule (phonology, orthography, and semantics), but I have not provided any assurance that these are the only types of idiosyncracy one will find. A fully explanatory theory of lexical rules should include an account of what

[11] One issue I have not dealt with is that of sub-regularities in rules such as the inflectional rules, where for some well-defined subset of verbs, a property such as the phonological form does not conform to the default, but is nonetheless rule-governed. For example, many of the strong verbs in English fall into a few subclasses with respect to the form of their corresponding past tense entries, where for each subclass a phonological regularity can be expressed that predicts the proper past tense form. I assume without further elaboration here that if such a sub-regularity is productive, it should be incorporated in the definition of the lexical rule; if not productive, then I do not have a mechanism to propose for capturing that sub-regularity.

the full range of exception types is, and why that catalogue is complete; but determining that range is in part an empirical question, outside the scope of the present study.

5.3 The function of lexical rules

In the introduction to this chapter I sketched a division of labor between the two mechanisms of inheritance and lexical rules, then in the preceding section identified one important kind of interaction between the two, to account for exceptions. Here I explore that division of labor in more detail, then develop a similar division separating the work of lexical rules from that done by syntactic rules.

5.3.1 Lexical rules and word classes

We might view inheritance within a hierarchy of word classes as a tool that eliminates redundancy along one dimension within the lexicon, while lexical rules provide the same service along another dimension. A given lexical item, by virtue of being a member of one or more word classes, shares inherited properties with other lexical items that belong to those same classes, but does not necessarily share a common morphological or semantic base (or indeed any of its idiosyncratic properties) with any of those other items.

That same lexical item, by participating in one or more lexical rules, has properties in common with a second set of lexical items, where the shared properties crucially include some or all of the idiosyncratic information which distinguishes that lexical item from others in its class. The members of this second set, related by lexical rules, all do share a single common semantic and morphological base (except for suppletions).

Of course, if a lexical rule relates two entries that both belong to a given word class (as happens with the verbal inflection rules), those two entries will share some inherited properties as well as the idiosyncracies. However, the lexical rule only establishes joint membership in that given class and the relationship of the idiosyncratic information in the two entries; all other properties shared by the two are established by inheritance within the word class hierarchy.

Both of these formal devices serve to express that which is common among (often overlapping) sets of fully-specified lexical entries, including properties that are morphological, syntactic, and semantic. In their capacity as redundancy mechanisms,

the two devices permit a parsimonious representation of the existing lexicon.

These two devices can also play a generative role: if a new transitive verb is added to the lexicon, the inheritance mechanism will assign the appropriate default properties to that verb, including the relevant lexical rules. These rules can then generate the related set of new lexical entries which share a morphological and semantic base with the new verb. Dowty 1981:80 makes this point about the dual roles played by lexical rules, though in the context of arguing that lexical rules differ from phrase structure rules only in that they play these roles. As will become clear, I draw a much sharper distinction between lexical rules and phrase structure rules than did Dowty.

5.3.2 Lexical rules and phrase structure rules

Phrase structure rules can also be viewed in either of two ways, either as tree admissibility conditions which license grammatical phrases, or as generative rules which produce the appropriate phrase structure trees for a given string of lexical items.[12] While they share with lexical rules this double life, the two types of rules have little else in common.

A lexical rule establishes a relation which holds for pairs of lexical items, predicting on the basis of one form the existence and content of the other, where either form may be used to predict the other, given the non-directional nature of lexical rules. Both members of the pair must be elements of the lexicon, so either member may also participate in other lexical rules, and of course, in phrase structure rules.

In contrast, a phrase structure rule of the sort introduced in section 2.4 establishes a relation between some single constituent and a set of lexical or phrasal constituents which help to determine the properties of that single constituent. The resulting constituent is not a member of the lexicon, so there can be no interaction between PS rules and lexical rules where the result of a PS rule participates in some lexical rule.[13] Thus a phrasal constituent may play a role in other PS rules, but not

[12] Cf. Gazdar 1981:156, following McCawley 1968.

[13] Cf. Wasow 1977:330 and Bresnan 1980:118, both of whom make this point quite clearly about the distinction between lexical and syntactic rules in how they interact with other lexical rules.

in any lexical rules.

But the differences between the two types of rules go well beyond the relation of the rule to the lexicon. As we have seen, lexical rules may permit idiosyncratic exceptions of several kinds, change assignment of thematic roles, or affect morphology. Phrase structure rules do none of these: No lexical item can be marked to ignore some phrase structure rule it would otherwise fit into, nor can any be marked to appear in a PS rule in spite of a conflict in syntactic properties.[14] No PS rule alters the thematic role assignment provided by the head of the phrase admitted or generated by the rule. And no PS rule alters the morphological form of its constituents; the constituents in syntactic rules must be combined without internal alteration to any of the constituents. That is, I take simple concatenation to be the only operation of syntactic combination permitted for PS rules.[15]

On the other hand, PS rules have some properties not shared by lexical rules: PS rules can apply recursively, but I take as a working hypothesis the position that lexical rules cannot, based on the observation that in each of the lexical rules studies here, the two related forms belong to distinct classes, precluding the re-application of the rule. This assumption forces a non-lexical analysis of constructions where the same affix can be attached more than once, with familiar examples including the English "anti-" prefix of "anti-anti-nuclear", and the Japanese causative suffix.

One other characteristic of lexical rules that may distinguish them from PS rules is their ability to establish complex semantic relations between two related forms. A lexical rule like the causative relates two lexical entries whose semantic properties are quite different, with the differences captured in the rule. It is at least an interesting hypothesis that PS rules, unlike lexical rules, never have idiosyncratic semantic effects. That is, it may be possible to maintain the constraint that the interpretation of the left-hand constituent of a PS rule is determined completely by the semantic properties of the right-hand constituents. Such a constraint would be

(Wasow draws the distinction between lexical rules and transformations, while Bresnan simply distinguishes lexical from non-lexical rules.)

[14]This general applicability of PS rules depends on barring the use of otherwise unmotivated diacritic features in lexical entries, a prohibition I assume here.

[15]This strong constraint on how PS rules can operate is by no means universally accepted; cf. the wrapping operations proposed by Bach 1981, Pollard 1984.

easy to satisfy in the simplest cases, where a head combines with a complement or an adjunct: the head can include as part of its subcat specification a description of the semantic operation to be used in combining the head's semantics with that of the complement or adjunct. Whether this strong constraint can be maintained for every phrase structure rule is a question I do not pursue further here.

Chapter 6

Adjectives and Unbounded Dependencies

In the previous chapter I provided a few examples of familiar lexical rules by way of explaining their form and function, their relationship to word classes, and their effect in phrase structure. To draw out some of the implications of this approach, and to provide the basis for a theory of constraints on lexical rules, I present in this chapter analyses of several less well-understood English constructions involving passives and unbounded dependencies. On the one hand, these analyses serve to further illustrate the interactions of the mechanisms I have proposed for representing lexical information, sometimes illuminating assumptions I have made but not discussed above. At the same time, the relative simplicity of the lexical rules proposed here and in the previous chapter will lead to the formulation of a restrictive theory of possible lexical rules, which I present in the final chapter.

Structure of the chapter

In section 6.1 I give an analysis of adjectival passives based on the recent insights of Levin and Rappaport 1986, showing that the relevant lexical rule does not require a constraint expressed in terms of thematic roles, as has been argued in the past. Sections 6.2 - 6.4 contain a lexical analysis of "tough-movement" adjectives,

one which extends easily to explain "too/enough" constructions and more complex constructions involving multiple gaps. In section 6.5 I expand on a proposal of C. Pollard's for a lexical account of "that-trace" or finite-VP facts, involving a second kind of lexical anchoring in unbounded dependencies.

6.1 Adjectival passives

To illustrate the nature of explanations provided by this account, I present an analysis of a familiar set of data involving verbal passives, adjectival passives, and datives. In constructing the analysis, I build on previous work with adjectival passives, done by Siegel 1973, Wasow 1977, 1980:294ff, Hust 1978, Bresnan 1982:16ff, Fabb 1984, Dryer 1985, and Levin and Rappaport 1986.

As Siegel 1973 and Wasow 1977 observed, there is a class of adjectives in English which have the same morphological form as verbal passives, but which are clearly distinct, showing the distribution of adjectives, not verbs. Both Wasow 1980 and Bresnan 1982, following Anderson 1977, propose a lexical rule which relates these adjectival passives to their verbal counterparts, with application of the rule constrained by the thematic role assigned to the subject of the verbal passive.[1] This thematic constraint, which requires the subject to be assigned the THEME role, was intended to account for contrasts like that in (1d-e).

(1) a. John sold used books to the students.
b. Used books were sold to the students.
c. The students were sold used books.
d. The used books remained unsold.
e. *The students remained unsold.

As Levin and Rappaport 1986 show, the relevant contrasts for adjectival passives can be accounted for without making reference to thematic roles at all, drawing instead on subcategorization properties of the verbal passives. In addition to providing greater descriptive adequacy, such an account allows us to retain a more restrictive theory of lexical rules, and avoids the difficulties involved in making precise the notion of thematic roles in general, and the THEME role in particular.

One set of examples which seem particularly awkward for the thematic account of adjectival passives is given in (2-3), where the adjectival passive of the verb *pay*

[1] Wasow's formulation relates the adjectival passive to the active verb, while Bresnan improves on this analysis by relating the adjectival passive to a class of intransitive participles which include the verbal passive. This latter formulation not only captures the morphological identity of the two passives (following Lieber 1980), but also offers an explanation for why the same thematic role is assigned to the verbal passive's subject and to the noun of which the adjective is predicated.

can appear as a modifier to nouns that must be assigned distinct thematic roles, making it hard for each to be the THEME.

(2) a. John did not pay his bills.
b. The bills were not paid.
c. The unpaid bills stacked up.

(3) a. John did not pay the waiter.
b. The waiter was not paid.
c. The unpaid waiters went on strike.

While posing a serious problem for the thematic constraint, these examples suggest that a quite different constraint is at work; to see the salient property of those verbal passives which do have a corresponding adjectival passive, consider the additional constrasts in (4 - 11):

(4) a. John did not send the letter to his sister.
b. The letter was not sent to his sister.
c. The letter was not sent.
d. The unsent letter sat on the table.

(5) a. John did not send his sister the letter.
b. His sister was not sent the letter.
c. *His sister was not sent.
d. *The unsent sister waited in vain.

(6) a. John did not hand the letter to his sister.
b. The letter was not handed to his sister.
c. *The letter was not handed.
d. *The unhanded letter stayed in John's pocket.

(7) a. John did not hand his sister the letter.
b. His sister was not handed the letter.
c. *His sister was not handed.
d. *The unhanded sister asked for the letter.

(8) a. John convinced the senators that he was right.
b. The senators were convinced that he was right.
c. The senators were (easily) convinced.
d. The convinced senators voted for the bill.

(9) a. John allowed the children to stay.
b. The children were allowed to stay.
c. *The children were allowed.
d. *The allowed children were happy.

(10) a. John placed the report on the table.
b. The report was placed on the table.
c. *The report was placed.
d. *The placed report has tall lettering on the cover.

(11) a. John misplaced the report.
b. The report was misplaced.
c. The misplaced report contained no surprises.

These contrasts are consistent with Levin and Rappaport's proposal that those verbs which have a corresponding adjectival passive are also those which have a verbal passive that is intransitive, having only one obligatory complement (the subject). It is this correlation which holds true for adjectival passives, rather than the earlier proposed correlation with thematic role assignment.

The rule in (12) expresses within the current framework the relationship between verbal and adjectival passives, making use of the insight provided by Levin and Rappaport. The rule states that for each intransitive passive verb, there is a corresponding adjectival form which has in its lexical entry the same idiosyncratic information that the verbal passive does, where the semantics of the adjective is based on that of the verbal passive. In particular, the thematic role that the verbal passive assigns to its subject will also be the role that the adjectival passive assigns, given the convention that two non-redundant lexical entries related by lexical rule will be identical except for those properties explicitly mentioned by the rule. This convention also ensures that the morphological form of the two passives will be the same, whether the verbal passive has an irregular form (like *taught*) or not.

In giving a preliminary formulation of this rule in (12), I accept without argument the need for lexical rules to be able to distinguish the immediate members of a class from those members which belong to subclasses of the class. From the discussion above, it should be clear that the adjectival passive rule only applies to strictly intransitive verbs, not to all members of the INCOMPLETE class, which

includes transitives, ditransitives, and so on. To mark this distinction between immediate members and more distant members of a class, I employ the notation of an exclamation mark following the name of a class. Thus the rule in (12) expresses a relation which holds only for immediate members of the INCOMPLETE class.

The other restriction on applicability of this rule is that the intransitive verbs be passives, as motivated by the evidence in (2-11) above. However, there is another class of intransitive verbs with corresponding adjectives, where the verbs are not passives. It is this additional data which will give rise to a more general formulation of the adjectival passive rule below, a formulation which will also aid in determining the class with which this rule is associated.

(12) LR-ADJ-PASS lexical rule (preliminary formulation)

LR-ADJ-PASS		(INCOMPLETE!)
LE2-Classes – ADJECTIVE	=	LE1-Classes – PASSIVE
LE2-Semantics	=	(STATE-OF LE1-Semantics)

Aside from the change of category, this rule mentions only semantics; all other idiosyncratic properties of both the verb and the adjective are identical. Any other differences between the two related lexical entries will be inherited, the result of their belonging to different classes.

This analysis shares with that of Bresnan 1983 the property that it relates adjectival passives to intransitives with the morphology of past participles, where the rule was motivated by passive intransitive verbs. As Bresnan (1978:8, 1982:29-32) shows, however, there is another class of adjectives related to intransitive verbs with past participle morphology, where the active verb is intransitive, as illustrated in (13 - 14) (examples taken from Bresnan 1982:30 and Levin and Rappaport 1986).

(13) a. fallen leaves
b. wilted lettuce
c. swollen feet
d. burst pipes

(14) a. *coughed patients
b. *cried children
c. *yawned students
d. *slept babies

Following Perlmutter 1978, who argued that there is syntactic motivation cross-linguistically for distinguishing these two types of intransitive verbs, Levin and Rappaport suggest that the contrasts in (13-14) can be explained by restricting the adjective-forming rule to *unaccusative* verbs in English, as in (13), excluding the *unergative* verbs in (14). As Perlmutter 1978 and others since then have shown for several languages, some intransitive verbs (unaccusatives) pattern with passives, while other intransitive verbs (unergatives) pattern with active transitive verbs. L. Levin 1987 provides a good summary of recent work on unaccusativity, and argues that the two groups of intransitives "should be represented as natural classes in a theory of grammar," a view that I adopt in the analysis given below.

While Levin and Rappaport do not provide independent evidence for such a syntactic distinction in English, some evidence does exist, as argued by Simpson 1983 on the basis of resultative constructions, and by Maling 1987 and L. Levin 1987 on the basis of constructions with pleonastic *there*. Neither of these constructions divides the set of intransitives exactly as does the adjectival construction illustrated in (13-14), but the overlap is significant; see L. Levin 1987 and several references cited there for discussion of mismatches in tests for unaccusativity. Levin concludes that virtually all constructions involving unaccusatives are in some way semantically restricted, so even though in some languages (arguably including English and Dutch) a syntactic account is necessary, it will not be sufficient to account for the distribution of unaccusatives. While acknowledging the need for additional semantic restrictions, I will only be concerned here with the syntactic aspects of unaccusatives in English, in pursuit of a more elegant analysis of adjectival passivesin English. Toward that end, I review one of the pieces of evidence for the unaccusative/unergative distinction, that given by Simpson 1983.

Simpson observes that resultative phrases like *flat* or *clean* can be predicated of objects of transitive verbs, and subjects of some intransitive verbs (including passives), but cannot be predicated of subjects of other intransitives (or of any

transitives). This distribution, typical of unaccusatives, is illustrated in (15), with several of the examples taken from Simpson 1983.

(15) a. He hammered the nail flat.
He wiped the board clean.

The nail was hammered flat.
The board was wiped clean.

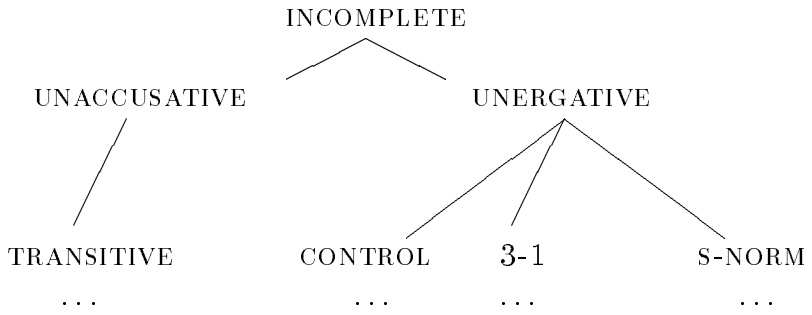
The overripe tomato fell flat.
Her feet swelled too large for her shoes.
The lettuce wilted completely limp.
The pipes burst wide open.

b. *He hammered the nails exhausted.
*She lit the candle visible.
*The old man coughed hoarse.
*The babies cried asleep.
*The audience laughed into the aisles.

This distinction in the acceptability of resultatives offers independent support for a division of English intransitives into two subclasses, one which includes unaccusative verbs like *fall* and *burst*, and the other for verbs like *cough* and *cry*. The unaccusative class permits an optional resultative which is not possible for the other class of intransitives, the unergatives. Since strictly transitive verbs also permit a resultative phrase, controlled by the object, they also belong to this unaccusative class, while the remaining classes of verbs inherit from the unergative class, since they lack resultatives.

On the basis of Simpson's evidence, the following restructuring of the INCOMPLETE part of the word class hierarchy of chapter 2 is motivated, showing the division of the INCOMPLETE class into two immediate subclasses, with TRANSITIVE inheriting from UNACCUSATIVE, and the other three old subclasses of INCOMPLETE now inheriting from the other new subclass UNERGATIVE.

(16) Subclasses of INCOMPLETE, restructured



In the class definition given for UNACCUSATIVE in (17), the only distinguishing subcategorization property specified is that members of this class permit an optional resultative phrase, whose syntactic and semantic properties I do not characterize fully. I note only that it must be predicative (to exclude attributive adjectives), and is subject to semantic control according to the same principles governing other complements and adjuncts, so that ordinarily the direct object will control if there is one, otherwise the subject controls.

(17) UNACCUSATIVE word class

UNACCUSATIVE	
Superclasses	INCOMPLETE
Complements	(Resultative)
Resultative-features	(PREDICATIVE +)
Lexical-rules	LR-ADJ-PASS

With this class in place, the rule in (12) can be generalized as in (12') to provide a more satisfactory account of the relationship between the relevant intransitive verbs and their adjectival counterparts.

(12') LR-ADJ-PASS lexical rule (final formulation)

LR-ADJ-PASS		(UNACCUSATIVE!)
LE2-Classes – ADJECTIVE	=	LE1-Classes – PAST-PARTICIPLE
LE2-Semantics	=	(STATE-OF LE1-Semantics)

This formulation expresses the two desired generalizations: (1) All intransitive passives should allow resultatives which are predicated of the passive's subject, since these passives are related to transitives which by hypothesis all belong to the UNACCUSATIVE class, and since all passive participles are members of the PAST-PARTICIPLE class, as presented in the previous chapter.[2]

(2) All transitive verbs should, as a default, have corresponding adjectives with the same morphological form as the past participle of the verbs, where each adjective assigns to its head noun the same thematic role assigned by the passive verb to its subject. Of course, as with any lexical rule there may be lexically idiosyncratic exceptions to this rule, either exceptions about morphological form, or exceptions to the applicability of the rule itself. Wasow 1980:307-9 identifies several candidate verbs which may be exceptions to the rule in (12'), but decisions about when to treat a verb as exceptional, and when to take such a verb as a sign that the rule needs to be more carefully constrained, are not ones to be made lightly, and would be a diversion here.

The generalization in (1), then, is expressed as a consequence of inherited information about the optional Resultative phrase, while that in (2) is the consequence of two interacting lexical rules, one the passive rule defined in section 3.1, and the other the rule given in (12').

This combination of subclass and lexical rule, then, extends to predict the following contrasts, on the assumption that those verbs with corresponding adjectives belong to the UNACCUSATIVE class, while those without them don't. The contrasts in (18-19) are noted in Levin and Rappaport 1986, and I give the others to

[2]The account given here would predict that resultatives should be possible for other than simple transitive verbs, such as ditransitives or object-equi verbs, since the active forms of these verbs also inherit from TRANSITIVE, and thus also from UNACCUSATIVE. Nothing presented here, then, predicts the ungrammaticality of the examples in (i); this suggests that a still more accurate structure for the INCOMPLETE part of the hierarchy would have a separate class for strictly transitive verbs, as a subclass of TRANSITIVE, and this subclass (rather than TRANSITIVE) would inherit from UNACCUSATIVE, then predicting the ungrammaticality of examples as in (i). I do not pursue this approach here, since it would require seeking independent support for such a class of strictly transitive verbs, taking us too far afield.

(i) *John set the ice cream in the oven melted.
*John required Bill to sing hoarse.

suggest the difficulty in drawing the relevant distinctions, beyond simply stipulating membership in either the unaccusative or unergative class.

(18)	a. vanished	John studies vanished civilizations.
	b. disappeared	*John looks for disappeared people.
	c. appeared	*John studies appeared holes in the ozone layer.
	d. reappeared	*John processed the reappeared children.
(19)	a. rested	John prefers (well-)rested children.
	b. slept	*John prefers (well-)slept children.
(20)	a. surfaced	John addressed the recently surfaced objections.
	b. arisen	*John addressed the recently arisen objections.
	c. risen	*John greeted the risen children.
		(cf. Bresnan 1982's *the risen Christ*)
(21)	a. returned	John welcomed the returned monarch.
	b. remained	*John visited the remained monarch.

Like other lexical rules, this one for relating unaccusative verbs to adjectives has exceptions, at least with respect to morphology. The following examples show distinct forms for pairs of entries related by this rule, where the past participle of an arguably unaccusative verb is not the same as the corresponding adjective, but is clearly morphologically related, at least historically. The two pairs in (22) come from Lieber 1981:199; I add those in (23).

(22)	a.	burned	?John discarded the burned toast.
			John admitted he had burned the toast.
		burnt	John discarded the burnt toast.
			*John admitted he had burnt the toast.
	b.	proved	*John only teaches proved theorems.
			John has proved another theorem.
		proven	John only teaches proven theorems.
			?John has proven another theorem.

(23) a. shrunk *John admired the shrunk pumpkin.
The pumpkin has shrunk considerably.
shrunken John admired the shrunken pumpkin.
*The pumpkin has shrunken considerably.
(cf. The pumpkin is somewhat shrunken.)
b. died *John buried the died bird.
The bird had died.
dead John buried the dead bird.
*The bird has dead.
(cf. The bird is dead.)

A few verbs like *open* present an interesting case, with two morphologically related adjective forms, one with the form of the past participle, the other not, as shown in (24).

(24) a. opened John saw the opened window.
The window has opened slowly.
The window was opened slowly by the burglar.
b. open John saw the open window.
*The window has open slowly.
(cf. The window is open.)

The simplest analysis of these forms is that the adjective *opened* is related by the adjectival passive rule in (12') to the passive of the transitive verb *open*, while the adjective *open*, with an irregular spelling, is related by (12') to the intransitive verb *open* (which is itself related to the transitive *open* by the causative lexical rule introduced in section 3.2).[3] Other transitive verbs affected by the causative rule should also have two related adjectival forms (if the intransitive verbs related by the causative rule are also unaccusatives), though most of these pairs of adjectives are morphologically identical (e.g., closed, shut), as expected given identity as the default relation in lexical rules.[4]

[3]Note the curious asymmetry of these two adjectives with respect to the negative prefix *un-*, seen in (i).

(i) a. The unopened window is cracked.
b. *The unopen window is cracked.

It is not clear why *unopen* is impossible, without a better understanding of the constraints on complements of *un-*.

[4]For curiosity's sake, note in (i) an example of a three-way morphological distinction among the past participle, the corresponding predicative adjective, and the related but distinct attributive

This analysis of adjectival passives provides a further illustration of the way in which word classes and lexical rules can interact, in this case leading to a more fine-grained distinction among intransitives, which resulted in a restructuring of the word class hierarchy. The exercise again raises questions about what constraints, if any, affect the interaction of lexical rules, and what classification of lexical rules is possible. To gain a broader base from which to develop answers to these questions, I work through a second extended example in the next section.

---

adjective (unusual since most adjectives of this kind can be used either attributively or predicatively).

(i) a. lived: These flowers have never lived long.
alive: These flowers are alive.
live: John prefers live flowers.

b. lived: *These flowers are lived.
*John prefers lived flowers.
(but cf. These flowers are long-lived.)
alive: *These flowers have never alive long.
*John prefers alive flowers.
live: *These flowers have live long.
*These flowers are live.
(but cf. This microphone is live.)

6.2 "Tough-Movement" adjectives

A second extended illustration of the kind of explanation available within this framework concerns the analysis of a class of adjectives that includes *easy, tough, difficult, amusing*, with some of the relevant data presented in (1-2).

(1) a. It is amusing to visit Disneyland.
b. It is amusing for kids to visit Disneyland. (two readings)
c. It is amusing for parents to visit Disneyland. (two readings)
d. It is amusing for parents for their kids to visit Disneyland.
e. *It is amusing for hotels for parents for their kids to visit Disneyland.
(ignoring reading where *hotels for parents* is a constituent)

(2) a. Disneyland is amusing to visit.
b. Disneyland is amusing for kids to visit. (one reading)
c. Disneyland is amusing for parents to visit. (one reading)
d. *Disneyland is amusing for parents for their kids to visit.
e. *Disneyland is amusing to visit Yosemite.

This construction, which has been of interest for a long time, first received discussion within the generative framework in Chomsky 1964:61-65, and has been a favorite cliff for linguistic rock climbers ever since.[5] While the earlier, transformational accounts attempted to capture the relationship between the constructions in (1) and those in (2), the more recent, lexically-based accounts have ignored the constructions in (1), concentrating instead on giving a base-generated account of the sentences in (2), which contain an NP gap. I follow these lexicalist pioneers in accounting for the examples in (2), but provide as well an account of their relationship to those in (1). I then show how this analysis extends to cover a broader range of constructions involving not only nominal variants of the *easy* adjectives, but also the *too/enough* constructions studied by Lasnik and Fiengo 1974.

[5]Explorers include Postal 1970, Bresnan 1971, Berman 1973, Lasnik and Fiengo 1974, Chomsky 1977, Fodor 1978, Nanni 1978 and 1980, Schachter 1981, Jacobson 1982:221-3, Maling and Zaenen 1982:253-4, Kaplan and Bresnan 1982:255-63, and GKPS 1985:150-152.

6.2.1 Two kinds of *amusing*

To begin with, adjectives like *amusing* as in (1) are predicates which have an infinitival complement, an optional prepositional phrase introduced by *for*, and an expletive *it* subject. The infinitival complement can be either a VP or an S introduced by the complementizer *for*; when it is a VP, the complement seems to allow arbitrary control, where one of the possible controllers is the For-PP, when present. Except for the optional For-PP, the properties of *amusing* are very much like those of *possible*, which was introduced in the previous chapter as a member of the S-INF-IT class in the COMPLEMENTATION hierarchy. To capture that similarity, I introduce a subclass of S-INF-IT for lexical entries like *amusing*, named EASY-IT in honor of its most infamous member, with the -IT suffix intended to help distinguish this class of predicates from a related class I introduce below.

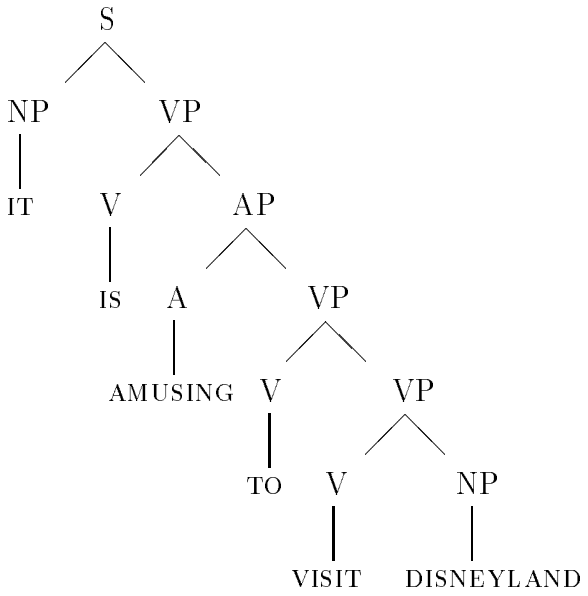
(3) EASY-IT word class

EASY-IT	
Superclasses	S-INF-IT
Complements	(For-PP)
For-PP-features	(CAT Preposition) (LEXICAL –) (PFORM For)
Lexical-rules	LR-EASY

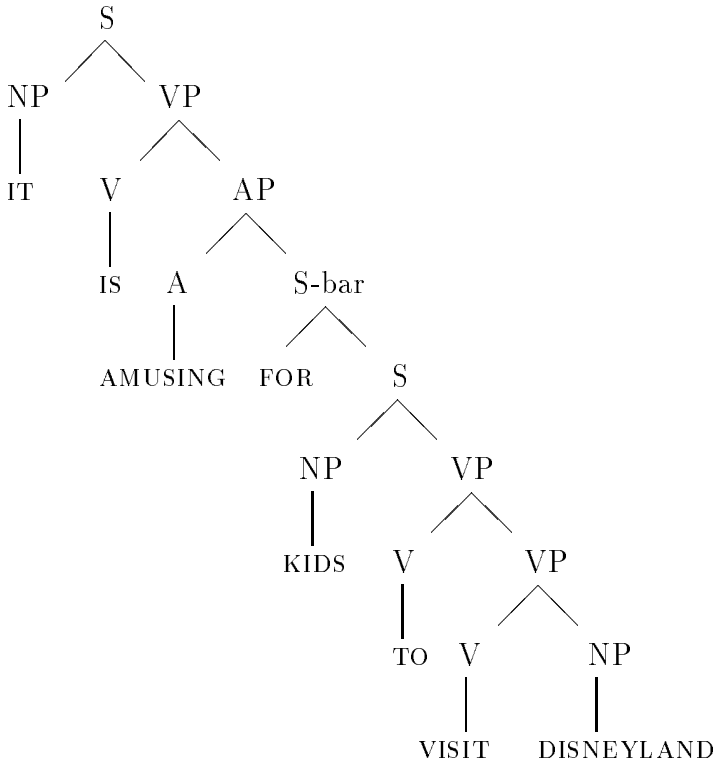
According to the definition in (3), members of this class will have, in addition to the inherited expletive *it* subject and infinitival (VP or S) complement, an optional For-PP. In addition, this class mentions the lexical rule that will relate members of this class to those of another class, one which will account for the examples in (2) above. I define this other class and the lexical rule below.

Adjectives that belong to the EASY-IT class will provide the right subcategorization information to permit sentences like those in (1a-d) above, but not (1e). The trees corresponding to (1a-d) are given in (4a-d), with the nodes labelled informally, for ease of reading.

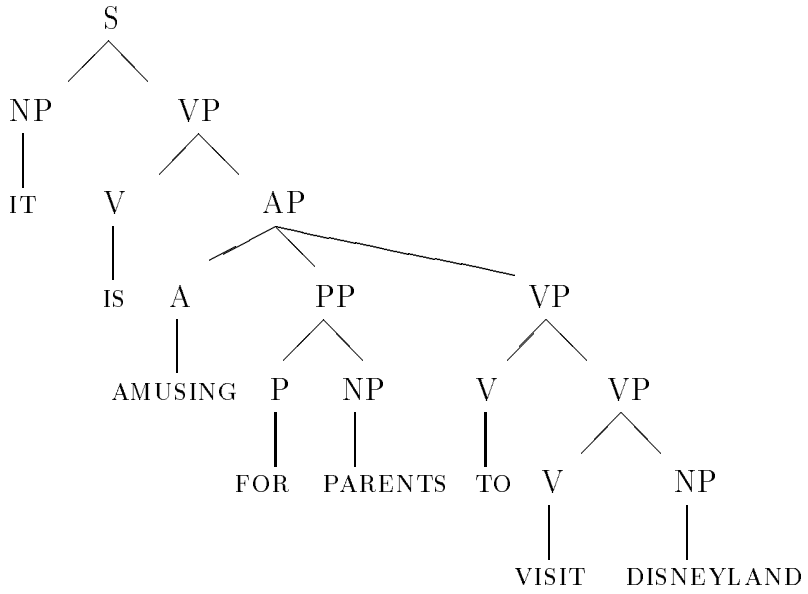
(4) a. It is amusing to visit Disneyland.



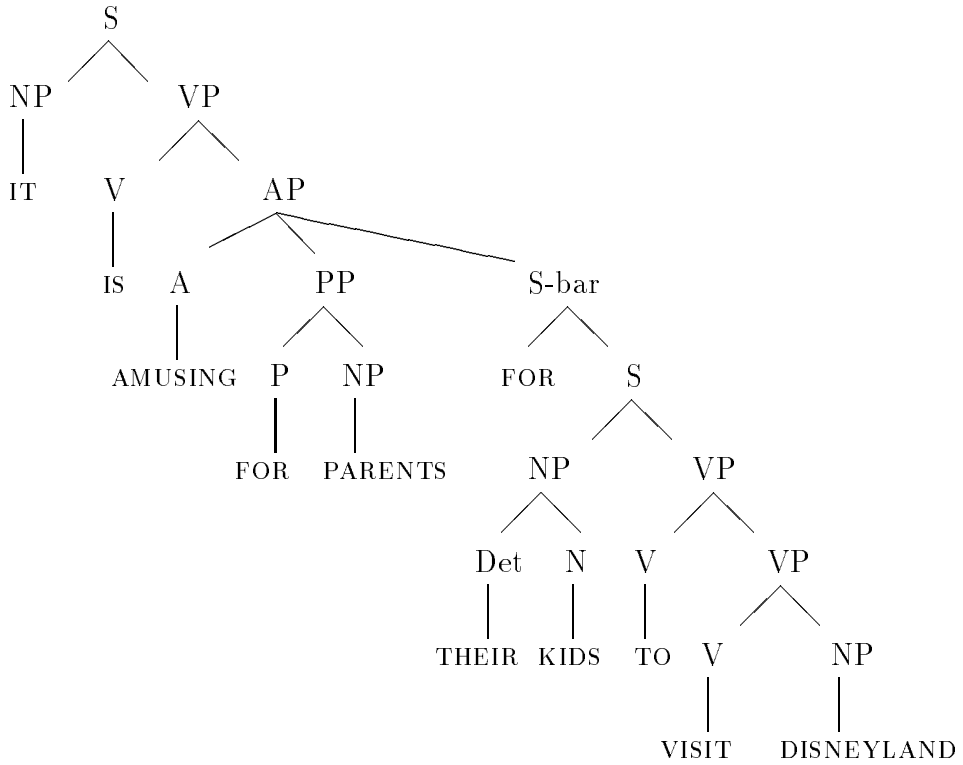
b. It is amusing for kids to visit Disneyland



c. It is amusing for parents to visit Disneyland.



d. It is amusing for parents for their kids to visit Disneyland.



Turning now to the examples in (2) above, I will assume that the adjective *amusing* used here has a lexical entry distinct from that in the examples of (1), an entry belonging to another class in the COMPLEMENTATION hierarchy, and related by lexical rule to the corresponding member of the EASY-IT class. However, this new subclass also contains adjectives which do not have corresponding entries in the EASY-IT class; I introduce them first, to motivate the proper definition of the class, together with its subclasses.

In their paper on this class of adjectives, Lasnik and Fiengo (1974:535) identified a set of adjectives like *pretty* and *delicious*, which they labelled "pure Object Deletion" adjectives, illustrated in (5).

(5) a. Mary is pretty to look at.
c. Sonatas are melodious to listen to.

Adjectives like these allow complements similar to adjectives like *amusing* in (2) above, but the *pretty* adjectives are different in two important respects. First, they do not have corresponding entries in the EASY-IT class, as seen in (6); and second, the *pretty* adjectives assign an independent thematic role to their subjects, as seen by the differences in entailment illustrated in (7).

(6) a. *It is pretty to look at Mary.
b. *It is melodious to listen to sonatas.

(7) a. Mary is pretty to look at.
entails
Mary is pretty.

b. Mary is difficult to recognize.
does not entail
Mary is difficult.

These two differences, which I do not try to correlate, indicate the need to separate into distinct classes the *pretty* adjectives and the *amusing* ones. What they have in common is that they both have three complements: a non-expletive subject, an infinitival VP with an NP gap, and an optional prepositional phrase introduced by *for* which, when present, controls the VP. These common properties indicate that

the two classes should be subclasses of a single class, which I label VP-SLASH, and take to be a subclass of EQUI, given the presence of a controlled complement for predicates like *pretty* and *amusing*.

(8) VP-SLASH word class

VP-SLASH	
Superclasses	EQUI
Complements	(For-PP)
For-PP-features	(CAT Preposition) (LEXICAL –) (PFORM For)
XComp-features	(VFORM Infinitive)
	(SLASH ((CAT Noun) (COMPLETE +)
	(CASE Accusative) (PREDICATIVE –)))

Members of this category still have an inherited subject, but keep the default (NFORM Norm) specification of equi predicates, so no mention of it is made here. The For-PP is marked to be optional, like that of the EASY-IT class. Since the default syntax for the XComp of control predicates is a VP, this class simply constrains it to be infinitival, with the additional specification that it must contain an accusative NP gap.[6] Since it is a subclass of OBJECT-EQUI, members of this class will treat the direct object as the default controller, though some additional provision needs to be made for getting arbitrary control when the For-PP is not present. Finally, members of this class will, as a default, assign a thematic role to their subjects, like other object-control predicates.

The (PREDICATIVE +) specification on the slashed NP is there to block sentences like those in (9a-b), which have corresponding grammatical sentences in (c-d).

(9) a. *A doctor is easy to be.
b. *A doctor is easy to become.
c. It is easy to be a doctor.
d. It is easy to become a doctor.

[6] Cf. Schachter's (1981:441) proposal, alike in spirit, for describing this kind of adjective, minus the case restriction, and Jacobson's (1982:222) similar proposal which does include a case restriction; GKPS (1985:150) essentially adopt Jacobson's approach, but specify Accusative case rather than non-Nominative as Jacobson does.

The ungrammaticality of examples (9a-b) results from the fact that both *be* and *become* subcategorize for a predicative phrase, so when there is a trace of that complement, the trace must carry the (PREDICATIVE +) property, which will be passed up in the value of SLASH on the intervening VP node, and eventually fail to match the (PREDICATIVE –) stipulation on the VP/NP subcat for *easy*.

While I will not provide a precise semantics for VP-SLASH predicates, it seems clear that the lexical entries for these predicates must ensure that the subject be co-indexed (or unified) with the slash-variable that corresponds to the NP-gap introduced in the VP complement. This declaration in the lexical entry, together with the default direct object control of the slashed VP, will allow the construction of the desired semantics for a phrase headed by an adjective like *pretty*.

As an example of a member of this class, the non-redundant entry for *pretty* will be as given in (10), specifying only the classes it belongs to, and the idiosyncracies of spelling, phonology, and semantics.

(10) Minimal lexical entry for *pretty*

PRETTY	
Superclasses	ADJECTIVE, VP-SLASH
Spelling	"pretty"
Phonology	/prIti/
Semantics	(PRETTY bearer:X exper:Y prop:Z)
Complements	
Subject-index	X
PP-For-index	Y
XComp-index	Z

Adjectives like *easy* or *amusing* as in (2) above have the same syntactic properties as do adjectives like *pretty*, so will also inherit from the VP-SLASH class. However, a separate class is needed, for the reasons given above, one involving a semantic difference, and the other involving the lexical rule that will finally account for the relationship between the two kinds of *easy* adjectives. Semantically, these *easy* adjectives are different from ordinary object-equi predicates in not assigning an independent role to their subjects. In this respect they are like raising predicates, though the role that gets assigned to the subject of *easy* is not the role of a controlled

complement's subject, but only the role from the slashed NP, as sketched above. To indicate this difference between *pretty* and *easy*, I use the same device used in the definition of the classes for raising predicates, explicitly blocking inheritance of a subject-index attribute in need of a value. Secondly, the definition of this subclass of VP-SLASH, which I label EASY-SLASH, identifies the rule which establishes the relationship between the two kinds of *easy*.

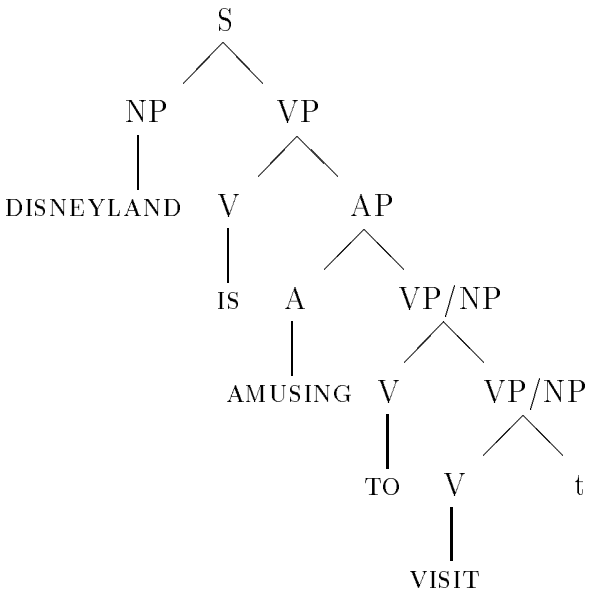
(11) EASY-SLASH word class

EASY-SLASH	
Superclasses	VP-SLASH
Complements	
Subject-index	DO-NOT-INHERIT
Lexical-rules	LR-EASY

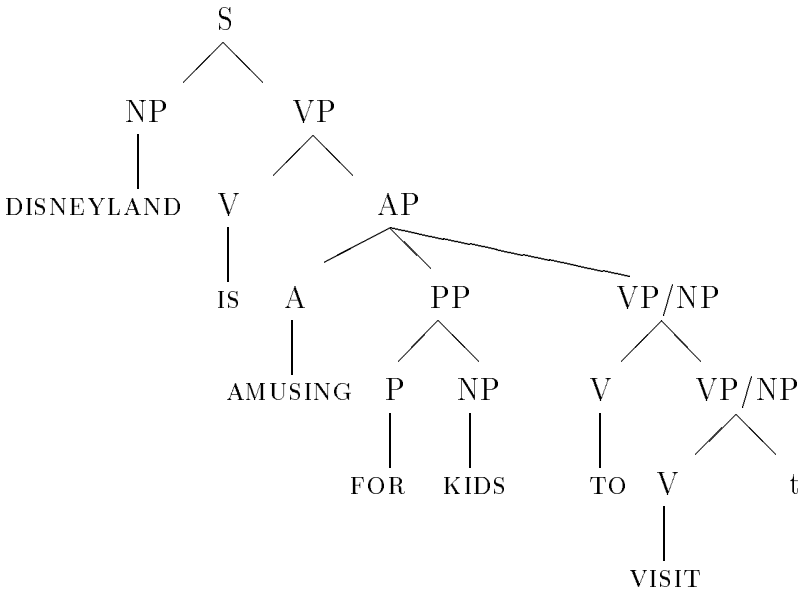
With the syntactic properties defined in VP-SLASH, from which it inherits, an adjective like *amusing* will appear in sentences like those in (2a-c) above, but not in (2d-e). The example in (2d) is ruled out because there is no provision for an S-bar with a *for* complementizer. Example (2e) is ruled out because the VP complement to *amusing* fails to have an NP gap, and hence does not satisfy the constraint specified in the definition in (8) above.

Trees for the examples in (2a-c) are given in (12a-c).

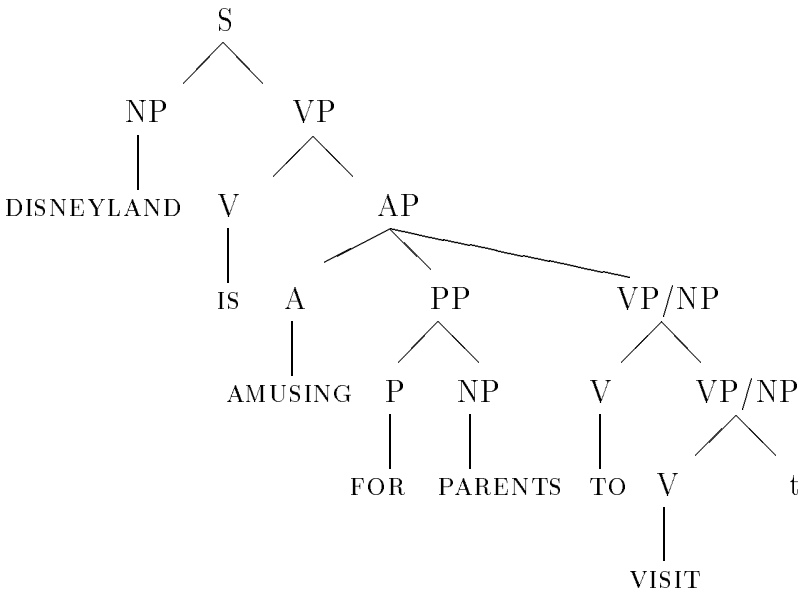
(12) a. Disneyland is amusing to visit.



b. Disneyland is amusing for kids to visit.



c. Disneyland is amusing for parents to visit.



Given the classes defined in (3), (8), and (11), the lexical rule that expresses the relation between pairs like the *amusing* of (1) and the *amusing* of (2) is quite simple, with all the work already done by inheritance. I give the definition of the rule as follows.

LR-EASY
LE2-Classes – EASY-SLASH = LE1-Classes – EASY-IT

This rule, introduced on the two classes mentioned in the rule, holds for each member of either class, unless the particular lexical item explicitly blocks applicability of the rule. Given the way in which the class hierarchy has been presented, there do not appear to be any examples of exceptions to this rule. Adjectives like *pretty* do not have to be marked as exceptions to the rule, since they are not members of one of the classes for which the rule is defined. Likewise, adjectives like *possible*, which might have been thought to be members of the EASY-IT class, have some properties distinct from adjectives like *amusing* in (1), making it clear that *possible* and others of its ilk belong to a superclass of EASY-IT, thereby remaining outside the scope of this lexical rule.

The principal difference between *amusing* and *possible*, aside from susceptibility to the LR-EASY rule, is that *easy* allows an optional For-PP complement in addition to the infinitival complement, while *possible* does not, as illustrated in (13).

(13) a. It is amusing for parents for their kids to visit Disneyland.
b. *It is possible for parents for their kids to visit Disneyland.

This difference follows from the definitions given earlier for the two classes S-INF-IT and EASY-IT, where *possible* is directly a member of the former, and *amusing* is a member of EASY-IT. Given the distinction between the two classes, the examples in (14) show that the LR-EASY lexical rule belongs as part of the definition of the EASY-IT class, not its parent class S-INF-IT.

(14) a. It is possible to visit Venice.
b. *Venice is possible to visit.
c. It is silly to hire Abernathy.
d. *Abernathy is silly to hire.

6.2.2 Nouns of the *easy* class

Lasnik and Fiengo 1974:568 also observe that there are some noun phrases which have properties like those of the *easy* adjectives, as in (15).

(15) a. Nureyev is a pleasure to watch.
b. This course is a breeze to pass.
c. Venice is a delight to visit.
d. It is a pleasure to watch Nureyev.
e. It is a breeze to pass this course.
f. It is a delight to visit Venice.

Since the class definitions given in (2) and (5) above made no mention of major category, these nouns are quite simple to represent: *pleasure* and *pleasant* both have entries that are members of the EASY-IT class defined in (2), and both also have entries which are members of the EASY-SLASH class, as predicted by the lexical rule given in (8). The differences between *pleasure* and *pleasant* follow from the fact

that the former is also a member of the COMMON-NOUN class, while the latter is instead also a member of the ADJECTIVE class.[7]

Just as there are four kinds of adjectives relevant to the discussion of the LR-EASY lexical rule, so are there four kinds of common nouns. The four kinds of adjectives, defined above in terms of the four classes EASY-IT, EASY-SLASH, S-INF-IT, and VP-SLASH, are exemplified by the two varieties of *easy*, *possible*, and *pretty*. If the types of nouns are to be parallel, there should be nouns that exhibit properties similar to *possible*, and others analogous to *pretty*.

Lasnik and Fiengo 1974:536 provide examples of nouns analogous to *pretty* and *melodious*, with one entry specifying a VP/NP complement, but no related entry with an unslashed infinitival complement, illustrated in (16-17).

(16) a. This room is a pigsty to behold.
b. Nureyev is a marvel to watch.

(17) a. *It is a pigsty to behold this room.
b. *It is a marvel to watch Nureyev.

Consistent with the approach taken for adjectives like *pretty*, these nouns will be members of the VP-SLASH class.

The fourth relevant set of nouns will be those which are analogous to adjectives like *possible*, which have an expletive *it* subject and an infinitival complement (S or VP) without a gap. Examples include nouns like *battle* and *relief*, seen in (18-19).

[7]These nouns will actually have to belong to a subclass of COMMON-NOUN which restricts the final obligatory argument to being the determiners *a/an*, in order to prevent examples like those in (i).

(i) a. *Venice is my delight to visit.
b. *Venice is the delight to visit.

There is at least one example of a mass noun that appears without a determiner: the word *fun*, illustrated in (ii), which seems to be taking on characteristics of an adjective for some (younger?) speakers, who find the examples in (iii) quite acceptable.

(ii) a. It is fun to visit Disneyland.
b. Disneyland is fun to visit.

(iii) a. ?It is very fun to visit Disneyland.
b. ?Disneyland is a fun place to visit.

(18) a. It was a major battle to replace John.
b. It was a real relief to take off my cast.

(19) a. *John was a major battle to replace.
b. *My cast was a real relief to take off.

These examples reveal a close parallelism between four groups of adjectives and four groups of nouns, with the relationships among the groups defined in terms of the four relevant classes, with the lexical rule LR-EASY.

Notice that on this account, each of the two entries for *pleasure* is not related to the corresponding entry for *pleasant* by some lexical rule; instead, the corresponding entries for *pleasure* and *pleasant* have similar properties because they have a word class in common (either the EASY-IT class or the EASH-SLASH class). What the lexical rule does is relate the one entry for *pleasure* to the other entry for *pleasure*, and likewise for the two entries for *pleasant*.

6.2.3 On *hard problems to solve*

One variation of the *easy to please* constructions that has received little analysis is illustrated in (20), where the adjective and its VP/NP complement are not contiguous.

(20) a. John is an easy man to please.
b. John is a man easy to please.

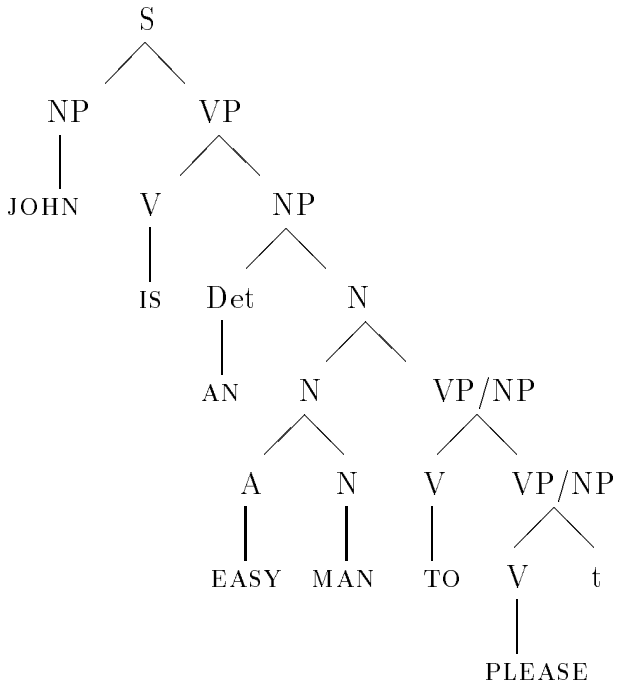
(21) a. *John is an easy to please man.
b. *John is an easy man.

(22) a. *John is an easy man to please Bill.
b. *John is a man easy to please Bill.

While the two examples in (20) are good, employing the *easy* which belongs to the EASY-SLASH class, the examples in (21-22) are ungrammatical. Given what I have said so far, the correct judgments are predicted for (20b), (21a), and (22a), but something more must be said in order to explain the grammaticality of (20a) and the ungrammaticality of (21b, 22b).

I will focus on explaining the grammaticality of (20a), assuming that the right syntactic structure for the sentence is the binary-branching structure given in (23), where *easy* forms a constituent with *man*, and where *to please* is sister to the phrase *easy man*. I have adopted the binary structure largely because it will simplify the exposition here; it might be equally defensible to hold that *easy*, *man*, and *to please* are all sisters of a single phrase.

(23) John is an easy man to please



What is awkward about this structure is that the head noun *man* does not by itself subcategorize for the VP/NP, with the intended reading.[8]

[8]There is a suspiciously similar construction, illustrated in (i), which might be expected to shed some light on the proper analysis of (20a), but which has a restricted enough interpretation to suggest that it should be treated separately, probably derived from the more general construction exhibited in (20a).

(i) a. John is a man to admire.
b. Mary is a woman to emulate.
c. This is a word to keep on the tip of your tongue.

These examples seem to mean something like *John is a good man to admire* or *Mary is a good*

Rather, it seems that when *easy* combines with *man*, the resulting phrase has a Complements list which contains not only the optional and obligatory complements that *man* started out with, but also the obligatory VP/NP complement and the optional For-PP controller required by *easy*. No mechanism presented so far provides for an adjunct combining with its head to affect the Complements list of that head or of the resulting phrase. Yet if the phrase structure proposed in (23) is correct, some kind of merging of subcat information between adjunct and head must be provided for.

The examples in (24-25) illustrate that the flow of information from an adjunct's list of subcats to the head's must be quite restricted; it would not do to simply merge the Complements list of any adjunct with that of the head in every case.

(24) a. *an eager man to please
b. *a fearful man of snakes
c. *a frightened man by snakes
d. *an angry man at John

(25) a. a man eager to please
b. a man fearful of snakes
c. a man frightened by snakes
d. a man angry at John

The above examples might suggest that what distinguishes *easy* from these other adjectives is that the VP/NP complement of *easy* is obligatory, while the PP complements of the above adjectives are optional. While there are not many adjectives

woman to emulate, where the semantic contribution of *good* has been incorporated into the N – VP/NP construction in (i). To test this, consider the examples in (ii), where the *good* reading should lead to an anomalous interpretation, and does (cf. the corresponding examples in (iii).

(ii) a. ?Mary is a person to underestimate.
b. ?Sharks are animals to tame.

(iii) a. Mary is an easy person to underestimate.
b. Sharks are difficult animals to tame.

Given the constrained interpretations of examples like those in (i-ii), it does not seem defensible to treat *easy man to please* as simply the modifier *easy* combining with *man to please*. In addition, such an analysis would leave unexplained the ungrammaticality of **John is an easy man*.

against which to test this hypothesis, the one clear case of an adjective that takes an obligatory complement counts against the idea:

(26) a. a man fond of snakes
b. *a fond man
d. *a fond man of snakes

Transferable subcats

The analysis I propose localizes in lexical entries the ability of a subcat to be transferred from adjunct to head. Just as subcats can be marked for the obligatory/optional distinction in a class definition or in a lexical entry, so can they be marked for a distinction I will term *transferable*. While as a default subcats will be Non-Transferable, those subcats which are identified by a class or lexical entry as Transferable will be subject to the following informally stated convention:

(27) **Transferable Subcat Convention**

> When a *transferable* subcat on a daughter in a local subtree is not associated with some sister in that subtree, the subcat becomes part of the corresponding subcat list of the head daughter in that subtree.

In the constructions studied here, this convention applies in cases where the lexical entry or phrase with a transferable subcat serves as an adjunct, so that the word or phrase's subcat list is not used directly. The intent of the convention in such cases is to make the transferable subcat a part of the head, so the head feature convention will ensure that the information is propagated to the mother node.

Having introduced this additional property of subcats, that they can be specified as *transferable*, I note that the default value for this property must be negative, since in general subcats from adjuncts do not pass to heads, as seen in (24) and (26) above. This default value will be overridden for the VP/NP and the For-PP subcats in the VP-SLASH class, to reflect the grammaticality of both examples in (28).

(28) a. That was a melodious sonata to listen to.
b. John is an easy man to please.

The informal notation I will use to indicate that a subcat is *transferable* is illustrated in the more complete definition of the VP-SLASH class that I give in (29). A subcat in the Complements list which is bracketed with vertical bars is one that has the non-default, positive value for the attribute Transferable, and is therefore subject to the convention given in (27) above.

(29) Revised definition of VP-SLASH

VP-SLASH	
Superclasses	EQUI
Complements	|(For-PP)| |XComp|
For-PP-features	(CAT Preposition) (LEXICAL –) (PFORM For)
XComp-features	(VFORM Infinitive)
	(SLASH ((CAT Noun) (COMPLETE +) (CASE Accusative) (PREDICATIVE –)))

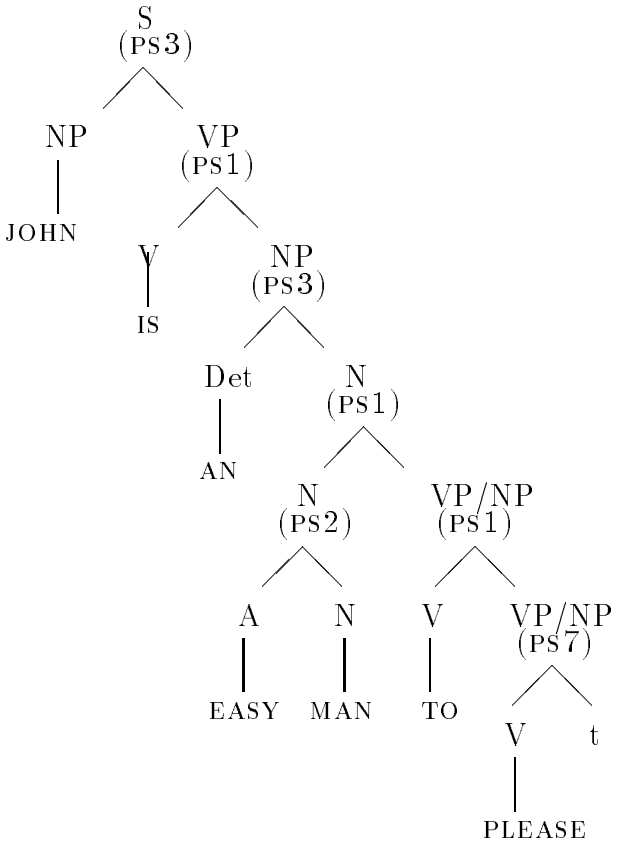
Members of this class, including the relevant lexical entry for *easy*, will inherit this non-default *transferable* property for both the XComp and the For-PP, so when *easy* combines as an adjunct with the head noun *man* (via the Adjunct PS rule), these two subcats will become part of the Complements list of *man*, by the convention in (27), and will then become part of the Complements list of the node *easy man*, accounting for the grammaticality of (20a) above.

For the feature-conscious reader, I note here that it is crucial that grammar rule (PS2) for adjuncts not have the left-hand side be stamped (LEXICAL –). Intuitively, if the head is lexical, and the only sister picked up is an adjunct, then the resulting constituent still retains everything on its Complements list (plus more, in some cases), so should properly still be labelled (LEXICAL +), given the definitions for the feature LEXICAL provided in the previous chapter. This is crucial because after the noun *man* has combined with the prenominal adjunct *easy*, the resulting phrase must be a legal candidate for the head of the post-head complements rule (PS1), which is restricted to heads that are (LEXICAL +). (The position is consistent with the definitions, even if it violates another intuition about what (LEXICAL +) might mean: that the constituent can be found in the lexicon. The feature wasn't defined

that way, and such a definition would not be straightforward to make precise in any case.)

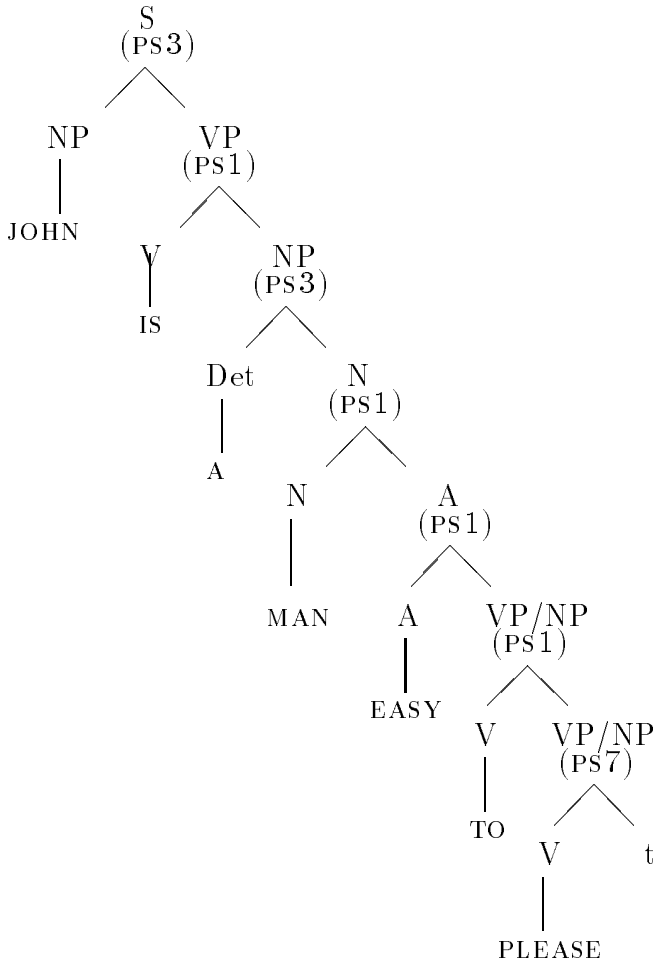
I repeat in (30) the tree given in (23), but here with each phrasal node annotated with the number of the relevant PS rule.

(30) John is an easy man to please.



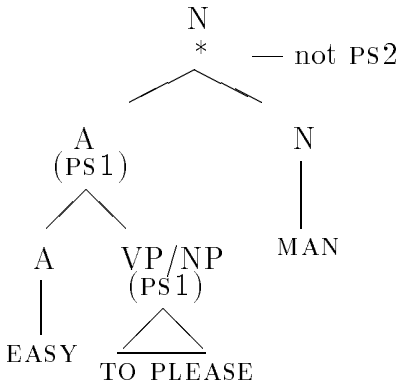
(20b) will have the following structure, with the nodes again annotated:

(31) John is a man easy to please.



(21a) will be ruled out because the following subtree will not be licensed, since the node dominating *easy to please* will not be lexical, and therefore will not be able to appear as the pre-head adjunct using rule PS2.

(32) *easy to please man



(21b) is excluded because *easy* has an obligatory VP/NP complement, which must be included as an obligatory complement of the phrase *easy man*, due to the convention adopted above about merging of subcat information between a head and its sister. Given this obligatory XComp for *easy man*, the phrase cannot serve as head in grammar rule (PS3), since that rule requires a head whose Complements list has exactly one obligatory complement remaining.

(22a) is excluded because the *easy* which requires an unslashed VP complement will not pass on its XComp subcat to the noun it modifies, since that XComp does not contain any binding features that license the incorporation. So *easy man to please Bill* will be excluded for the same reason that *eager man to please* is excluded: nothing licenses the post-nominal infinitival VP.

(22b) is probably best excluded on semantic grounds, since the subject of *easy to please Bill* is an expletive pronoun, the wrong sort to unify with the head noun being modified. The mechanism that blocks this example will also serve to explain the lack of ambiguity in a noun phrase like *an easy problem*, an ambiguity that might be expected given that there are two distinct lexical entries for *easy*, both adjectives, and therefore apparently both eligible as adjuncts for the common noun *problem*. On the assumption that a noun must serve semantically as the subject of adjectival adjuncts, those adjuncts must specify some thematic role for the noun to play. Thus any adjective which requires an expletive subject should give rise to a semantically ill-formed expression when it appears as an adjunct to a noun. Thus,

what prevents the EASY-IT *easy* from serving as an adjunct to *problem* is the fact that this *easy* requires an expletive *it* subject,

Summary

What this analysis reveals is the utility of the word class hierarchy with its inheritance of properties as a mechanism for expressing complex relationships among lexical items, allowing with equal facility the representation of regularities, subregularities, and lexical idiosyncracy. As has been illustrated before, regularities can be expressed with either of two mechanisms, one being the inheritance of properties defined for word classes, and the other being lexical rules. In the analysis given here, the lion's share of the work is left to the inheritance mechanism, so that once the word classes were properly defined for the two kinds of *easy* adjectives, the lexical rule linking the two forms was trivial to state, The same word classes and the same rule also allow a straightforward account of the quite similar relationship among some pairs of common nouns like *pleasure*, taking advantage of the fact that a given lexical item can inherit from several word classes, drawing some properties from each. On this account, *pleasant* and *pleasure* show similar properties because they inherit from some of the same word classes.

The word classes defined here also provide the basis for an account of constructions like *easy man to please*, though this account requires the postulation of a convention for passing certain subcat information from adjunct to head. In the next section I draw additional support for these classes, the LR-EASY lexical rule, and this Transferable Subcat Convention, by giving a closely related analysis for constructions employing the words *too* and *enough*.

6.3 "Too" and "enough"

Jackendoff 1972:227 noticed that in addition to the familiar adjectives and nouns discussed above, there are two other words in English which also appear in phrases that contain a VP with an NP gap. Lasnik and Fiengo 1974:536 take up these constructions in much more detail, and I repeat some of their examples in (1-2) showing this property of the lexical items *too* and *enough*.

(1) a. The mattress is thin.
b. *The mattress is thin to sleep on.
c. The mattress is too thin to sleep on.

(2) a. The football is soft.
b. *The football is soft to kick.
c. The football is soft enough to kick.

While adjectives like *thin* and *soft* do not take VP/NP complements, these complements do appear in adjective phrases which contain not only an adjective like *soft* but also either *too* or *enough*. Informally, it seems that the combinations *too soft* and *soft enough* have properties that mirror those of members of the EASY-SLASH class described above. To see this, compare the examples in (3) with those in (4).

(3) a. Timbuctoo was hard to get to.
b. Timbuctoo was hard for Stanley to get to.
c. *Timbuctoo was hard to get to Cairo.

(4) a. Timbuctoo was too isolated to get to.
b. Timbuctoo was too isolated for Stanley to get to.
c. *Timbuctoo was too isolated to get to Cairo.

In principle, one could use either of two mechanisms to combine *too* with *isolated*: a lexical rule which augments the Complements list of *isolated* and adds the prefix *too*, or a phrase structure rule that concatenates *too* and *isolated*, providing some mechanism for merging the subcat information from *too* with that of *isolated*. I will argue here that the lexical rule approach, while initially more attractive, fails to make the right predictions about slightly more complex data, and must be abandoned in favor of a phrasal analysis for these constructions.

6.3.1 Not lexical rules

Any analysis of the adjectival modifiers *too* and *enough* should be able to account for at least the data in (5-6), perhaps also shedding light on the proper treatment of constructions like those in (7-10), which employ other adjectival modifiers such as *more/-er*, *so*, and *as*. The presence of each of these modifiers in a phrase is correlated with the presence of another modifying phrase that follows the head adjective.

(5) a. The Constitution is too long to memorize.
b. The Constitution is too long for kids to memorize.
c. The Constitution is too long to memorize it.
d. The Constitution is too long for kids to memorize it.

(6) a. The C. is long enough to assign as extra credit.
b. The C. is long enough for me to assign as extra credit.
c. The C. is long enough to assign it as extra credit.
d. The C. is long enough for me to assign it as extra credit.

(7) a. The C. is longer than the Gettysburg Address.
b. The C. is many pages longer than the Gettysburg Address.
c. The C. is longer than the Gettysburg Address is.
d. The C. is many pages longer than the Gettysburg Address is.

(8) a. The C. is the longest document of any in this file.
b. The C. is the longest document I ever memorized.

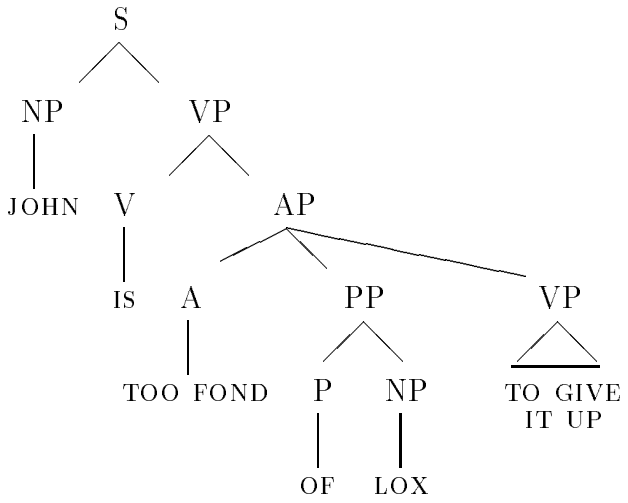
(9) a. The C. is so long that I could not memorize it.

(10) a. The C. is as long as ten Gettysburg Addresses.
b. The C. is as long as the Gettysburg Address is.

In each of these, the adjective *long* combines with another form, and in doing so, it picks up one or more additional complements, optional and obligatory, with the properties of those complements determined by the form that *long* combined with. The challenge here is to provide a way of merging the subcat information from the adjective with the subcat information correlated with *too*, *enough*, *-er*, and so on.

On the lexical rule approach, *too long* would be an adjective related to *long*, but with additional properties.[9] As a lexical item, *too long* with its augmented list of optional and obligatory complements would then serve as the head in the Post-head Complements rule, ready to pick up all complements (except the subject) in the expected fashion. This would give rise to phrase structures like that in (11), where *too fond* would be a single lexical item, an adjective subcategorizing for the obligatory Of-PP complement that *fond* requires, and also subcategorizing for the infinitival VP complement that is associated with *too*.

(11) John is too fond of lox to give it up.



While the use of lexical rules to combine *too* and *fond* might be expected to handle the examples in (5-6), there is evidence that shows the lexical rule approach is not general enough to handle even slightly more complex constructions. Using a lexical rule to combine *too* with *fond* predicts that the resulting lexical item will always be the head of its phrase, in order to make use of the now-augmented list of complements that the lexical rule created. But this is not the case; consider the examples in (12-13).

[9] Again, there would seem to be two choices: either *too long* inherits from an additional class that supplies the new properties, or the lexical rule stamps on those properties. Depending on the choice, one might in fact have to have two lexical rules for *too*, and two very similar rules for *enough*, plus a third rule for *enough* to handle the *that S* complement.

(12) a. This article was too carelessly written to publish.
b. This article was written too carelessly to publish.

(13) a. He was too thoroughly embarrassed by his blunder to speak again.
b. He had slipped too often on that sidewalk to trust his footing.

The lexical rule that would combine *too* with *careless* (or *carelessly*) would produce a new lexical item which included on its subcat list the VP/NP which is correlated with the presence of *too*. This could account for (12b), but does not predict the grammaticality of (12a), where *too carelessly* is not contiguous with *to publish*. The problem is that *too carelessly* in (12a) is not the head of the phrase in which the VP/NP complement licensed by *too* appears. Thus, while one could generalize a lexical rule for *too* so that it combined with either adjectives or adverbs, one would not want to propose an independent lexical rule that made a lexical item for *too carelessly written*. The generalization, not expressible by lexical rule, is that the presence of *too* in a phrase introduces a subcat which may be used in that local phrase, or passed up to be used in a containing phrase.

Given that the lexical rule is not general enough, some other mechanism must be provided to merge information in the subcat list of an adjunct with that of the head. In fact, such a mechanism was proposed in the discussion above of constructions like *easy man to please*, where I introduced a lexically constrained convention for transferring subcat information from an adjunct to its head. That same convention will serve as the basis for an account of the *too/enough* constructions that is quite analogous to the one given above for more familiar *easy* adjectives.

One might well argue that a lexical rule could still be used to combine *too* or *enough* with an adjective or adverb, leaving to the Transferable Subcat Convention the task of propagating information about the subcats introduced by *too/enough*. For the sake of simplicity, I do not adopt this split approach here; first, because several lexical rules would be needed to handle the two kinds of *too/enough*; and second, because there are no idiosyncratic exceptions for this construction.

6.3.2 Phrasal analysis of "too/enough" constructions

The convention for *transferable* subcats provides the basis for giving a phrasal analysis of the examples in (5-6) which will also account for the examples in (12). To start with, I assume there will be two distinct lexical entries for *too*, one for examples (5a-b), and one for (5c-d). The first of these entries, given in (15), identifies *too* as an adverb that inherits from the same class that *easy* does, to capture the similarities noted in (3-4) above.

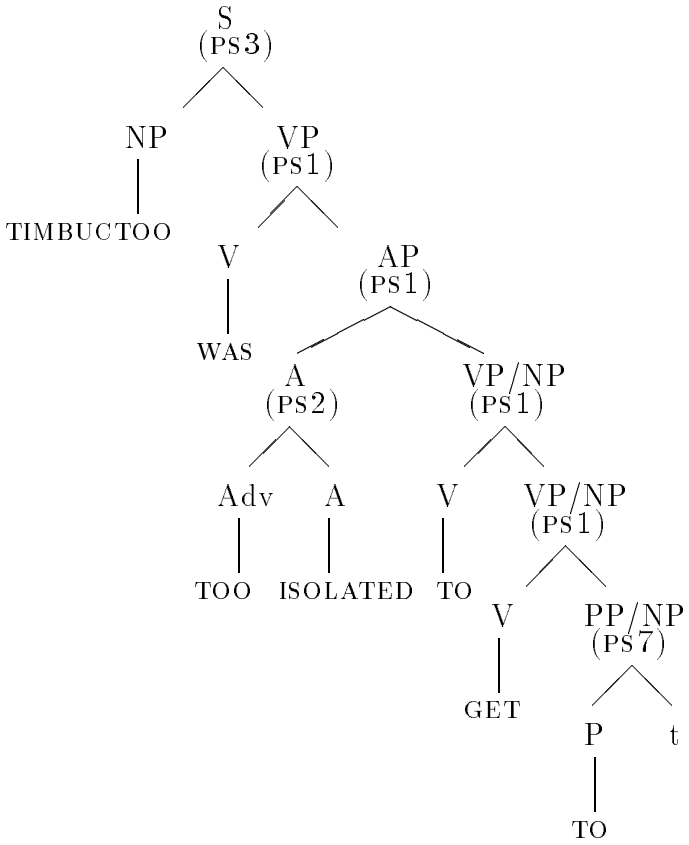
(15) Non-redundant entry for *too* as in *too long (for me) to memorize*

TOO-SLASH	
Superclasses	ADVERB, EASY-SLASH
Spelling	"too"
Phonology	...
Semantics	...

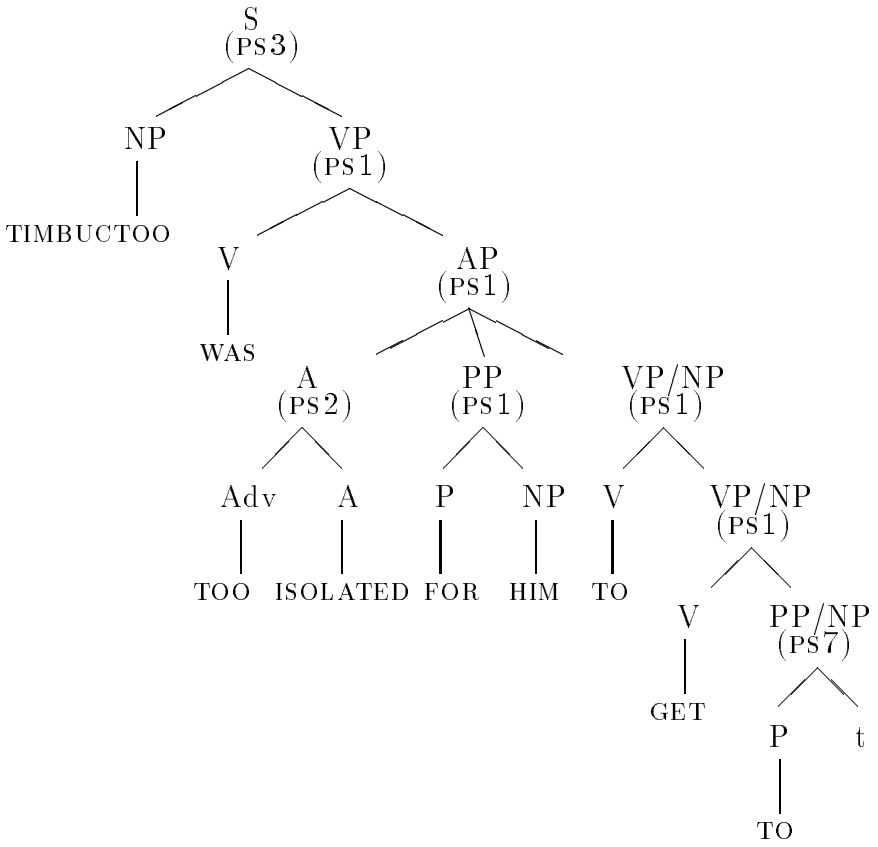
Like *easy*, this entry for *too* inherits two transferable subcats (the For-PP and the VP/NP), so they do not have to be introduced as values of the Complements attribute in the lexical entry. Though I do not give a precise semantics for this entry, it must be the case that when this *too* appears as a modifier of an adjective like *isolated*, the subject of the adjective must be unified with the subject subcat of *too*, and hence with the NP gap that subject is co-indexed with. Formalizing this semantic property is well outside the scope of this study, but I will assume that this subject-subject unification can be effected.

Now assuming that the class definition for adjectives includes mention of an adverbial adjunct, needed for simple cases like *completely isolated*, the adverb *too* can appear as a modifier of *isolated*, giving rise to structures like that in (16a-b), annotated as usual with the relevant PS rules.

(16) a. Timbuctoo was too isolated to get to.



(16) b. Timbuctoo was too isolated for him to get to.



Since the first entry for *too* belongs to the EASY-SLASH class, the other entry for *too*, given in (17), is predicted by the LR-EASY lexical rule to exist, and to have just the properties that it does.

(17) Entry for *too* as in *too long (for me) to memorize it*

TOO-BASIC	
Superclasses	ADVERB, EASY-IT
Spelling	"too"
Phonology	...
Semantics	...

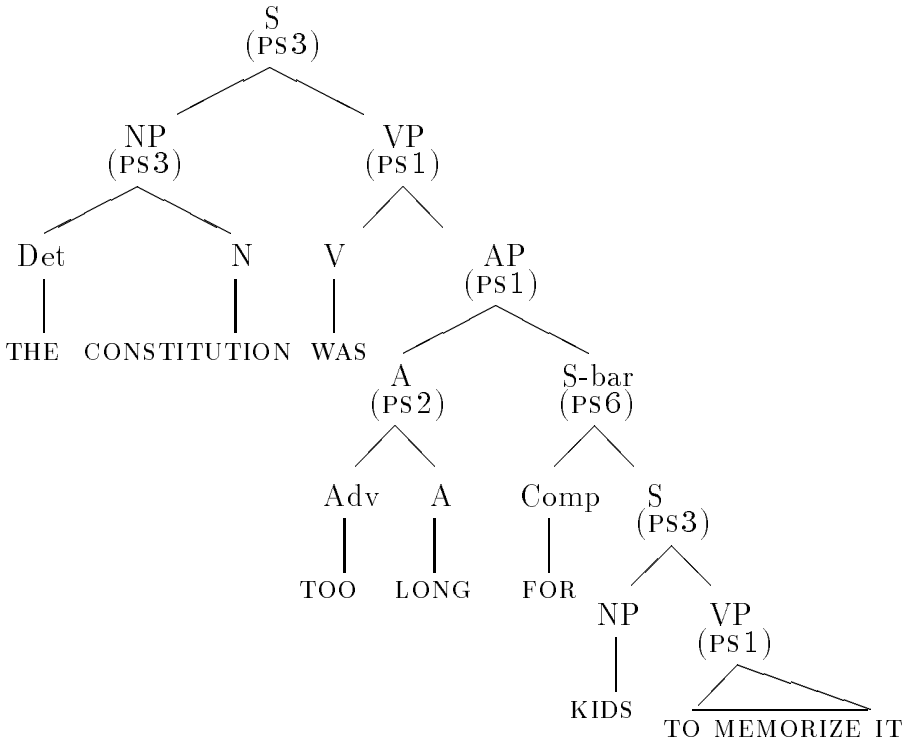
As illustrated above in (5c-d), this second *too* takes an infinitival complement which is either a VP or a for-S, like other members of the S-INF-IT class. That this

complement can be a for-S is illustrated in (18a), analogous to the the examples in (18b-c) with other, more familiar members of this class.

(18) a. John is too obscure for there to be a book about him.
b. It is easy for there to be a flaw in that argument.
c. It is possible for there to be a flaw in that argument.

The structure for sentence (5d) using this second entry for *too* is given in (19).

(19) The Constitution was too long for kids to memorize it.



The lexical entries for *enough* will look very much like those for *too*, but with one or two minor differences. One unusual characteristic of *enough* is that it follows the head it modifies, rather than preceding it as do most lexical adjuncts. Maintaining the assumption that what distinguishes pre-head from post-head adjuncts is the value for the LEXICAL attribute, and that the default value for this feature is positive, the lexical entries for *enough* must override that default. While this may seem counter-intuitive, there are no ill effects from making *enough* (LEXICAL –),

since it never needs to occur as the head of a phrase using the PS1 rule. This distinction will ensure that *too* precedes its head, while *enough* follows its head, consistent with the examples in (20).

(20) a. The river was shallow enough to cross.
b. *The river was enough shallow to cross.
c. The river was too deep to cross.
d. *The river was deep too to cross.

I give in (21-22) the two entries for *enough* that are analogous to the entries for *too*.[10]

(21) Entry for *enough* as in *short enough (for me) to memorize*

ENOUGH-SLASH	
Superclasses	ADVERB, EASY-SLASH
Spelling	"enough"
Phonology	...
Atomic-features	(LEXICAL –)
Semantics	...

(22) Entry for *enough* as in *short enough (for me) to memorize it*

ENOUGH-BASIC	
Superclasses	ADVERB, EASY-IT
Spelling	"enough"
Phonology	...
Atomic-features	(LEXICAL –)
Semantics	...

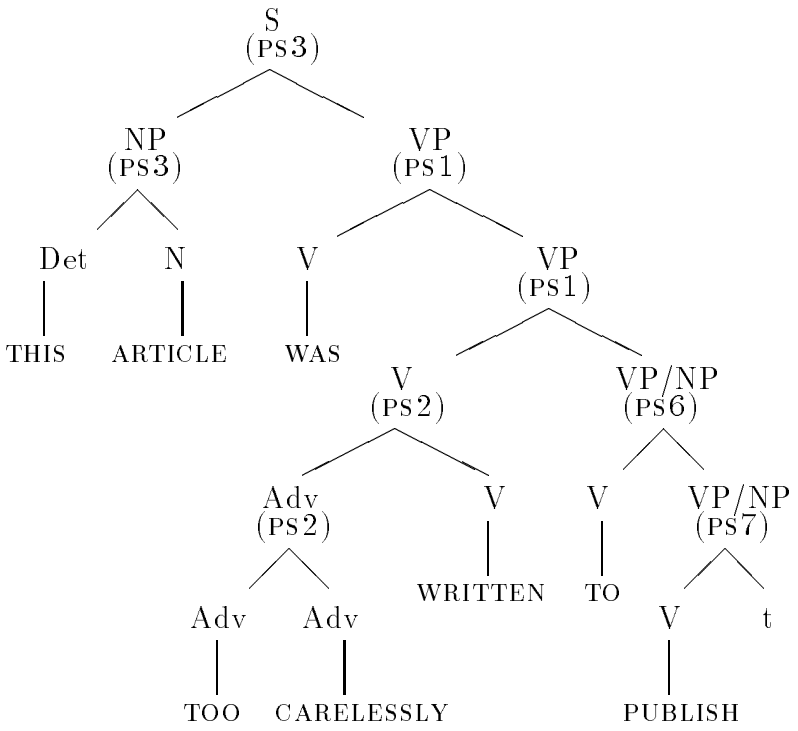
This analysis of *too* and *enough*, while motivated by the simple examples in (5-6) above, also predicts the grammaticality of examples like those in (12), given that adverbs can themselves be modified by certain adverbs, as illustrated in (23).

[10] For completeness, I note that there will need to be yet a third entry for *enough*, to account for the example given in (9a) above, where the verbal complement of *enough* is a that-S rather than infintival (a variant noted by Jackendoff 1972:228). This third entry will be just like the other two, except that instead of belonging to the class EASY-SLASH or EASY-IT, the entry will identify S-NORM as its other superclass (besides ADVERB), thereby inheriting the default that-S complement.

(23) a. John very slowly crossed the street.
b. Mary picked up her fork extremely quickly.

Assuming that *too* and *enough* belong to this subclass of adverbs, the phrase structure given in (24) follows directly.

(24) This article was too carelessly written to publish.



6.3.3 More than "enough"

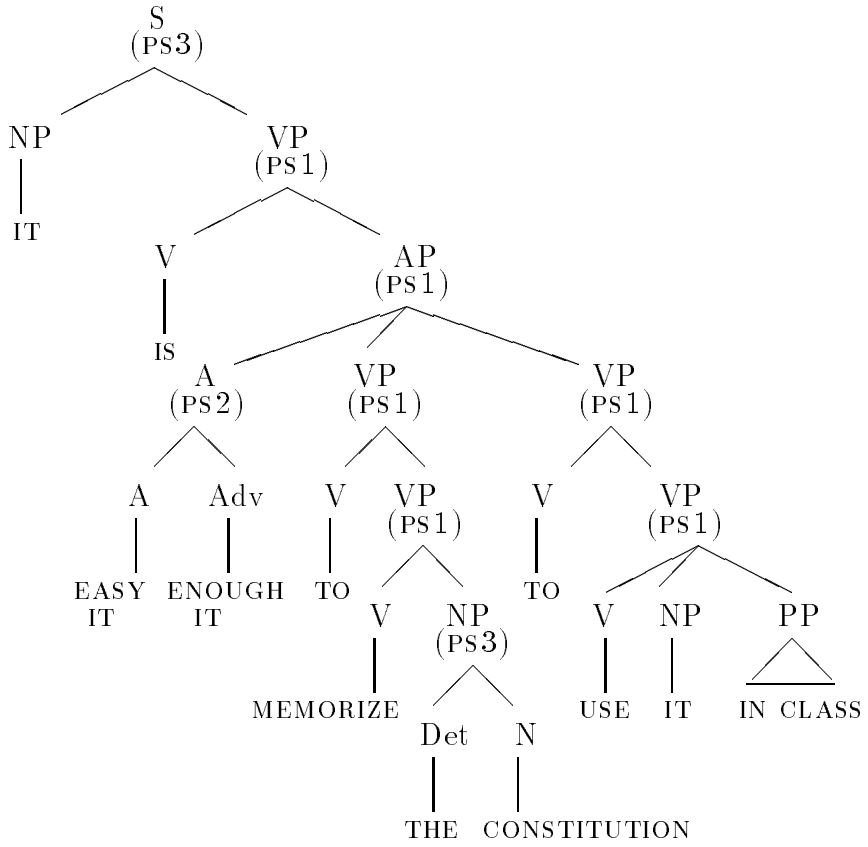
Having sketched the pieces of this analysis, we can now examine the way in which this treatment of *too* and *enough* interacts with that presented above for *easy* adjectives. Consider the examples in (25).

(25) a. It is easy enough to memorize the Constitution to use it in class.
b. *It is easy enough to memorize the Constitution to use in class.
c. The Constitution is easy enough to memorize to use it in class.
d. The Constitution is easy enough to memorize to use in class.

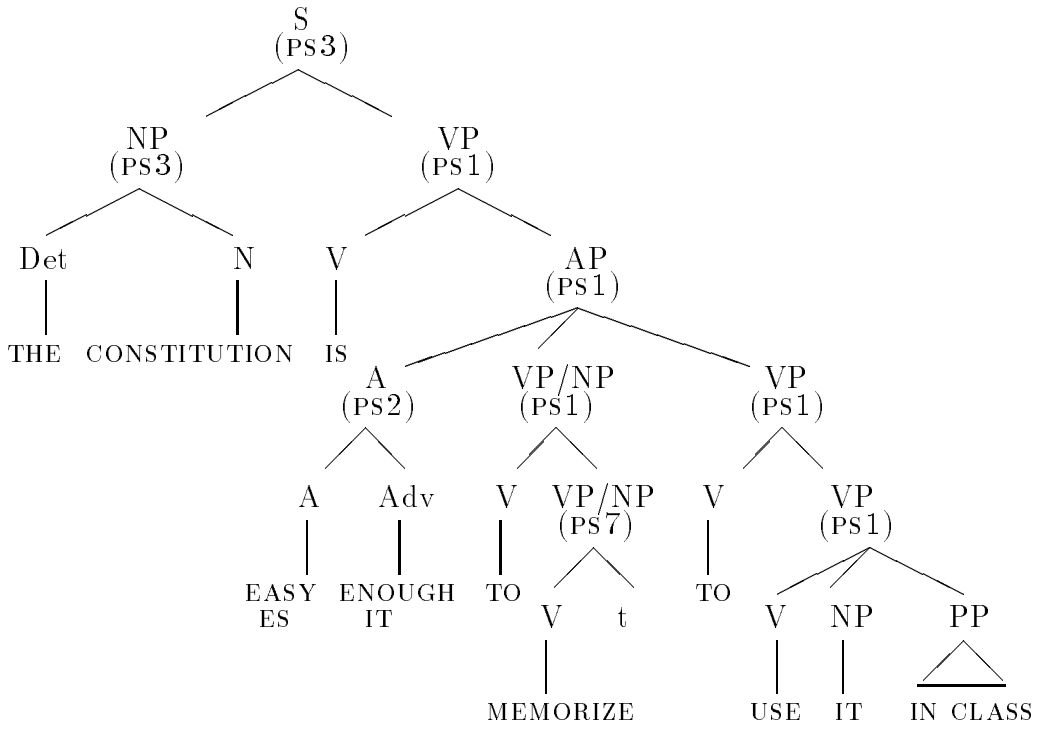
These examples illustrate the four possible combinations of the two lexical entries for *easy* and the two entries for *enough*. Of the four, only (25b) is ungrammatical; more precisely, it is semantically ill-formed. Intuitively, (25b) is blocked because the *easy* of this example must belong to the EASY-IT class, given the unslashed complement *to memorize the Constitution* and the expletive *it* subject. However, the *enough* here must belong the the EASY-SLASH class, given the VP/NP complement *to use in class*, but this *enough* has a non-expletive subject subcat, co-indexed with the NP gap in its VP/NP complement. That co-indexing will be lost if the subject subcat of *easy* is unified with the subject subcat of *enough*, since the expletive *it* has no index. On the other hand, (25c) is fine since in this case it is the subject subcat of *enough* which has no index, so no information will be lost when the non-expletive subject of this *easy* tries to be co-indexed with *enough*'s expletive subject. Again, making these semantic intuitions precise is outside the scope of this discussion.

I give the trees for (25a,c,d) annotated with the rules used, and indicating which of the two varieties is being used for both *easy* and *enough*, using the subscript *IT* for the non-slashed entries and *ES* for EASY-SLASH.

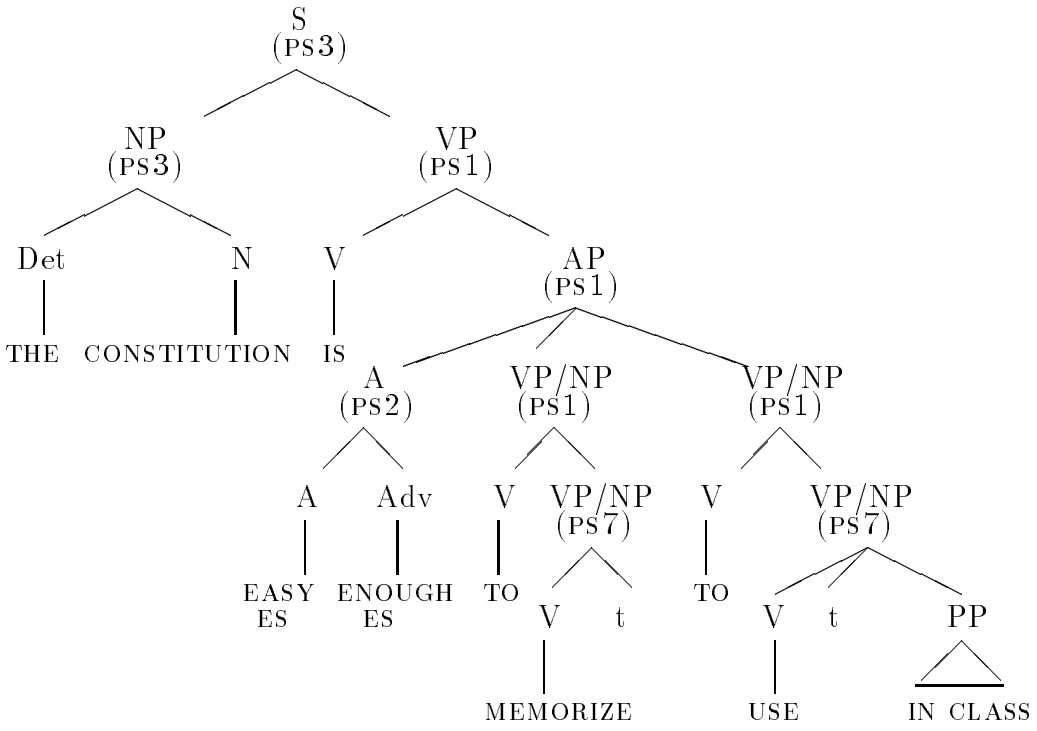
(53) It is easy enough to memorize the Constitution to use it in class.



(54) The Constitution is easy enough to memorize to use it in class.



(55) The Constitution is easy enough to memorize to use in class.



Summary

I began this section by considering the possibility of accounting for *too/enough* constructions by providing a special set of lexical rules that applied to adjectives, constructing for each adjective in the language a new adjective whose morphology (and spelling) was the result of prefixing *too* or suffixing *enough* to the original adjective, and whose Complements list was augmented to include the additional complements associated with the relevant *too* or *enough*. I spared myself the task of formulating this collection of lexical rules, by observing that such an analysis was incapable of being extended to account for the grammaticality of more complex examples.

Instead, I showed how the analysis of *easy* constructions given in the previous section could also account for *too/enough* constructions without additional stipulations, once the proper lexical entries were proposed for *too* and *enough*. I also

suggested in passing that this kind of analysis should also be able to account for constructions involving *more*, *as*, and *so*, all of which appear in constructions headed by adjectives, and including post-head complements which are licensed by the presence of one of these words. While it does seem that such an extension of this analysis holds promise, I do not pursue it here.

What this exercise with *too* and *enough* shows is that the machinery of word classes and lexical rules, while originally motivated as a means of capturing generalizations about familiar complementation properties and relationships among verbs and nouns, can be used to provide a revealing account of the properties of unusual, two-of-a-kind lexical items like *too* and *enough*. Indeed, many of the properties of these apparently idiosyncratic lexical items turn out to be just what they should be, once the entries are identified as belonging to independently motivated word classes.

Of course, to make the account above completely satisfying, one would need to provide a formal representation of the semantic properties of *too* and *enough*, properties I have only sketched here. As usual, I direct the reader to the work of Pollard and Sag for a presentation of a semantic framework consistent with the assumptions I have made.

6.4 Sonatas and double gaps

One final class of examples involving *easy* adjectives is familiar from Jackendoff 1972:227, who noted the existence of sentences like those in (1), where the head of the verb phrase following the adjective takes two complements, providing two possible sites for an NP gap inside the VP.

(1) a. It is easy to play sonatas on this violin.
b. Sonatas are easy to play on this violin.
c. This violin is easy to play sonatas on.

On the analysis given above, the *easy* of (33a) belongs to the EASY-IT class, and the *easy* in both (33b) and (33c) is the corresponding adjective in the EASY-SLASH class, related by the LR-EASY rule to the *easy* of (33a).

What is interesting about these examples is that they interact with the analysis of unbounded dependencies, providing motivation for treating the SLASH feature as stack-valued, taking as its value an ordered list of category specifications, where (informally) the most recent specification pushed onto that list is the first to be pulled off of the list when a filler is found. To see this interaction, consider the examples in (2-3).

(2) a. Which sonatas is it easy to play on this violin?
b. Which violin is it easy to play these sonatas on?

(3) a. Which violin are these sonatas easy to play on?
b. *Which sonatas is this violin easy to play on?

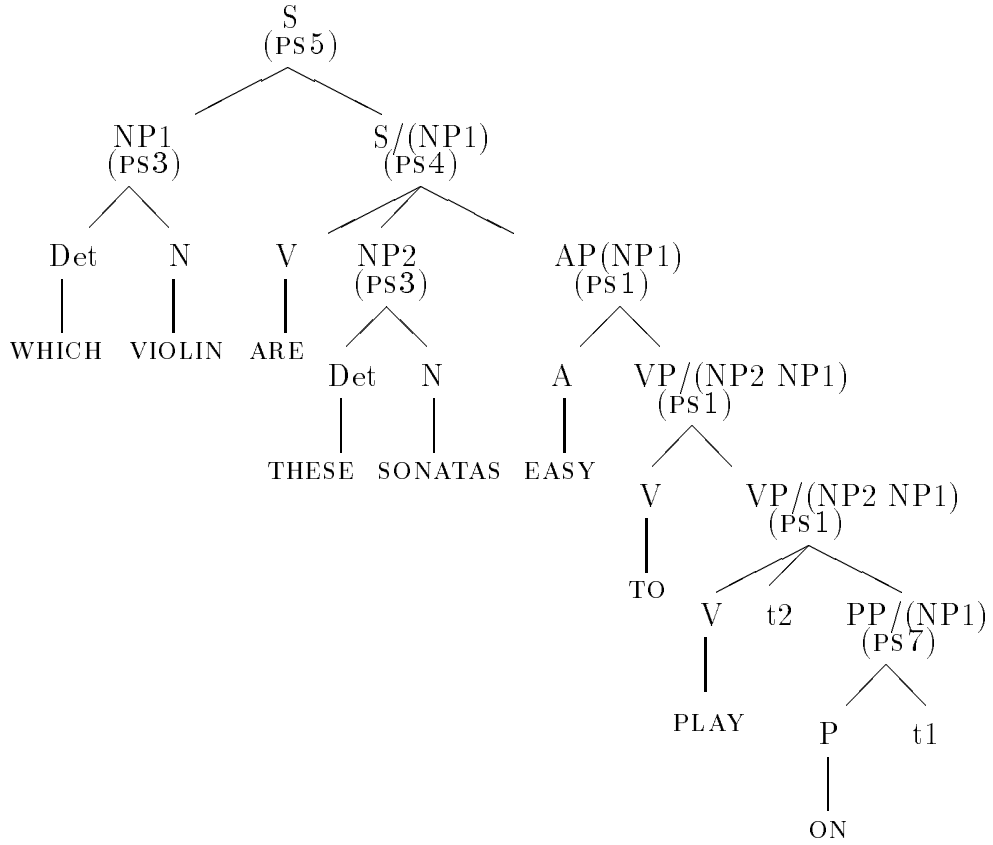
The examples in (2) are both acceptable; both employ the EASY-IT adjective *easy*, and both have an NP gap inside the VP complement of *easy*. Their grammaticality is just as expected given the definition of the EASY-IT class given above.

What I will account for here is the difference between the two examples in (3), where the WH-question corresponding to (1b) is good, but the question in (3b) corresponding to (1c) is not. Intuitively, the difference is that the filler-gap dependencies in (3a) are nested, while those in (3b) are not. The account I give will provide a formal mechanism for ensuring the nesting, a mechanism that makes a

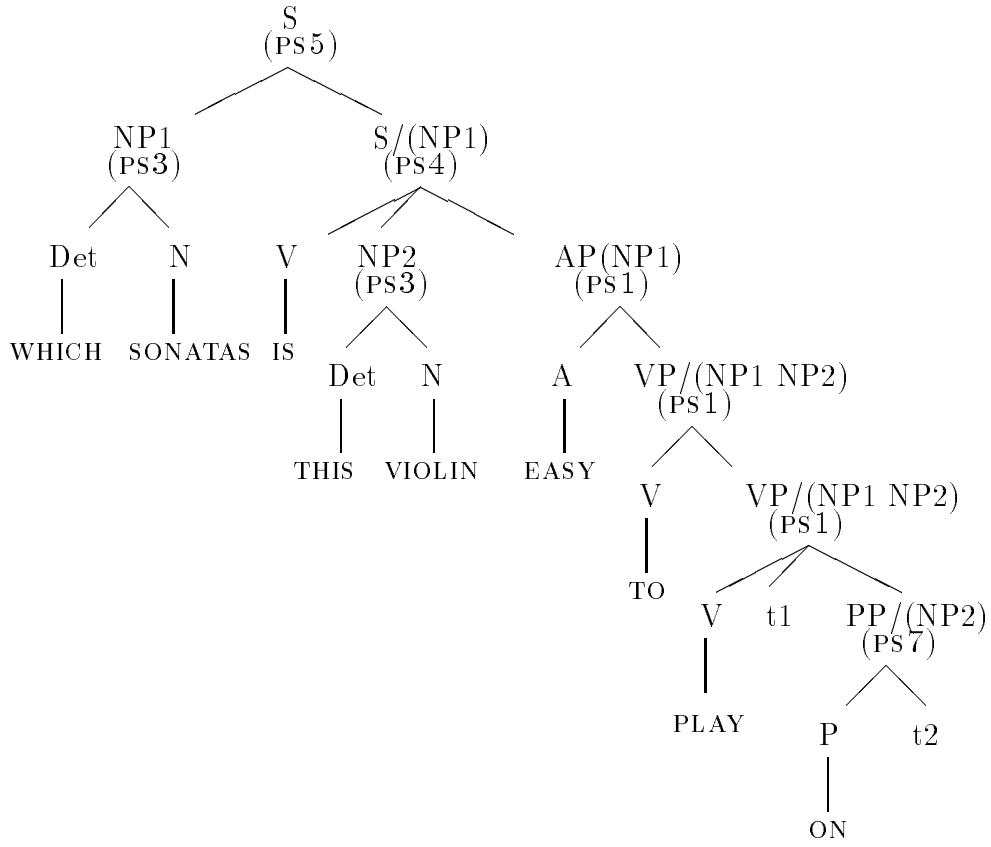
clear (and correct) prediction about more complex examples, as I will show. If, as Maling and Zaenen 1982:247ff have argued, such nesting of dependencies is not a language-independent property, the mechanism I introduce will have to be loosened, so individual languages, such as the Scandinavian languages, can provide for crossed dependencies.

To help see the relevant difference between the two examples in (3), I provide the phrase structures that would be provided for each, if both were grammatical. In each of the trees, I index separately the two distinct SLASH dependencies.

(4) Which violin are these sonatas easy to play on?



(5) *Which sonatas is this violin easy to play on?



The structure for (4) makes it clear that the SLASH attribute must be permitted to have more than a single category description as its value. Since the node in (4) dominates *play t on t*, a verb phrase with two distinct NP gaps, the description of that node must allow information about both of those gaps to be carried up through that VP node. If the SLASH attribute is to contain information about both kinds of gaps at once, the two need to be kept distinct. And as (5) suggests, information must also be conveyed about the order in which information about each of the two gaps was collected. As the nesting hypthosis predicts, the category description that was most recently added to the SLASH attribute is the first gap to be accounted for, either by matching against a lexically specified subcat that calls for such a gap, or by matching the slash requirement in a PS rule, such as the top-level linking rule. That is, there is a "last in, first out" ordering needed for the category descriptions

in the SLASH attribute; I will adopt the common term *stack-valued* to describe an attribute whose value is this kind of "last in, first out" list.

Assuming such a stack-valued SLASH feature, to account for the grammaticality of (4) and the unacceptabiltiy of (5), I turn now to some more complex examples, to test the predictions made by treating SLASH in this way. Note that on this analysis, the indices that I include in the trees below are merely for the convenience of the reader; the machinery that accounts for each gap does not do so with indices, but by making use of the stack-valued property of SLASH.

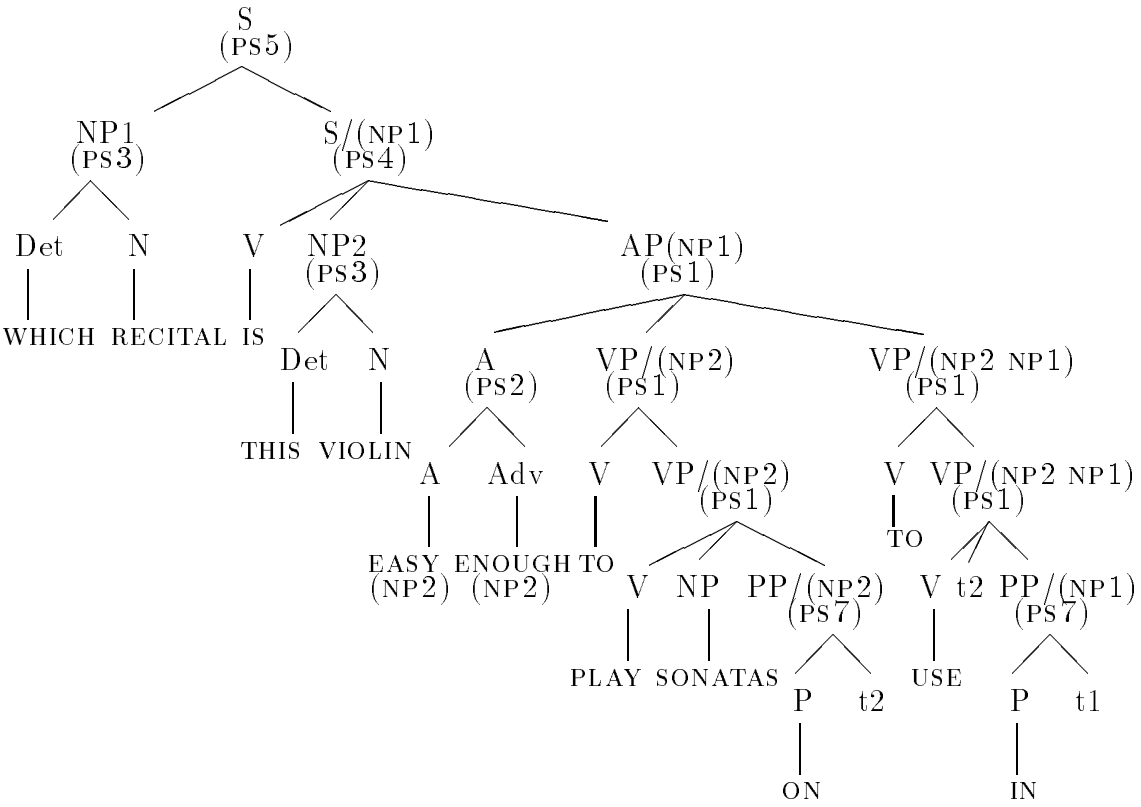
The additional complexity arises from again combining the *easy* constructions with the *too/enough* constructions, as in the previous section. On the assumption that SLASH is a stack-valued feature, the contrasts in (6-7) are as predicted, since the dependencies in the (b) examples are nested, consistent with the stack-valued mechanism, unlike the (c) examples, which show crossed filler-gap dependencies.

(6) a. These sonatas are too easy to play on the violin to use in the recital.
b. Which recital are these sonatas too easy to play on the violin to use in?
c. *Which violin are these sonatas too easy to play on to use in the recital?

(7) a. This violin is easy enough to play sonatas on to use in the recital.
b. Which recital is this violin easy enough to play sonatas on to use in?
c. *Which sonatas is this violin easy enough to play on to use in the recital?

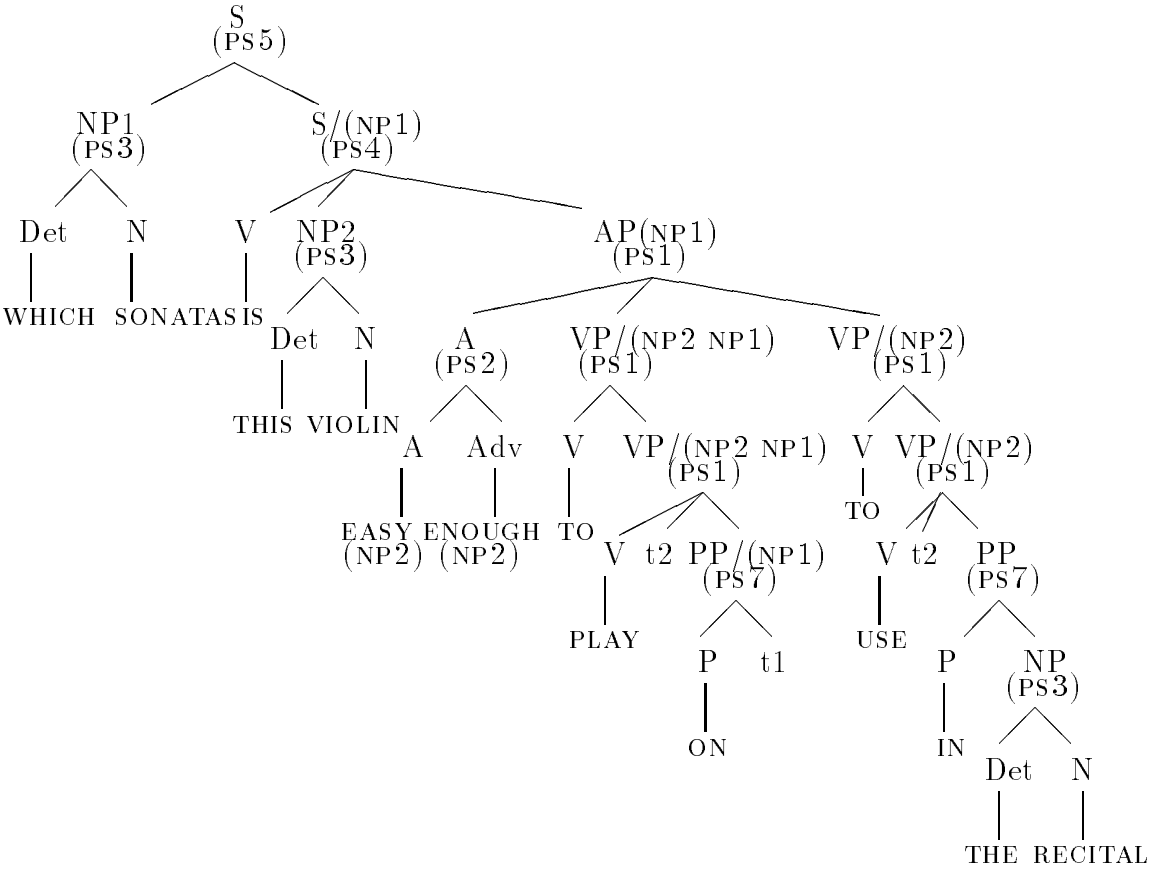
To see the difference in nested vs. crossed dependencies, and the way in which the stack-valued property of the SLASH attribute allows the first but not the second, consider the phrase structures for (7b-c) given in (8-9).

(8) Which recital is this violin easy enough to play sonatas on to use in?



In this structure the node dominating *easy enough* requires two VP complements, each containing an NP gap (NP2), where both of those gaps are co-indexed with the subject of *easy enough*, which is *this violin*. Since *easy enough* lexically specifies the existence of the NP2 gaps, the NP2 specification is pulled off of the SLASH stacks for each of the complement VPs, which is fine since the NP2 specification is first on each of the two SLASH lists. This leaves only the NP1 gap to be passed up in the usual fashion by the Foot Feature Principle, until it reaches the top-level linking rule, where it can be co-indexed with the filler *which recital*.

(9) *Which sonatas is this violin easy enough to play on to use in the recital?



This structure makes it clear that the only possible interpretation that can be given to (7c) is the anomalous one where there is some violin which is easy to perform on some sonatas, and that violin is to be used in the recital. This reading is forced by the stack-valued nature of the SLASH attribute, since the node dominating *easy enough* will license only the first NP specification in the SLASH attribute of each of its sister VPs, and insists that this specification be co-indexed with the subject *this violin*. That leaves only the NP1 gap in the first VP to be carried up out of the adjective phrase by the Foot Feature Principle, to be matched with the filler *which sonatas* at the top of the tree.

Summary

Double gap constructions are rare in English, making these otherwise peripheral constructions quite important in constructing an adequate account of unbounded dependencies and the mechanisms needed. Given the lack of sharp judgments for several of the examples presented above, the validity of the treatment of SLASH given here will be better tested by looking to data from languages where double gaps play a more central role. Such constructions are found in unbounded dependencies in Scandinavian languages, as discussed in Maling and Zaenen 1982. While I do not pursue an analysis of these constructions here, the intent is that stack-valued treatment of the SLASH attribute developed on the basis of the English data will be language-independent.

6.5 Tensed-VP constructions

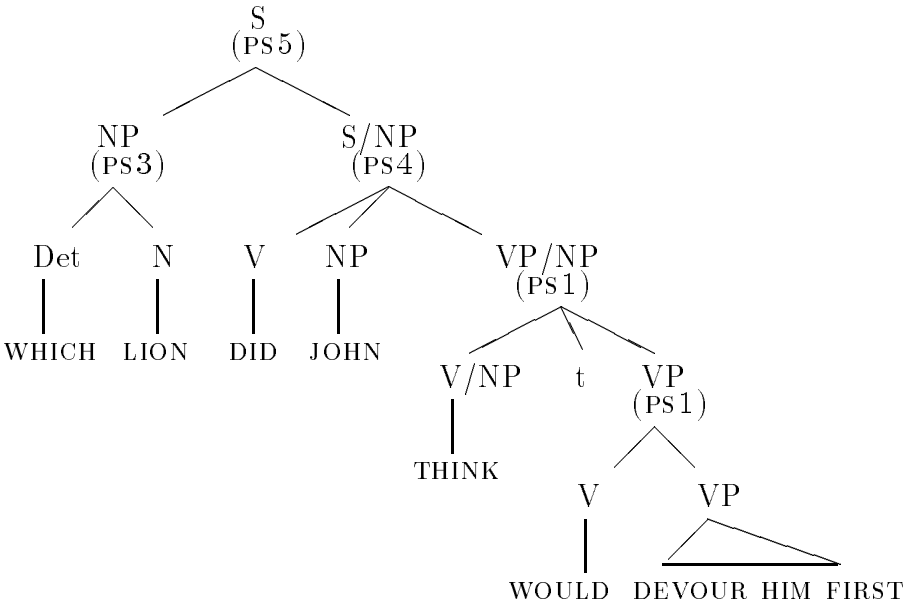
In the chapter on phrase structure, I presented the single PS rule for introducing traces in English, but noted that it would not account for examples like those in (1), where it would appear that there is a trace in place of the subject in the complement of *think*.

(1) Which lion did John think would devour him first?

To provide an account of this kind of construction without losing the explanatory power of a single phrase structure rule introducing traces as sisters of lexical heads, I borrow heavily from work of C. Pollard and T. Wasow, who developed the main features of the analysis I present here. While this analysis differs in some important respects from that of GKPS 1985, it enjoys the strengths of the GKPS analysis while also explaining two additional phenomena not accounted for by GKPS, one involving the agreement between filler and tensed VP, the other involving strong crossover constraints.

I begin by presenting in (2) the structure that I assume for the example in (1), where the verb *think*, which ordinarily takes a finite sentential complement, is followed instead by a finite VP.

(2) Which lion did John think would devour him first?



What Pollard and Wasow suggest is that the lexical entry for *think* in (1-2) specify two complements in addition to the subject: a traced NP complement and a finite VP complement. The presence of the trace in the lexical entry anchors one end of the unbounded dependency in (2), and also explains the number agreement which must hold between the filler and the head of the finite VP complement, as I show below. This entry for *think* will be related by a lexical rule to the ordinary entry for *think*, illustrated in (3), where it is followed by a finite S.

(3) Did John think the thinnest lion would devour him first?

In giving an account of this construction, I first sketch the range of predicates like *think* which can head a phrase containing a finite VP, then characterize in terms of word classes the properties of both entries for *think*, and the lexical rule that relates them.

The examples in (4-7) and (8-11) show that the predicates appearing in constructions like (2), which I will call Tensed-VP constructions, are a subset of those which can take finite sentential complements. The generalization seems to be that if a predicate can take a bare finite S, one without the complementizer *that*, then the

predicate can also appear in a Tensed-VP construction. Most verbs treat the complementizer as optional, but verbs followed by a PP seem to insist on the presence of the complementizer, as do some adjectives with expletive *it* subject.

(4) a. Who did you say would win?
b. You said John would win.
c. You said that John would win.

(5) a. Who did you tell me would win?
b. You told me John would win.
c. You told me that John would win.

(6) a. *Who did Mary announce to her family would hire her?
b. *Mary announced to her family you would hire her.
c. Mary announced to her family that you would hire her.
d. Who did Mary announce to her family that you would hire?

(7) a. *Who did it please Mary would hire him?
b. *It pleased Mary you would hire him.
c. It pleased Mary that you would hire him.
d. Who did it please Mary that you would hire?

(8) a. Name a swimmer who John is sure will win.
b. John is sure the tall swimmer will win.
c. John is sure that the tall swimmer will win.

(9) a. Name a swimmer who John is certain will win.
b. John is certain the tall swimmer will win.
c. John is certain that the tall swimmer will win.

(10) a. *Name a swimmer who it is doubtful can win.
b. *It is doubtful a tall swimmer can win.
c. It is doubtful that a tall swimmer can win.

(11) a. *Name a swimmer who it is unfortunate has won.
b. *It is unfortunate a mean swimmer has won..
c. It is unfortunate that a mean swimmer has won.

Since all predicates which can take finite S complements allow (or require) the presence of the complementizer *that*, I defined the S-NORM class back in chapter

two to capture this default, making the XComp have an obligatory complementizer. Those predicates for which the complementizer is optional, then, belong to a subclass of S-NORM which I label BARE-S and define as in (12), specifying that the sentential complement may or may not have a complementizer.

(12) BARE-S word class

BARE-S	
Superclasses	S-NORM
Complements	
XComp-features	(COMP That None)
Lexical-rules	LR-Tensed-VP

Now BARE-S predicates will include *think, say, tell, sure,* and *certain*, while *announce, please, unfortunate,* and *doubtful* belong directly to the S-NORM class. This classification not only captures the difference in optionality of *that*, but predicts that only the verbs and adjectives of the BARE-S class will have corresponding entries that can appear in the Tensed-VP constructions as in (1). This generalization is captured by associating the desired lexical rule with the BARE-S class, leaving us with the task of defining a class for these corresponding entries.

Entries like that for the *think* of (1) must introduce a complement subcat which is the controller of the finite VP complement, but which does not get instantiated in a phrase. The SLASH specification, since it must be propagated by one of the feature conventions, cannot be part of the invisible complement's subcat specification, so must be included in the features of the entry itself. To see why the invisible subcat cannot itself be the bearer of the SLASH, recall that *easy* adjectives subcategorize for a VP/NP, and the reason that SLASH value on the VP is not propagated by the Foot Feature Principle is because it was lexically specified. Maintaining this constraint on the Foot Feature Principle, we can still propagate the SLASH information for verbs like *think* by placing that SLASH specification on the lexical head, since the Head Feature Convention simply propagates all head features (including SLASH) to the mother.

To indicate the special nature of the controller subcat, I make use of the same attribute which encodes whether a given subcat is obligatory or optional, marking it

instead as *filled*, to be interpreted as follows: this subcat, like any other complement subcat, will be associated with a Complement daughter in a phrase structure rule like PS1, but will not also be associated with an actual word or phrase, as are ordinary right-hand members in a PS rule. In this respect the subcat will behave exactly like the Traced-complement element in the PS rule given above as PS7, which introduces ordinary traces. I indicate that a subcat is *filled* rather than optional or obligatory by specifying this value for the attribute STATUS, which I have implicitly assumed for all subcats, using the notational convenience of parentheses as shorthand for the value Optional, distinct from the default Obligatory value.

With this introduction, I now present in (13) the word class to which the *think* of (1) belongs, which I give the label FINITE-VP.

(13) FINITE-VP word class

FINITE-VP	
Superclasses	OBJECT-EQUI
Category-features	(SLASH α)
Complements	
XComp-features	(VFORM Finite)
DObject-features	α
DObject-status	Filled
Lexical-rules	LR-Tensed-VP

The use of the α notation here is exactly the same as its use in the two phrase structure rules that account for unbounded dependencies: the features assigned as the value of SLASH for the lexical entry are to be copies of those assigned to the direct object. Some of the direct object's feature values are, of course, the defaults provided by parent classes such as TRANSITIVE. Additional values will be supplied by the control mechanism which will ensure that the syntactic features of the controlled complement's subject are unified with those of its controller, the direct object. Since this controlled VP is finite, it will include agreement features, which will be copied on the direct object trace via control, and therefore also copied in the value of SLASH on the lexical head, given the alpha-matching stipulation in (13).

As one might have expected, the lexical rule relating members of the FINITE-VP class with members of the BARE-S class is maximally simple, mentioning only the two relevant classes.

LR-TENSED-VP
LE2-Classes – FINITE-VP = LE1-Classes – BARE-S

This machinery accounts for the judgments illustrated in (14), where the filler and the finite VP complement must agree in number. It also predicts the difference between the two examples of (15), since the direct object for object-equi verbs is by default (NFORM NORM), which fails to unify with the subject of the controlled VP in (15b), specified to be (NFORM THERE).

(14) a. Which candidate do you think is going to win?
b. *Which candidate do you think are going to win?
c. *Which candidates do you think is going to win?
d. Which candidates do you think are going to win?

(15) a. This horse, I'm just positive will win.
a. *There, I'm just positive will be lions at the zoo.

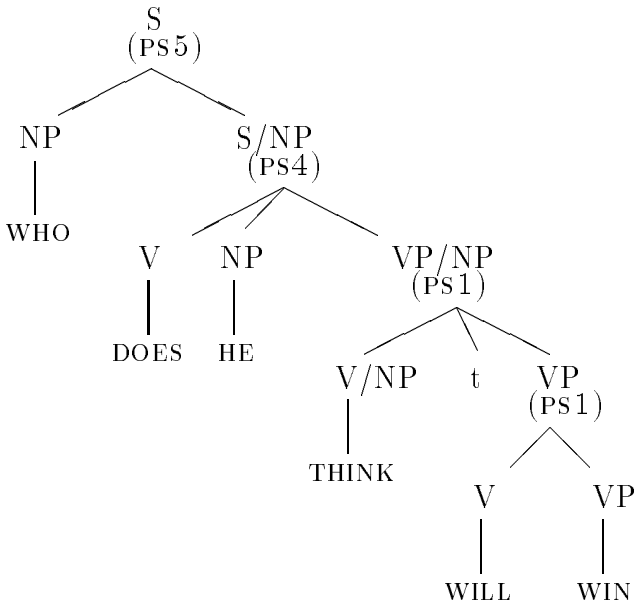
In addition to these facts about agreement, there is independent support for the presence of a trace in the Tensed-VP constructions based on facts about strong crossover, as T. Wasow points out. The examples in (16) can both be treated as violations of the same constraint, an insight due to Jackendoff which can be formulated in terms of the notion of c-command, following Reinhart 1976 and Wasow 1972. This constraint, essentially the same as the one now familiar in Government Binding theory as Principle C, says: "A referring expression must not be co-indexed with an element that c-commands it."

(16) a. *Who-1 does he-1 think will win.
b. *He-1 thinks John-1 will win.

What Wasow 1972 argued was that if (16a) is to be ruled out in the same way as (16b), then (16a) must include a trace in the verb phrase *think will win*, where traces are treated as referring expressions. With a trace present, (16a) is excluded because

the co-indexed *he* c-commands the position of the trace, which is itself co-indexed with *who*, as illustrated in (17).

(17) Who does he think will win.

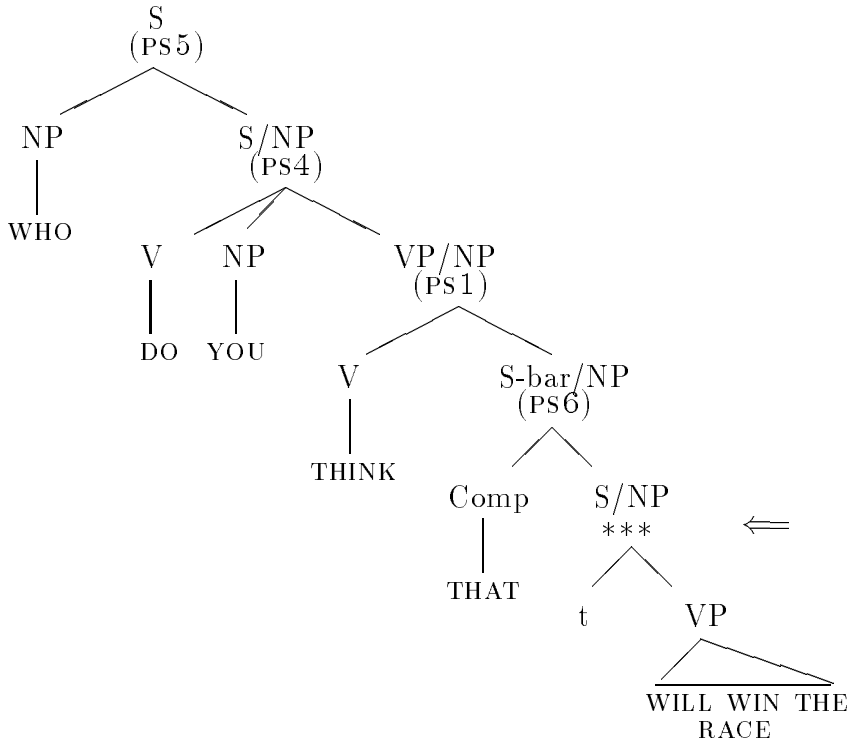


Since the presence of a direct object trace in examples like (16a) is necessary to ensure agreement between filler and VP complement by general principles, there are then two independent reasons for this trace. It should be noted here that the analysis of Tensed-VP constructions proposed by GKPS 1985 did not introduce any trace at all, leaving no obvious explanation for the unacceptability of (16a). Nor did that analysis provide any account of the agreement facts noted in (14).

What the GKPS account and this one both provide is an explanation for the familiar "that-trace" data illustrated in (18), excluding (18a) because the presence of the complementizer *that* shows that the complement of *think* must be an S, which means the trace would have to be in subject position, but my analysis provides no way of introducing a trace sister to a VP head. I illustrate in (19) what the structure of (18a) would have to be, indicating the phrase which could not be admitted by the grammar given here.

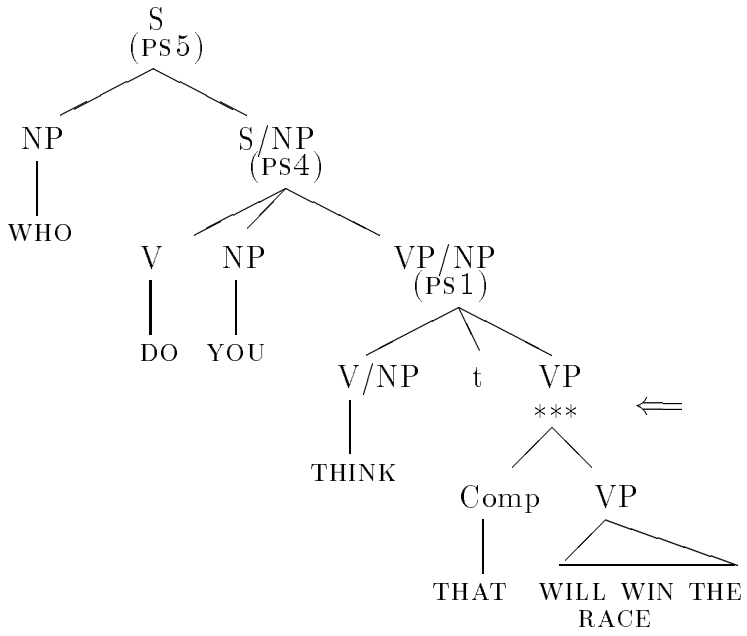
(18) a. *Who do you think that will win the race?
b. Who do you think will win the race?
c. Which race do you think that John will win?

(19) *Who do you think that will win the race?



Notice that an alternative structure, given in (20), is not possible for English because the lexical entry for the complementizer *that* requires a sentential sister, not one that is a VP.

(20) *Who do you think that will win the race?



However, it may be that some languages differ from English in having complementizers that select for either S or VP. On this view, the lack of "that-trace" effects noted by Perlmutter 1971 for Dutch, and by Maling and Zaenen 1978 for Icelandic, can be represented as a lexical property of the complementizers in these two languages, rather than treating the examples in (21) as containing an S/NP with a subject trace.

(21) a. Wie vertelde je dat gekomen was? (Dutch)
who said you that come was
Who did you say that had come?

b. Hver sagδir θu, aδ vaeri kominn til Reykjavikur? (Icelandic)
who said you that was come to Reykjavik-gen.
Who did you say that had come to Reykjavik?

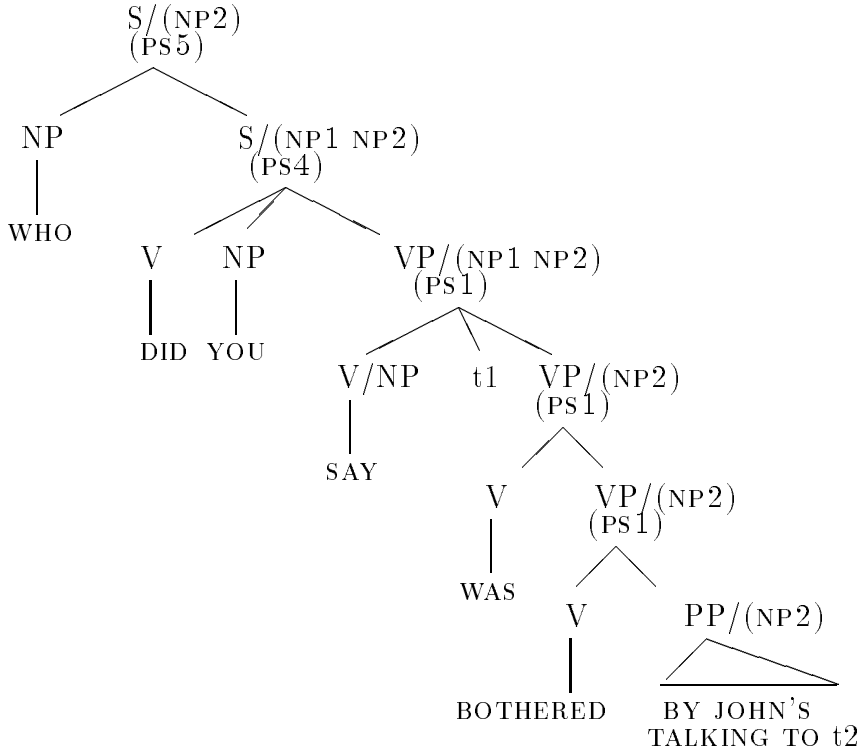
The structure for (21a), then, would be analogous to that in (20), where the relevant entry for *vertelde* would introduce a trace and a finite VP, but the finite VP

would be marked for the complementizer *dat*. The same story would hold for *sagðir* and the Icelandic complementizer *að*, assuming that both languages have a lexical rule corresponding to the LR-Tensed-VP rule given above for English. If this kind of analysis can be provided for Dutch and Icelandic, it locates in the lexical entries for the complementizers the reason for the difference in grammaticality between (18a) and the examples in (21).

6.5.1 Parasitic gaps

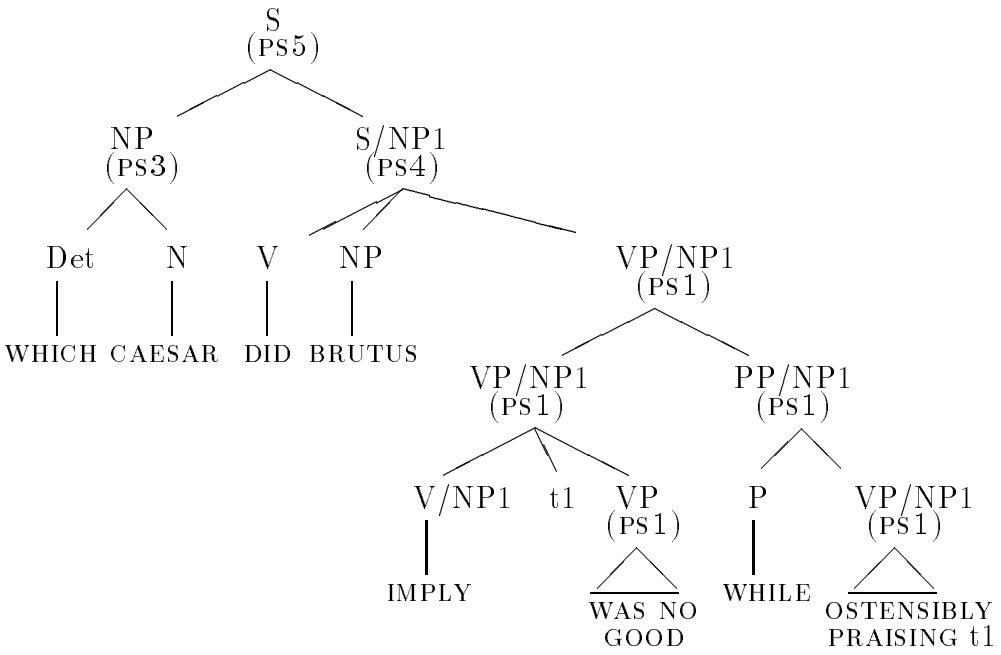
One of the virtues of the GKPS analysis was its correct predictions about some rather subtle distinctions in the grammaticality of parasitic gap constructions. GKPS 1985:165f explain that the example in (22), originally from Engdahl 1983:21, is predicted to be ungrammatical given the assumption that SLASH is a HEAD feature. The analysis I have sketched here makes the same prediction, since I share their assumption about the distribution of SLASH and do not treat the complement headed by *was* as slashed. To see why (22) is bad, examine the tree in (23), where the prepositional phrase *by John's talking to t* contains an NP gap, which I index NP2, that is never given a filler. A sentence can, of course, contain multiple gaps, as we saw with the examples involving *easy* adjectives, but each of the gaps in a well-formed sentence must be licensed by a filler, a PS rule, or a lexical entry. The NP1 gap introduced by the verb *say* does get its filler, the string-initial *who*. But the NP2 gap is not licensed anywhere, so (22) is syntactically ill-formed.

(22) *Who did you say was bothered by John's talking to.



In contrast the example in (24), also originally from Engdahl 1983, is well-formed, and predicted to be so on the analysis given in GKPS.

(24) Which caesar did Brutus imply was no good while ostensibly praising?



The more general analysis given here also predicts the grammaticality of (24), and assigns the structure sketched above, where the *while* phrase modifies the VP *imply was no good*, and where both phrases contain a gap with the same index, the result of treating SLASH as a head feature.

Chapter 7

Summary

One central theme of this work has been to articulate the division of labor within the lexicon and the grammar, introducing the notion of word classes within the lexicon, and linking them to lexical rules for both derivation and inflection, then separating the work of these lexical rules from that of phrase structure rules with their associated feature principles. While English was the chief source of data used in developing this structured lexicon, I have attempted at several points to suggest how this work will extend to explanations of phenomena in other languages as well.

In the course of developing the form and content of lexical entries, I proposed several properties of subcategorized-for elements that a lexical entry may specify, either idiosyncratically or by virtue of that entry's membership in a word class that determines one of these properties. Of the four introduced here, three are familiar, and one is novel: first, a complement or adjunct subcat may be marked as either obligatory or optional; second, the subcat may consist of a disjunction, as with the Determiner/Possessive-NP for common nouns; third, an optional subcat's semantic argument position may be existentially bound (perhaps in several distinct ways); and fourth, a subcat may be marked as transferable, as I proposed in the analysis given for *tough*-adjectives. Further work may well reveal other properties that lexical entries can assign to subcats; the mechanisms proposed here for representing such properties, making use of inherited defaults that may be overridden, should extend naturally to additional such properties.

I also motivated a strong constraint on lexical representation which only allows a lexical entry to specify for a given subcat whether or not that complement or adjunct lacks its final obligatory argument. I incorporated this restriction into the syntactic feature mechanisms for representing categories, using the two features LEXICAL and COMPLETE, an approach which has the additional benefit of allowing us to dispense with the familiar but relatively unconstrained X-bar device for distinguishing lexical and phrasal categories.

Given a relatively rich hierarchy of word classes that capture many of the generalizations within the lexicon, the task of formulating lexical rules is simplified, with each rule expressing a relationship holding between minimally-specified members of two such word classes. In practice, the formulation of these rules interacts closely with the definition of the relevant word classes, resulting in a balance of power that sharply constrains (in the right ways, one hopes) the expressive power of this framework.

Further study will be required to determine what kinds of syntactic or semantic constraints may be needed in formulating lexical rules, beyond those implicit in the relevant class definitions. The examples of rules given here did not require such ad hoc constraints, suggesting that it may be possible to exclude in principle the use of such constraints on lexical rules, requiring that any restriction on application of a lexical rule be expressed in word class definitions which are in general independently motivated.

Additional work is also needed to determine whether there are distinct clusters of lexical rules that are formally distinct; this notion of classifying lexical rules is familiar from the traditional distinction between inflectional and derivational rules, and also from more recent proposals such as Wasow's (1980) major/minor rule dichotomy. I have not examined enough rules in this study to provide an adequate base from which to develop such a classification; rather, I sought to illustrate with a few detailed examples the considerations relevant to the proper use of word classes, lexical rules, and phrase structure rules, as tools necessary for providing precise analyses of data within a coherent, constrained syntactic framework.

Finally, much remains to be done within the word class hierarchy on several

fronts, including the capturing of default information about thematic roles within these classes; and the separation of language-particular and language-independent properties within the hierarchy. Work in both of these areas will enhance the generality of the framework I have developed here.

References

Anderson, S. 1977. "Comments on the paper by Wasow," P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax.* New York: Academic Press, 361-77.

Anderson, S. 1982. "Where's morphology?," *Linguistic Inquiry* 13:571-612.

Bach, E. 1981. "Discontinuous constituents in generalized categorial grammar," V. A. Burke and J. Pustejovsky, eds., *Proceedings of the 11th Annual Meeting of the North Eastern Linguistic Society.* University of Massachusetts at Amherst: Dept. of Linguistics, 1-12.

Bach, E. 1983. "On the relationship between word grammar and phrase grammar," *Natural Language and Linguistic Theory* 1:65-89.

Baker, C.L. 1979. "Syntactic theory and the projection problem," *Linguistic Inquiry* 10:533-581.

Berman, A. 1973. *Adjectives and Adjective Complement Constructions in English.* Doctoral dissertation, Harvard University.

Borsley, R. 1983. "A Welsh agreement process and the status of VP and S," G. Gazdar, E. Klein, and G. Pullum, eds., *Order, Concord, and Constituency.* Dordrecht: Foris Publications, 57-74.

Bresnan, J. 1971. "Sentence stress and syntactic transformations," *Language* 47: 257-281.

Bresnan, J. 1978. "A realistic transformational grammar," M. Halle et al., eds., *Linguistic Theory and Psychological Reality.* Cambridge, Mass.: MIT Press, 1-58.

Bresnan, J. 1980. "Polyadicity: Part I of a theory of lexical rules and representations," T. Hoekstra, H. v.d. Hulst, and M. Moortgat, eds., *Lexical Grammar.* Dordrecht: Foris Publications, 97-121.

Bresnan, J. 1982a. "The passive in lexical theory," J. Bresnan, ed., *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: MIT Press, 3-82.

Bresnan, J. 1982b. "Control and complementation," *Lingusitic Inquiry* 13:343-434.

Chomsky, N. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Chomsky, N. 1970. "Remarks on nominalization," R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar.* Waltham, Mass.: Ginn and Company, 184-221.

Chomsky, N. 1977. "On Wh-movement," P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax.* New York: Academic Press, 71-132.

Croft, W. 1986. *Categories and Relations in Syntax: The Clause-Level Organization of Information.* Doctoral dissertation, Stanford University.

Culicover, P. W., T. Wasow, and A. Akmajian, eds., 1977. *Formal Syntax.* New York: Academic Press.

Dowty, D. 1978a. "Applying Montague's views on linguistic metatheory to the structure of the lexicon," D. Farkas, W. Jacobsen, and K. Todrys, eds., *Papers from the Parasession on the Lexicon.* Chicago: Chicago Linguistic Society, 97-137.

Dowty, D. 1978b. "Governed transformations as lexical rules in a Montague Grammar," *Linguistic Inquiry* 9:393-426.

Dowty, D. 1981. "Quantification and the lexicon," M. Moortgat, H. v.d. Hulst, and T. Hoekstra, eds., *The Scope of Lexical Rules.* Dordrecht: Foris Publications, 79-106.

Dryer, M. S. 1985. "The role of thematic relations in adjectival passives," *Linguistic Inquiry* 16:320-326.

Engdahl, E. 1983. "Parasitic gaps," *Linguistics and Philosophy* 6:5-34.

Fabb, N. 1984. *Syntactic Affixation.* Doctoral dissertation. MIT, Cambridge, Massachusetts.

Flickinger, D. 1983. "Lexical heads and phrasal gaps," M. Barlow, D. Flickinger, and M. Wescoat, eds., *Proceedings of the Second West Coast Conference on Formal Linguistics.* Stanford: Stanford Linguistics Department, 89-101.

Flickinger, D., C. Pollard, and T. Wasow 1985. "Structure-sharing in lexical representation," *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, 262-267.

Gazdar, G. 1981. "Unbounded dependencies and coordinate structure," in *Linguistic Inquiry* 12:155-184.

Gazdar, G., G.K. Pullum, and I.A. Sag 1982. "Auxiliaries and related phenomena in a restrictive theory of grammar," *Language* 58:591-638.

Gazdar, G., E. Klein, G.K. Pullum, and I.A. Sag 1985. *Generalized Phrase Structure Grammar.* Cambridge, Mass.: Harvard University Press.

Green, G. 1974. *Semantics and Syntactic Regularity.* Don Mills, Ontario: Fitzhenry and Whiteside, Ltd.

Gruber, J. 1976. *Lexical Structures in Syntax and Semantics.* Amsterdam: North-Holland Publishing Company.

Harlow, S. 1983. "Celtic relatives," *York Papers in Linguistics* 10:77-121.

Hoekstra, T., H. van der Hulst, and M. Moortgat 1980. *Lexical Grammar*. Dordrecht: Foris.

Hudson, R. 1984. *Word Grammar*. Oxford and New York: Blackwell.

Hust, J. "Lexical redundancy rules and the unpassive construction," *Linguistic Analysis* 4:61-89.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Jackendoff, R. 1973. "The base rules for prepositional phrases," S. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston, Inc., 345-56.

Jackendoff, R. 1975. "Morphological and semantic regularities in the lexicon," *Language* 51:639-671.

Jacobson, Pauline 1982. "Evidence for gaps," P. Jacobson and G.K. Pullum, eds., *The Nature of Syntactic Representation*. Dordrecht: D. Reidel, 187-228.

Kajita, M. 1968. *A Generative-Transformational Study of Semi-Auxiliaries in Present-Day American English*. Tokyo: Sanseido.

Kaplan, Ron and Joan Bresnan 1982. "Lexical-Functional Grammar: A formal system for grammatical representation," J. Bresnan, ed., *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: MIT Press, 173-281.

Kiparsky, P. 1968. "Linguistic universals and linguistic change," E. Bach and R. Harms, eds., *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 171-202.

Kiparsky, P. 1982. "Lexical morphology and phonology," I.S. Yang, ed., *Linguistics in the Morning Calm*. Seoul: Hanshin.

Klein,E. and I.A. Sag 1985. "Type-driven translation," *Linguistics and Philosophy* 8:163-202.

Koutsoudas, A., G. Sanders and C. Noll 1974. "The application of phonological rules," *Language* 50:1-28.

Ladusaw, W. 1982. "Semantic constraints on the English partitive construction," D. Flickinger, M. Macken, and N. Wiegand, eds., *Proceedings of the First West Coast Conference on Formal Linguistics.* Stanford: Linguistics Department.

Lakoff, G. 1970. *Irregularity in Syntax.* New York: Holt, Rinehart, and Winston.

Lasnik, H. and R. Fiengo 1974. "Complement object deletion," *Linguistic Inquiry* 5:535-571.

Levin, B. and M. Rappaport 1986. "The formation of adjectival passives," *Linguistic Inquiry* 17.

Levin, L. 1987. "Toward a theory of relation-changing rules in LFG," unpublished manuscript, University of Pittsburgh.

Levin, L., M. Rappaport, and A. Zaenen, eds. 1983. *Papers in LFG.* Bloomington, Ind: Indiana University Linguistics Club.

Lieber, R. 1980. *On the Organization of the Lexicon.* Doctoral dissertation. MIT, Cambridge, Massachusetts.

Lieber, R. 1981. "Morphological conversion within a restrictive theory of the lexicon," M. Moortgat, H. v.d. Hulst, and T. Hoekstra, eds., *The Scope of Lexical Rules.* Dordrecht: Foris Publications, 161-200.

Maling, J. 1987. "Existential sentences in Swedish and Icelandic," unpublished manuscript, Brandeis University.

Maling, J. and A. Zaenen 1978. "The nonuniversality of a surface filter," *Linguistic Inquiry* 9:475-497.

Maling, J. and A. Zaenen 1982. "A phrase structure account of Scandinavian extraction phenomena," P. Jacobson and G.K. Pullum, eds., *The Nature of Syntactic Representation.* Dordrecht: D. Reidel, 229-282.

Marchand, H. 1969. *The Categories and Types of Present-Day English Word-Formation.* Munich: Beck.

McCawley, J. 1968. "Concerning the base component of a transformational grammar," *Foundations of Language* 4:243-269. Reprinted in Grammar and Meaning. New York: Academic Press, 35-58.

Morrill, G.V. 1985 (ms). "Are the feature percolation conventions ..." Unpublished manuscript, School of Epistemics, University of Edinburgh.

Nanni, D. L. 1980. "On the surface syntax of constructions with easy-type adjectives," *Language* 56:568-81.

Oehrle, R. 1975. *The Grammatical Status of the English Dative Alternation.* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Perlmutter, D. 1971. *Deep and Surface Structure Constraints in Syntax.* New York: Holt, Rinehart, and Winston.

Perlmutter, D. 1978. "Impersonal passives and the Unaccusative Hypothesis," *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society.* Berkeley: University of California Linguistics Department, 157-189.

Pollard, C. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and Natural Language.* Doctoral dissertation, Stanford University.

Pollard, C. 1985. "Phrase structure grammar without metarules," Goldberg, MacKaye, and Wescoat, eds., *Proceedings of the Fourth West Coast Conference on Formal Linguistics.* Stanford: Stanford Linguistics Department, 246-261.

Pollard, C. and I. A. Sag 1987 (ms). *Information-based Syntax and Semantics.* Draft in preparation. Stanford: Stanford Linguistics Department.

Postal, P. 1971. *Cross-over Phenomena.* New York: Holt, Rinehart, and Winston.

Pullum, G.K. 1982. "Free word order and phrase structure rules," J. Pustejovsky and P. Sells, eds., *Proceedings of the Twelfth Annual Meeting of the Northeastern Linguistic Society.* Amherst: Graduate Student Linguistics Association, University of Massachusetts, 209-220.

Pullum, G.K. and D. Wilson 1977. "Autonomous syntax and the analysis of auxiliaries," *Language* 53:435-63.

Quirk, R. and S. Greenbaum 1973. *A Concise Grammar of Contemporary English.* New York: Harcourt, Brace, Jovanovich, Inc.

Reinhart, T. 1976. *The Syntactic Domain of Anaphora.* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Rosenbaum, P.S. 1967. *The Grammar of English Predicate Complement Constructions.* Cambridge, Mass.: MIT Press.

Rosenberg, S. 1983. "HPRL: A language for building expert systems," IJCAI 1983 *Proceedings*, 215-217.

Sag, I. 1982. "Coordination, extraction, and Generalized Phrase Structure Grammar," *Linguistic Inquiry* 13:329-336.

Sag, I. 1986 (ms) "A semantic theory of obligatory control," unpublished manuscript. Stanford: Stanford Linguistics Department.

Sag, I. and C. Pollard 1986 (ms.) "Head-Driven Phrase Structure Grammar: An informal synopsis," class notes. Stanford: Stanford Linguistics Department.

Schachter, P. 1981. "Lovely to look at," *Linguistic Analysis* 8:431-48.

Selkirk, E. 1981. "English compounding," M. Moortgat, H. v.d. Hulst, and T. Hoekstra, eds., *The Scope of Lexical Rules.* Dordrecht: Foris Publications, 229-278.

Sells, P. 1983. "Relative clauses in Irish and Welsh" *York Papers in Linguistics* 10:159-172.

Shieber, S. 1986. *An introduction to unification-based approaches to grammar.* Stanford: Center for the Study of Language and Information.

Siegel, D. 1973. "Nonsources for UN-passives," J. Kimball, ed., *Syntax and Semantics*, vol. 2. New York: Seminar Press.

Simpson, J. 1983. "Resultatives," Levin, L., M. Rappaport, and A. Zaenen, eds., *Papers in LFG.* Bloomington, Ind: Indiana University Linguistics Club, 143-158.

Wasow, T. 1972. *Anaphoric Relations in English.* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Wasow, T. 1977. "Transformations and the lexicon," P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax.* New York: Academic Press, 327-360.

Wasow, T. 1980. "Major and minor rules in lexical grammar," T. Hoekstra, H. v.d. Hulst, and M. Moortgat, eds., *Lexical Grammar.* Dordrecht: Foris Publications, 285-312.

Wasow, T. and S. Roeper 1972. "On the subject of gerunds," *Foundations of Language* 8:44-61.

Wasow, T., I.A. Sag, and G. Nunberg 1982. "Idioms: An interim report," S. Hattori and K. Inoue, eds., *Proceedings of the XIIIth International Congress of Linguists.* The Hague: CIPL, 102-115.

Williams, E. 1981. "On the notions 'lexically related' and 'head of a word'," *Linguistic Inquiry* 12:245-274.