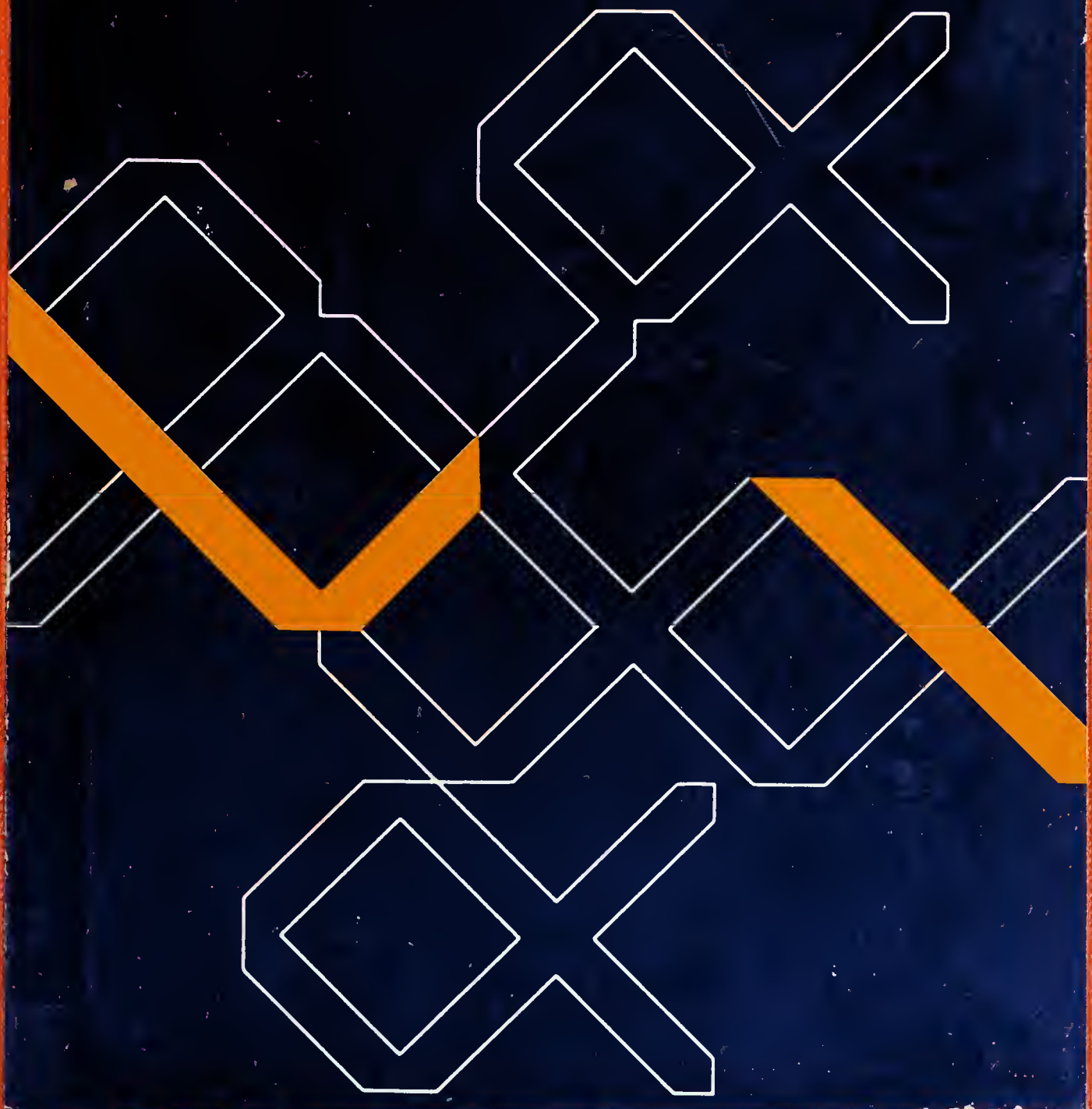
Linguistic Inquiry
Monograph Two

$\bar{X}$ Syntax: A Study of Phrase Structure

Ray Jackendoff



$\bar{X}$ Syntax: A Study of Phrase Structure

Linguistic Inquiry Monographs
Samuel Jay Keyser, general editor

1. *Word Formation in Generative Grammar*
   Mark Aronoff
2. *$\bar{X}$ Syntax: A Study of Phrase Structure*
   Ray Jackendoff
3. *Recent Transformational Studies in European Languages*
   S. Jay Keyser, editor
4. *Studies in Abstract Phonology*
   Edmund Gussmann

X̄ Syntax: A Study of Phrase Structure

Ray Jackendoff

The MIT Press, Cambridge, Massachusetts,
and London, England

Fifth printing, 1991

Printed and bound in the United States of America

Library of Congress catalog card number 81-80193

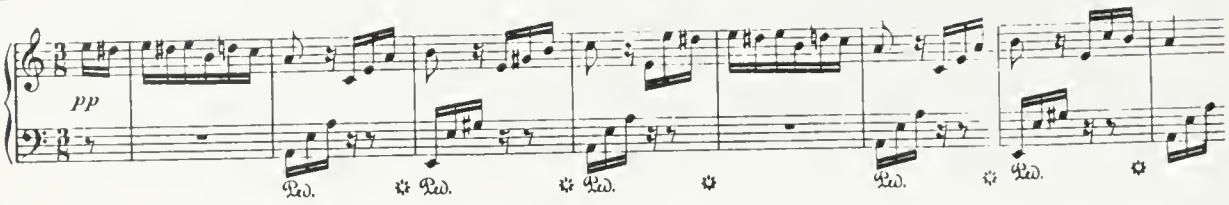
ISBN 0-262-60009-9

Foreword

We are pleased to present this monograph as the second in the series *Linguistic Inquiry Monographs*. These monographs will present new and original research beyond the scope of the article. Because of their originality it is hoped they will benefit our field by bringing to it perspectives that will stimulate further research and insight.

Samuel Jay Keyser

For Elise



and for Amy, who came in time for the page proofs.

Contents

Preface

In the fall of 1967, Noam Chomsky gave the series of lectures at MIT in which he introduced the Lexicalist Hypothesis, one of the first suggestions that the expressive power of a generative grammar is not concentrated primarily in its transformations. Many of us who heard those lectures found the new theory tremendously exciting, and we looked forward to its impact on the study of syntax.

Its immediate effect was more one of spirit than of substance. "Remarks on Nominalization", the written form of the lectures, was the first authoritative stand against the growing school of Generative Semantics; "lexicalist" came to be a term describing anyone who wanted to retain deep structure in transformational grammar, or who opposed transformational derivation of lexical items, or who proposed surface structure rules of semantic interpretation. "Remarks" and Chomsky's next major paper, "Deep Structure, Surface Structure, and Semantic Interpretation", became rallying points for a lexicalist–interpretivist school, mostly consisting of students of Chomsky's from 1965 and later.

But although "Remarks" had a great effect on the approach to syntactic description, the most important issues it raised received surprisingly little attention for several years. Chomsky's arguments against the transformational derivation of lexical items and for the cross-categorial application of transformations were widely appreciated and extended. But his proposals on the organization of the lexicon and the theory of phrase structure, although intriguing, were sketchy and programmatic, and this seems to have discouraged further investigation. Many of us, of course, had inherited a misplaced bias against phrase structure from the early work in generative grammar, which had little of interest to say about phrase structure grammar, stressing only its insufficiency for linguistic description; this prejudice too stood in the way of serious research. More glamorous issues were demanding attention as well: in the late sixties, the relation of syntax to semantics was the main arena; in the seventies, constraints on the transformational component.

In the past few years, though, the structure of the lexicon has come to be of growing concern, as phonology, syntax, and semantics have all put pressure on it to solve their problems. And the gradual reduction in the power of transformations has led syntactic theory, like other aspects of our lives these days, towards simpler ways—in this case back to phrase structure. Thus the scope of Chomsky's original proposals is only just now beginning to be appreciated.

My original conception of this monograph was as a general overview of the

Lexicalist Hypothesis and of its implications for all components of the grammar. But as I got involved in working it out, the importance of phrase structure stood out more and more strongly, until it became clear to me that I would have to devote the entire book to it. And still it kept growing, until the original four or five chapters became the present ten, taxing my patience to a degree I hope I have spared the reader.

In its present form, the first two chapters of the book introduce the motivations for a richer account of the base. Chapter 3 is the conceptual core of the book: it elaborates Chomsky's sketches in "Remarks" into a highly constrained and explicit theory of constituent structure. The remainder of the book justifies the theory and works out its consequences through a detailed examination of English phrase structure, giving a much more wide-ranging and systematic survey of this aspect of the language than has to my knowledge appeared in recent years. Thus the book is conceived of as a contribution both to the study of English syntax and, because of the implications of English for universal grammar, to linguistic theory as a whole.

As is usual for a book that has been so long in preparation, parts of it have appeared in various forms elsewhere. The most rudimentary version was a paper called "Speculations on Presentences and Determiners", published by Indiana University Linguistics Club and now mercifully out of print. "Introduction to the $\bar{X}$ Convention", also published by IULC, was the real germ of the book; readers familiar with that paper will find passages from it scattered throughout chapters 2, 3, and 4, though they will notice that I have since adopted three levels of bars or primes in major phrasal categories rather than two. A paper published in *Language*, "Morphological and Semantic Regularities in the Lexicon", was originally intended as a chapter of this book, but it was crowded out by the insidiously encroaching demands of the phrase structure component. Finally, a paper entitled "Constraints on Phrase Structure Rules", consisting mostly of excerpts from chapters 3, 4, and 9, is appearing in *Studies in Formal Syntax*, edited by Akmajian, Culicover, and Wasow (1977).

This monograph has benefited from the advice of numerous colleagues and students over the years. My discussions on this material with Adrian Akmajian, John Bowers, Joan Bresnan, Noam Chomsky, Peter Culicover, Joe Emonds, Jay Keyser, Joan Maling, Ann Reed, Lisa Selkirk, and Henk van Riemsdijk have been particularly fruitful, and I am grateful for their suggestions and encouragement. I especially want to thank Akmajian and Culicover, in whose homes I precipitously arrived as a houseguest in the summer of 1975 when chapters 6 and 7 were at white heat, for their patience and forbearance in the face of my obsessive enthusiasm.

I also need to express my gratitude to my wife Elise, who, though she arrived on the scene after the book was already well along, quickly came to an understanding of what I found fascinating in this stuff and found herself beginning to care that I cared about it. I have an overwhelming amount to thank her for, perhaps not least of which is her sobering comment on the whole enterprise: "Impressive, but I'd rather read *Newsweek*. . . ."

Brandeis University, August 1977

$\bar{X}$ Syntax: A Study of Phrase Structure

1: Preliminaries

1.1. Goals

Like all work in generative grammar, this study is concerned with describing what it is that a human being knows when he knows how to speak a language. In particular, we take it to be essential to separate those aspects of a speaker's knowledge that are universal from those that are language-particular, under the hypothesis that the language-particular features must be learned by the speaker, but the universal parts may be innate, i.e. determined by the structure of the human organism itself. Success in separating universal from language-particular components enables us to make interesting claims about the nature of the mind.

Given the fact that children learn languages in a relatively short period of time and from rather fragmentary data, we would prefer a theory of linguistic structure in which the job of the language learner is as easy as possible, that is, in which the universal component is maximized. The extreme case, of course, would be if the entire language were innate; but this hypothesis is clearly false, since children learn different languages depending on the linguistic environment in which they are raised. Therefore, the theory must strike the proper empirical balance between the diversity of human languages and the need for them to be learnable.

This vision of the goals of linguistic theory originates in the work of Chomsky; it is explicated in great detail in many of his writings, particularly *Aspects of the Theory of Syntax*. What distinguishes the school of generative-transformational grammar from other extant linguistic theories is the emphasis given to these goals at every level of linguistic description, from the high-level formulation of general theory to the most intricate low-level details of particular languages. The present study, it is hoped, is no exception.

The particular area of inquiry we will address here is the theory of phrase structure. Traditional grammar is concerned with defining parts of speech and constituent structure, with such concepts as the grammatical relations subject and object, and with the relations between modifiers and heads. The present study is an attempt to deal with these traditional issues in the framework provided by the goals and methodology of generative-transformational grammar.

The major questions we will be attempting to answer are these: What are the lexical categories (parts of speech) available to linguistic theory? How are they associated with syntactic structures, and what is the relationship between lexical categories and the syntactic constructions in which they take part? What relationships exist among lexical categories and how can they be expressed? What relationships exist among larger syntactic categories and how can they be expressed?

These questions will be dealt with in the context of an attempt to describe the grammar of English. Of course, in some respects, this choice of data severely limits the range of possible linguistic phenomena to be used as input in formulating the theory. In taking this approach rather than a relatively superficial survey of the grammars of many languages, I am presupposing two methodological biases: first, that it is only by investigating a system very thoroughly that one can begin to understand it; second, that evidence from a single language can in fact provide substantial evidence for linguistic universals. These biases are of course not empirical claims, but only strategies for arriving at empirical claims. No matter how a claim is arrived at, it must of course be tested against a great many details of a great many languages.

1.2. Assumptions about the Structure of Linguistic Theory

By the time of *Aspects of the Theory of Syntax* (1965), it was generally agreed that linguistic theory must provide five distinct components, each with its own characteristic types of rules and characteristic interactions with the other components. In other words, the organization of the grammar into components, the rule types, and their interactions are not something that the child must learn in learning a particular language, but rather are part of his hypothesis about what he is to learn. The field has seen many disputes as to exactly the nature of these components, what aspects of the language each accounts for, how the rules are formalized and constrained, and how the rules interact within components and across components; the settling of these questions is an empirical matter. But the general separation of the theory into components will be taken for granted here.

The five components of the grammar are the *lexicon*, the *categorial* or *phrase structure* component, the *transformational* component, the *phonological* component, and the *semantic* component. The first two are often considered subcomponents of a larger component called the *base* component; the first three, or the second and the third, are often lumped together as the *syntactic* component. Because of the concern of this study with the nature of phrase structure, we will be dealing primarily with the form of the categorial component. We will make no attempt to discuss the other components at all exhaustively. But since any decision about the form of the base is nearly always intimately tied up with the other components, particularly the transformations and the semantics, we will feel no obligation to ignore them.

The lexicon is the repository of idiosyncratic information about individual formatives of the language. Each formative is assigned a lexical entry which describes its

phonological, semantic, and syntactic properties. The lexicon also contains a class of rules called lexical redundancy rules, which may include word formation rules, morpheme structure rules, and rules of allomorphy.

The categorial component includes a set of context-free phrase structure rules, which define the hierarchical syntactic categories of the language and simultaneously determine the left-to-right order of the immediate constituents of each category. The phrase structure rules generate a set of phrase markers whose initial symbol is *S* and whose terminal symbols are the lexical categories of the language.

The base also includes a lexical insertion rule which substitutes lexical formatives for terminal symbols in the phrase markers generated by the categorial component. In the *Aspects* theory, lexical insertion is constrained by the strict subcategorization and selectional restrictions of the heads of phrases. We will assume here a less restricted lexical insertion rule which checks only the lexical category and the strict subcategorization restrictions; selectional restrictions are regarded as well-formedness conditions on semantic interpretations, as argued in Jackendoff (1972, chapter 1). We will also assume that terminal symbols can be left unfilled by lexical items, resulting in an "empty node" in the sense of Emonds (1970); and that a terminal symbol may dominate a phonologically null item PRO, used as an anaphoric element in elliptical constructions.[1] An empty node must be filled at some point in the derivation in order for the derivation to be well-formed; PRO must receive a reading by the rules of anaphora or other specific rules of interpretation, in order for the semantic representation to be well-formed. The phrase markers with lexical formatives generated by the base component of a language *L* constitute the possible *deep structures* of sentences of *L*.[2]

The deep structures are taken as input to the transformational component, which relates them to the surface structures of *L*, and thence via the phonological component to the phonetic representations of *L*.[3] We will assume here that the transformations of *L* are extrinsically ordered and that their application is governed by the principle of the transformational cycle, as follows: certain syntactic categories are designated by linguistic theory as *cyclic nodes*. In the Standard Theory of *Aspects*, the cyclic nodes are the S nodes; it is now generally acknowledged that at least NPs are also cyclic nodes. In applying the transformations to a sentence, apply in order $T_1, \ldots, T_n$ first to the structures dominated by the most deeply embedded cyclic nodes, that is, to those

[1] PRO takes over the function of the empty node Δ used in Jackendoff (1972), which did not have the same meaning as Emonds's empty node. However, Wasow (1975) argues that they need not be distinct entities.

[2] It should be clear from these assumptions that I will not be concerned with theories of transformational grammar which use unordered base strings or posttransformational lexical insertion. Nor will I concern myself here with providing explicit arguments against such theories. However, the analysis arrived at here constitutes an implicit challenge to their proponents, in that it deals with a range of facts and issues generally ignored in discussion of the significance of deep structure in generative grammar.

[3] There has been considerable evidence that earlier levels of syntactic derivation also serve as input to the phonology. See Bresnan (1971), for example.

which do not dominate any other cyclic node. After $T_1, \ldots, T_n$ have applied to these domains, they apply again in order to the structures exhaustively dominated by cyclic nodes, all cyclic nodes dominated by which have already governed an application of the cycle. The cycle is repeated under the same conditions until the highest S node is reached.

We will assume furthermore the principle of the *strict cycle* (Chomsky (1973, 243)): "No rule can apply to a domain dominated by a cyclic node *A* in such a way as to affect solely a proper subdomain of *A* dominated by a node *B* which is also a cyclic node." Without this convention, it would be easy to circumvent the ordering of transformations within embedded contexts.

We will take it for granted that there is a class of transformations which take place on the final (highest) cycle only. In the best possible case, the last-cyclic transformations have the same conventions on application as the cyclic transformations, i.e. they apply in the highest clause only and cannot affect only a subordinate cyclic domain. However, last-cyclic transformations which violate this condition have appeared in the literature, and I will leave the issue open.

It has often been suggested that there is also a class of precyclic transformations, which naturally must have different conventions of application than the cyclic rules. However, no convincing arguments for such rules have appeared within the particular variant of transformational grammar being assumed here; clearly the theory would be more highly constrained if they were unnecessary.

In the Standard Theory of *Aspects*, the semantic component is taken to be an interpretive component which derives semantic representations from the deep structures of *L*. However, if taken literally, this assumption turns out to lead to the Generative Semantics position that deep structures and semantic representations are isomorphic, a position advocated by Lakoff (1971a), McCawley (1971), and others. Since in this theory the form of underlying phrase markers must reflect purely semantic considerations, all the aspects of syntactic generality in a language expressible by the phrase-structure rules of the Standard Theory are lost.

Alternatively, one can give up the assumption that deep structure is the only syntactic source of semantic information, adopting instead what Chomsky has called the "Extended Standard Theory". Arguments for this modification to the Standard Theory appear in Chomsky (1970b) and Jackendoff (1972), among others, and I will not go into those arguments here. In the Extended Standard Theory, various aspects of the interpretation of a sentence are derived from its deep structure, from its surface structure, and possibly from its end-of-cycle derived structures. Although this theory, with its variety of rule types and rather complex organization, is less immediately attractive than the Standard Theory or Generative Semantics, it appears to be rather tightly constrained, in that the formal universals organizing the grammar are considerable.

A recent alternative to the Extended Standard Theory has been provided by the

"trace theory" of movement rules, developed in Chomsky (1973; 1975), Wasow (1971), and Fiengo (1974). In this theory, all movements of constituents leave behind anaphoric "traces" which serve to mark the previous position of the moved constituent. Certain very general constraints on movement rules follow immediately from independently motivated conditions on anaphora which are imposed on the traces. While the consequences of this theory for the transformational component are still not well understood, it has been pointed out by John Goldsmith that the trace theory permits all semantic interpretation to take place at the level of surface structure. Jackendoff (1975b) shows that this property permits a number of disparities in the semantic component of the Extended Standard Theory to be resolved. Furthermore, it restores to the semantic component the conceptual simplicity of the Standard Theory. We will leave open which version of the Extended Standard Theory is to be preferred, as it is largely irrelevant to our concern with phrase structure.

1.3. Lexicalism and Interpretivism

Chomsky's paper "Remarks on Nominalization" (1970a, dating originally from 1967) contrasts the (at that time) usual transformational derivation of nominals such as *the derivation of nominals* from sentences such as *Someone derives nominals* with the "Lexicalist Hypothesis", in which nominals are generated directly by the base component. Among the consequences of the Lexicalist Hypothesis is the claim that the theory of grammar must include a way to refer to more than one syntactic category, using a single term of the structural description of a grammatical rule. Thus the Lexicalist Hypothesis not only changes the nature of the relations among formatives such as *derive* and *derivation*, but also has profound consequences for the role of syntactic categories in linguistic theory. Chapter 2 reviews Chomsky's evidence for the Lexicalist Hypothesis and adds further evidence; subsequent chapters explore its consequences for the categorial component of the grammar.

It is perhaps important to emphasize the logical independence of the Extended Standard Theory and the Lexicalist Hypothesis, since these positions have been identified by and large with the same individuals and since both have been set in opposition to Generative Semantics. There is no a priori reason why the existence of surface structure semantic interpretation has any connection with the choice of a lexical or transformational source for derived nominals. One could well imagine a lexicalist version of the Standard Theory or a transformationalist version of the Extended Standard Theory. In practice, however, the two hypotheses turn out to supplement each other, for what I think are two separate reasons.

The first reason is methodological. The proponents of Generative Semantics have attempted to constrain the theory of grammar by restricting the number of rule types, categories, and derivational levels; and they are willing to accept as the price of this constraint a high degree of abstractness and complexity in syntactic derivations. Those

who have adopted the lexicalist and surface structure interpretivist positions, however, are willing to accept a wider range of rule types, categories, and rule interactions within the theory of grammar, provided each component can be sufficiently constrained. They often are rather skeptical about the ability of traditional logic to provide a model of any part of semantic representation. This methodology turns out to be congenial both to the Lexicalist Hypothesis and to the Extended Standard Theory of semantic interpretation.

But there are empirically based ties between the two hypotheses as well. For instance, if one wants to maintain that the interpretation of sentences containing negation is determined solely from their deep structure, one must permit transformations which drastically alter the shape of lexical formatives; Klima (1964) posits such changes as *neg + any → no*, *neg + often → seldom*, and *neg + many → few*. Though one could maintain a narrow Lexicalist Hypothesis for nominalizations and still admit such derivations, they certainly violate the spirit of the Lexicalist Hypothesis, in that they account for what are at best semiproductive lexical processes by means of a transformation. Only by adopting a surface structure interpretive theory of negation (e.g. Jackendoff 1969a; 1972, chapter 8) can we maintain a lexicalist theory of negative morphemes. Hence surface structure interpretation is a necessary condition for a desirable extension of the Lexicalist Hypothesis; the two theories thus complement each other empirically as well as methodologically.

2: Motivation for the $\bar{X}$ Convention

This chapter provides the background against which the rest of this study is set. Section 2.1 is concerned with Chomsky's treatment of the relation between verbs and their nominalizations; section 2.2 with adjectives and their nominalizations; and section 2.3 with adjectives and adverbs. Our approach is to show that none of these relationships among lexical items can be accounted for transformationally. The task of expressing these relationships must therefore fall on other components of the grammar.

One kind of relationship which it becomes necessary for the grammar to express is *cross-category generalization*. We recapitulate Chomsky's development of the $\bar{X}$ Convention as a way to incorporate cross-category generalization, and we show how this leads to a theory of the base.

2.1. Chomsky's Arguments

Chomsky's paper "Remarks on Nominalization" has as its most immediate concern the syntax of derived nominals such as *John's criticism of the book* and *John's refusal of the offer*. The point of departure is the considerable number of differences between derived nominals and "gerundive nominals" such as *John's criticizing the book* and *John's refusing the offer*. Basically, gerundive nominals behave like regular transforms of sentences, but derived nominals do not.

Here are some of the differences. Gerundive nominals occur with aspectual verbs, but derived nominals do not.

(2.1) John's having criticized the book
Byrne's having been refusing the offer just when Nixon arrived

(2.2) *John's have(ing) criticism of the book
*John's have(ing) been refusal of the offer

Gerundive nominals use adverbs before the gerund; derived nominals use adjectives before the nominalization.

(2.3) John's sarcastically criticizing the book
John's emphatically refusing the offer

(2.4) John's sarcastic criticism of the book
John's emphatic refusal of the offer

Derived nominals can take part in partitive expressions and can be pluralized; gerundive nominals cannot.

(2.5) *some of John's criticizing the book
*John's many refusings the offer
(2.6) some of John's criticisms of the book
John's many refusals of the offer

But gerundive nominals can take a range of adverbials that are prohibited in derived nominals.

(2.7) John's criticizing the book too often
John's refusing the offer in a suspicious manner
(2.8) *John's criticism of the book too often
*John's refusal of the offer in a suspicious manner

Gerundive nominals, but not derived nominals, can undergo Subject Raising, Dative Shift, and Particle Movement.

(2.9) its being certain that John will win $\overset{\text{Subject Raising}}{\Rightarrow}$
John's being certain to win

John's giving the book to Bill $\overset{\text{Dative Shift}}{\Rightarrow}$
John's giving Bill a book

John's looking up the information $\overset{\text{Particle Movement}}{\Rightarrow}$
John's looking the information up

(2.10) the certainty that John will win $\not\Rightarrow$
*John's certainty to win

John's gift of a book to Bill $\not\Rightarrow$
*John's gift of Bill a book

John's looking up of the information $\not\Rightarrow$
*John's looking of the information up

Derived nominals may use *the* in the absence of a subject, while gerundives may not; compare (2.10) with **the being certain that John will win.*

The properties of gerundive nominals are exactly what one would expect if they were transforms of sentences. But transformations which predict this behavior in

producing derived nominals from sentences will necessarily be rather difficult to state and order.[1]

A second kind of data distinguishing gerundive from derived nominals concerns syntactic productivity and regularity. While all sentences (except those whose subjects are too complex to have a possessive form) have gerundive nominals, many, for example those in (2.11), have no derived nominals.

(2.11) It is easy to please John.
John amused the children with his stories.
It seems that John is gone.

(2.12) its being easy to please John
John's amusing the children with his stories
its seeming that John is gone

(2.13) *the easiness to please John
*John's amusement of the children with his stories
*its seem that John is gone

Conversely, many derived nominals do not correspond precisely to well-formed sentences.

(2.14) *John doubted about their proposal.
*It is probable John's leaving.
*John advised to Bill.

(2.15) *John's doubting about their proposal
*its being probable John's leaving
*John's advising to Bill

(2.16) John's doubts about their proposal
the probability of John's leaving
John's advice to Bill

Transformations which perform all the minor syntactic adjustments necessary to produce (2.16) from well-formed sentences will be numerous and highly idiosyncratic, if they can be stated at all.

A third argument concerns semantic regularity. Derived nominals typically have meanings rather idiosyncratically related to the parallel verb, whereas gerundive nominals have absolutely predictable semantic relationships with the verb. Chomsky mentions examples such as *laughter, marriage, construction, actions, activities, revolution, belief, doubt, conversion, permutation, trial, residence, qualifications, specifications* (p. 19/189),[2] *intelligence, deed,* and *writings* (note 10, p. 19/note 11, p.

[1] In particular, Ross (1967) argues from the existence of a transformation producing derived nominals to the conclusion that Dative Shift and Particle Movement are last cyclic: if they were cyclic, they would have no way of knowing that the sentence they operate in was to be turned into a derived nominal on the next cycle. Jackendoff (1972, chapter 9) argues that this conclusion leads to an ordering paradox, for Dative Shift must take place in cyclic derivations such as *John is believed by Harry to have been given a book by Bob.*

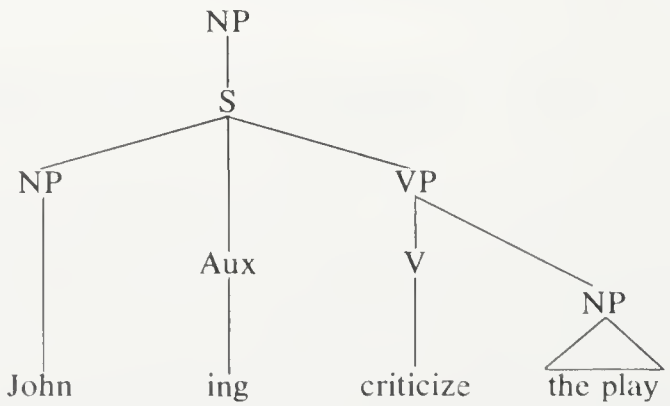
[2] I will give the page references on Chomsky's *Studies on Semantics in Generative Grammar* followed by those in Jacobs and Rosenbaum (1970).

217). Other examples are easy to find. The point of these examples is that a transformational derivation of at least some readings of these words, from sentences involving the verbs *laugh, marry,* etc., must involve the presence of semantic material idiosyncratic to the particular nominalization. For example, *John's deeds* does not mean 'things which John did' but rather at least 'fairly significant things which John did'. The derivation of the political sense of *revolution* from *revolve* is fairly obscure; similarly for the sense of *construction* in *the verb–particle construction in English*. In short, a transformational derivation of these nominals entails an incredible proliferation of idiosyncratic transformations.

Finally, there are a great number of nouns which occur in structures similar to those observed in derived nominals, but for which there is no corresponding verb at all. Chomsky cites constructions like *the weather in England, the story of Bill's exploits, the message from Bill to Tom about the meeting, a war of aggression against France, John's attitude of defiance towards Bill, his habit of interrupting* (p. 30/196), *the author of the book, the assistant vice-chancellor of the University* (p. 32/196), and *John's poems/sonnets/Alexandrines* (*about Homer*) (note 10, p. 20/note 11, p. 217). A consistent transformationalist approach must derive these nouns from verbs too. Following such an approach, Lakoff (1971b) comes to derive *king* from a verb by analogy with *ruler*; he then invokes the notion of "positive absolute exception" to the nominalization rule, in order to guarantee that the verb underlying *king* never appears at the surface. Such a solution is indicative of the unnaturalness of the transformational theory.

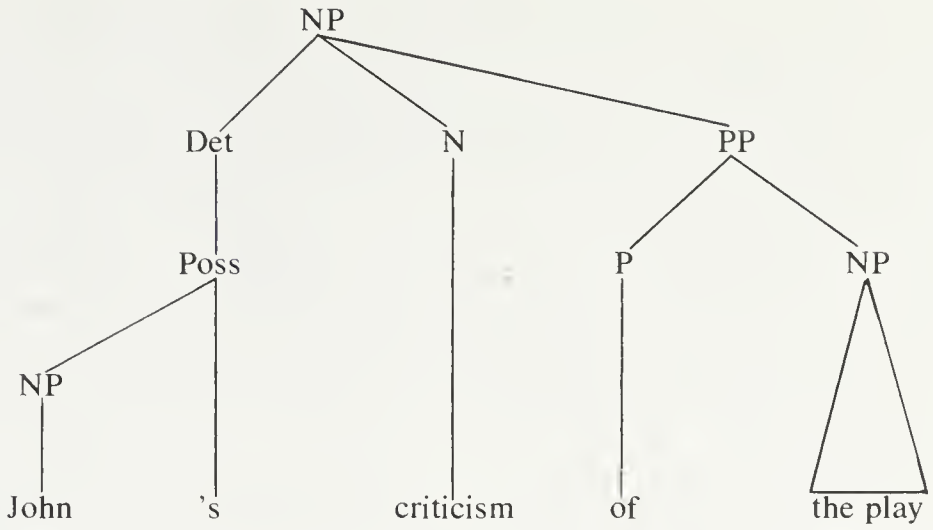
To describe the differences between gerundive and derived nominals in a systematic way, Chomsky claims that gerundives have an underlying structure like (2.17), containing an S,[3] but that derived nominals have an underlying structure like (2.18), in which the head is a noun and no S node appears.

(2.17)



[3] We will propose a somewhat different structure in section 3.6, but for the moment we adopt Chomsky's formulation.

(2.18)



In order to propose a structure like (2.18), in which *criticism* is a deep-structure noun, we must provide a mechanism in the lexicon to show the relation between *criticism* and the verb *criticize*. Chomsky proposes that *criticize* and *criticism* form a single lexical entry, unmarked for the syntactic feature differentiating nouns from verbs. When inserted under a verb node, this entry will direct the use of the phonological form *criticize*; when inserted under a noun node, it will be realized as *criticism*. In an alternative formalism argued for in Jackendoff (1975b), *criticize* and *criticism* are separate lexical entries whose relationship is specified by a lexical redundancy rule.

Whichever formalism one chooses, there is an important innovation in this theory: the relations among surface lexical formatives can be expressed by other than transformational means. As Chomsky points out (p. 17/188), early work on transformational grammar had no other means but transformations to relate lexical items; hence the transformational position on nominalizations was the only possible theory. But with the introduction of a lexicon and syntactic features into linguistic theory, as in *Aspects*, a lexicalist position could be formulated, in which these regularities were expressed in the lexicon. As the evidence seems to make the transformationalist position appear problematic, it behooves us to explore the consequences of the lexicalist position.

As a first consequence, we arrive at immediate explanations for the syntactic differences between gerundive and derived nominals pointed out above. Gerundives, being sentences in deep structure, are expected to contain aspect; since *-ing* substitutes for Tense and Modal, their absence is accounted for. Derived nominals, being NPs, will not contain auxiliary elements. Likewise, it is typical for sentences to exhibit the preverbal adverbs which appear in gerundive nominals, and typical for NPs to exhibit the prenominal adjectives which appear in derived nominals. The verb of a sentence cannot be pluralized or quantified, but the noun of an NP may; hence the contrast between gerundive and derived nominals of (2.5)–(2.6). The time and manner adverbials of gerundives are typical of sentences but not of NPs, so derived nominals do not

permit them. We expect Subject Raising, Dative Shift, and Particle Movement in Ss but not in NPs, hence the differences in (2.9)–(2.10). Only NPs are expected to use the definite article *the*, so only derived nominals can use it.

The differences in productivity and regularity also make more sense within the Lexicalist Hypothesis. If gerundives are simple transforms of sentences, we correctly expect unfailing productivity of forms and a systematic semantic relationship to sentential forms. On the other hand, lexical relationships are notoriously irregular and idiosyncratic, and it is therefore not at all surprising that derived nominals should be as erratic as they are. Furthermore, if derived nominals are nouns, we would expect (as observed) that there are NPs of the same structure as derived nominals, not related in any way to sentences, simply because there happen to be no verbs lexically related to their head nouns.

Thus the Lexicalist Hypothesis makes a principled differentiation between the two types of nominalizations, explaining their differences as an automatic consequence of their structures. Given the proposed structures, the properties of the two types of nominalizations could not be too much different than they are without complicating the grammar. On the other hand, the Transformationalist Hypothesis must resort to a great deal of descriptive complexity to yield the observed properties of derived nominals. These properties are thus taken to be accidental, in that there is no logical connection between them, and in that the grammar would be much simpler if in fact the differences between gerundive and derived nominals did not exist. Stated differently, in the Transformationalist Hypothesis, the language learner must learn a large number of related transformations to produce derived nominals; but in the Lexicalist Hypothesis, the language learner needs to learn only the differences in deep structures (2.17) and (2.18) in order to produce all the rest of the differences from independently learned rules.

To make this argument fully convincing, we must of course show that all the rules applying to derived nominals in the lexicalist account are independently motivated. To begin with, examine the alleged base form (2.18). It has often been claimed that the possessive form in *John's house*, for example, is derived by reduction from a deep structure relative clause construction such as *the house which John has*. Under this theory, possessive NPs do not appear in the base. However, (2.18) claims that the NP *John* occupies prenominal position in the base form, so the base must apparently be enriched to accommodate this structure. Chomsky shows, however, (pp. 37–38/200–201) that this enrichment of the base is independently necessary: the "inalienable" sense of *John's leg*, in which the leg is part of John's body rather than a wooden leg he happens to own, cannot be paraphrased by *the leg that John has*. For example, we cannot say *The leg that John has hurts* meaning *John's leg hurts*. Further, expressions such as *John's forte* have only an inalienable sense; **the forte which John has* is nonsensical. This suggests that at least the inalienable sense of possession is not derived from a reduced relative clause, but rather is generated in the base.

Further evidence comes from phrases like *John's book* and *John's proof of the*

theorem. Here appropriate relative clause paraphrases are *the book which John wrote* and *the proof of the theorem that John discovered*. Insofar as the relative clause describing John's relationship to the object is dependent on the nature of the object and is often still open to variation, a transformation reducing the relative clause to a possessive will violate the recoverability requirement on deletion transformations. Suppose, however, that the possessives are generated in the base. One projection rule for the possessive position can specify a rather loose notion of "intrinsic connection" between the possessive NP and the object denoted by the larger NP. This notion would be sharpened by the semantic nature of the larger NP: if it is written material, intrinsic connection denotes the writer; if it is an idea, intrinsic connection denotes the discoverer; and so forth. Placing the burden of specifying intrinsic connection on the semantic component (or preferably on real-world knowledge) instead of on the syntax eliminates the violation of recoverability. It simultaneously explains the creativity in the use of intrinsic connection: for example, *John's chair* may denote the chair that John owns (alienable possession), or, by intrinsic connection, the chair that John built, designed, or habitually sat in.[4] Of course, a complete account must be much clearer about the nature of intrinsic connection that I can be here. My conjecture, however, is that it is a rather fundamental cognitive function, of which the examples above are simply special cases.

Thus there is some justification for including the possessive position in the base, independently of its use in derived nominals. What about the prepositional phrase following the head noun in (2.18)? Again, it is often claimed that PPs in this position are reduced relative clauses, and so this position is unnecessary in the base. For example, *a book about Harry* and *a picture of the lizard* are easily derived from *a book which is about Harry* and *a picture which is of the lizard*. Furthermore, the reduction rule has good independent motivation, since it provides a straightforward source for such phrases as *the child sleeping in the alley* and *a book yellow with age*. Chomsky points out, however, that at least some PPs in NPs cannot be derived from plausible relative clauses, for example *the weather in England, the weather in 1965, the author of the book, his attitude of defiance, his advantage over his rivals, the reason for his refusal,* and *his habit of interrupting*. Since there are no constructions such as **the weather which is in England,* for example, it seems plausible simply to expand the base rules to permit certain PPs to be generated under NP. But this extension is precisely what we need to generate (2.18) as a base form.

As further justification for this expansion of the base rules, we observe that many adjective phrases can include prepositional phrases or complement clauses which cannot be reduced relative clauses, for example *afraid of Bill, eager to please, pleased with his work, apprehensive about the job,* and *sorry to be of trouble*. This suggests that

[4] Cf. "Remarks", pp. 45/205–206. One might with good reason try to interpret possession as a special type of intrinsic connection: there seems little justification to syntactically distinguish *own* or *have* from *build* or *write* purely for the purposes of stating a reduction transformation, when a suitable semantic generalization might eliminate the reduction rule altogether.

all three categories N, V, and A permit a range of following phrases. We will show in chapter 4 that prepositions likewise exhibit a range of associated structures.

To express this generalization, Chomsky introduces the variable X to stand for any lexical category symbol, and uses the notation $\bar{X}$ (henceforth here X′) to denote the node immediately dominating X. The general base rule schema which expresses the existence of material to the right of the head of a phrase is (2.19).

(2.19) $X' \rightarrow X - Comp$

Comp (complement) is an abbreviation for some sequence of nodes: since it never seems to be referred to as a constituent, it does not stand for a node. (2.19) is an abbreviation for the four rules (2.20).[5]

(2.20) $V' \rightarrow V - Comp_V$
$N' \rightarrow N - Comp_N$
$A' \rightarrow A - Comp_A$
$P' \rightarrow P - Comp_P$

Chomsky further suggests that the material preceding the head of a major category phrase is introduced by a base rule of the form (2.21). $\bar{\bar{X}}$ (henceforth here X″) designates the node which is two nodes above a major category node, and [Spec, X′] (the specifier of X′, henceforth noted as $Spec_X$) is an abbreviation for the material preceding the head.

(2.21) $X'' \rightarrow Spec_X - X'$

(2.21) is an abbreviation of the rules shown in (2.22):

(2.22) $V'' \rightarrow Spec_V - V'$
$N'' \rightarrow Spec_N - N'$
$A'' \rightarrow Spec_A - A'$
$P'' \rightarrow Spec_P - P'$

Again, the material in the specifier varies from category to category. Chomsky associates $Spec_V$ with the traditional auxiliary node, $Spec_N$ with the traditional determiner, and $Spec_A$ with traditional degree phrases. We will discuss $Spec_P$ in chapters 4 and 6. Chomsky does not say explicitly whether he considers the specifier to be a constituent or, like Comp, an abbreviation for a sequence of constituents. His diagrams show it as a constituent; however, we will argue here that it should not be.

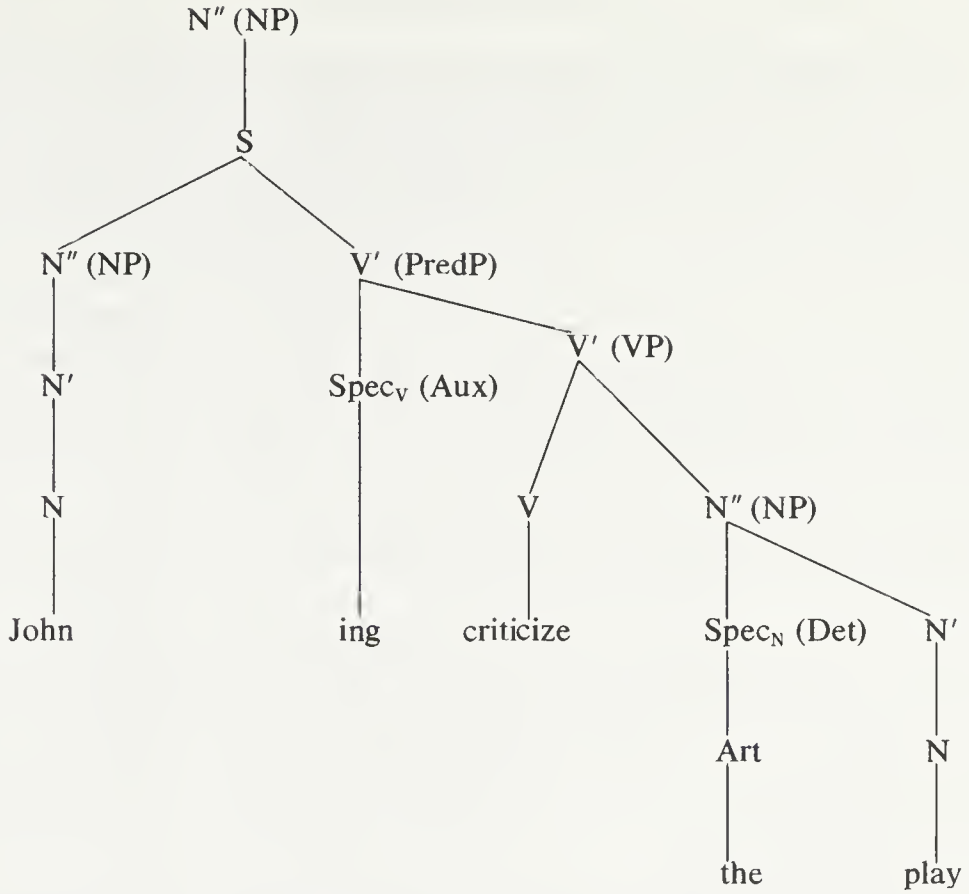
To complete the base rules, Chomsky uses the initial rule (2.23).

(2.23) $S \rightarrow N'' - V''$

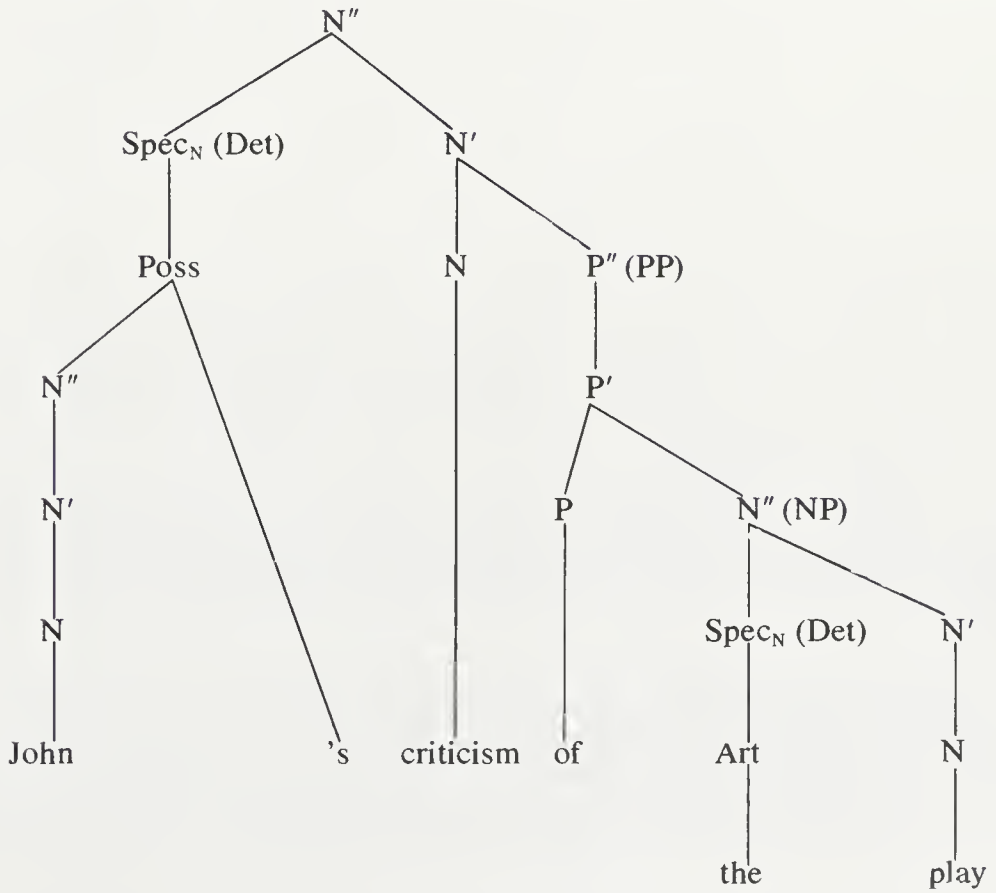
Thus the trees (2.17) and (2.18) can be rewritten as (2.24) and (2.25). I enclose the traditional names of the nodes in parentheses.

[5] Chomsky does not subscript the various Comps by category. However, since the the Comp varies from one category to another, I have included the subscript to indicate the possibility of variation.

(2.24)


N″ (NP)
S
N″ (NP)
V′ (PredP)
V′ (VP)
N′
Spec$_V$ (Aux)
N
V
N″ (NP)
John
ing
criticize
Spec$_N$ (Det)
N′
Art
N
the
play

(2.25)


N″
Spec$_N$ (Det)
N′
Poss
N
P″ (PP)
N″
P′
N′
P
N″ (NP)
N
Spec$_N$ (Det)
N′
John
's
criticism
of
Art
N
the
play

Chapter 3 suggests some modifications to Chomsky's translation of the traditional nodes into the $\bar{X}$ notation. However, for the moment we will assume his analysis.

Return to the base structure (2.18), now further analyzed as (2.25), for the derived nominal *John's criticism of the play*. One of the stronger arguments for deriving this nominal from a sentence is that the selectional restrictions parallel those of the sentence *John criticized the play*: the range of noun phrases that can occupy possessive position in the derived nominal is identical with the range of possible subjects in the sentence, and the range of noun phrases following *of* is identical with the range of direct objects in the sentence. In order to capture this generalization for a wide range of nominals, we do not want to state twice in the grammar the selectional restrictions shared by *criticize* and *criticism*; nor do we want to state twice the set of rules which enforce these selectional restrictions, once for NP and once for S. One way to avoid stating the selectional restrictions and rules twice is to derive the nominal from the sentence, for then the selectional restrictions will only be stated on the verb, and the rules for enforcing the restrictions will be stated only for the domain S.

However, Chomsky's use of a more complex lexicon, including related lexical items, eliminates the problem of stating selectional restrictions twice. The lexicon will specify, as part of the relation between *criticize* and *criticism*, the similarity in selectional restrictions. So even in the lexicalist approach the selectional restrictions need only be stated once.

As for the rules which enforce selectional restrictions, we first observe that such rules must exist in noun phrases as well as sentences, since selectional restrictions for *weather* and *attitude* must prohibit, for example, **the weather in our idea* and **my shirt's attitude of hunger*. Furthermore, Chomsky suggests that along with the generalization of base rules as in (2.19), there is a corresponding generalization in the projection rules. For example, the notion "subject of" can, where semantically appropriate, be generalized to possessive NPs, and "object of" can be generalized to postnominal NPs. This generalization ensures that the rules enforcing selectional restrictions on *criticize* and *criticism* are in fact the same rules in Ss and NPs, and they need not be stated twice. Thus, if this generalization can be expressed, and if the lexical relations can be stated successfully, the Lexicalist Hypothesis can relate the selectional restrictions of derived nominals to those of sentences without recourse to a transformational derivation.

A second possible difficulty for the Lexicalist Hypothesis concerns the Passive transformation. We observe that the relation of the derived nominals (2.26a) and (2.26b) seems to be the same as that between the active sentence (2.27a) and its passive (2.27b).

(2.26) a. the enemy's destruction of the city
b. the city's destruction by the enemy

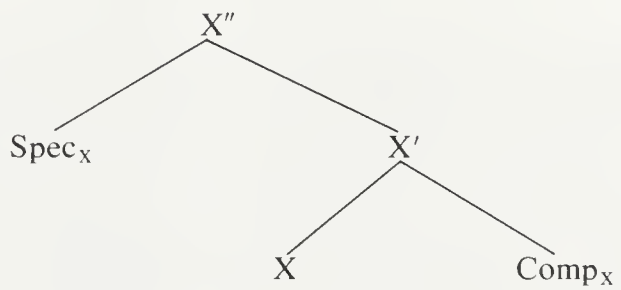
(2.27) a. The enemy destroyed the city.
b. The city was destroyed by the enemy.

In order to capture the similarity of the two relations, the grammar should not have to state, in addition to the standard Passive, a transformation having very similar effects to the Passive but in the domain NP. Within the Transformational Hypothesis this generalization is quite simple; all four phrases above are derived from the same underlying form, but (2.26a) undergoes Nominalization, (2.27b) undergoes Passive, and (2.26b) undergoes Passive followed by Nominalization. (2.26b) thus is the nominalization of a passive sentence.

In the Lexicalist Hypothesis this solution is not available, since there is no Nominalization transformation. Instead Chomsky suggests that, like the base rules and the projection rules, some transformations, in particular the Passive, may apply over the domain NP as well as S. If this is the case, (2.26b) is derived from the same underlying form as (2.26a), but it undergoes the generalized form of the Passive. Hence its relation to (2.26a) is the same as the relation of (2.27b) to (2.27a), precisely as required, and no extra rules need be added to the grammar. Section 4.7 will discuss the generalized Passive in more detail.

The general nature of the claims made by the $\bar{X}$ Convention are now clear. The structural schema (2.28), in which X represents any lexical category, is claimed to constitute a linguistically significant generalization of the structures associated with major categories.

(2.28)



That is, we expect there to exist rules whose structural descriptions refer to a range of structures including more than one value of X. For example, we expect to find rules whose domains include V″ and N″ or perhaps A′ and N′: but we do not expect to find rules whose domains include nodes at different levels, for example P′ and A″.

The fact that some rules generalize to more than one major category of course does not mean that *all* rules generalize over more than one major category. If that were the case, there would be little reason to distinguish among major categories. Rather, we use nongenerality of rules to distinguish one category from another. For example, the rule of Cleft Sentence Formation can be used as a test to distinguish N″ and P″ from A″ and V″.

The existence of cross-category generalizations is therefore not taken to be evidence that one category is a subclass of some other category, as suggested by some, for example, Lakoff (1971b), Ross (1969), and Postal (1971, chapter 18). Such a tactic may merely shift around the names assigned to categories, since the distinctions must

still be made. For example, if we claim that all adjectives are deep structure verbs, we must still assign a syntactic feature ±F which distinguishes surface structure verbs from surface structure adjectives, and which is referred to by all rules in which adjectives and verbs differ. In effect, then, we may have simply changed the traditional name "adjective" into "+F" and the traditional name "verb" into "−F", not a very insightful proposal.

Of course, in a theory of grammar which does not include the $\bar{X}$ notation or some similar device, there is no way to express cross-category generalization directly. Thus one is forced to resort to a supercategory, with all the syntactic difficulties such a proposal entails. Chomsky's claim, then, is that cross-category generalizations play an important role in grammar, and that the theory of grammar thus must contain a way of expressing them.

The traditional method of expressing a classificatory system which includes cross-classification is with a set of distinctive features. For the purpose of stating generalized rules in the $\bar{X}$ notation, chapter 3 will introduce a set of syntactic distinctive features. But first, let us examine other evidence for the Lexicalist Hypothesis, of a similar nature to Chomsky's but involving different categories.

2.2. Adjectives and Nouns

Bowers (1968a), from which much of the material in this section is derived, discusses the relationship between adjectives such as *high, long*, and *deep*, and their nominals *height, length*, and *depth*, comparing the transformationalist and lexicalist positions. The basic fact to be accounted for is that the genitive phrase in nominals like *the building's height, the railroad's length, the lake's depth, the slope's steepness, the table's width, the paper's thickness*, and *the plane's lateness* is restricted in precisely the same way as the subject of the corresponding sentences *The building is high, The railroad is long, The lake is deep, The slope is steep, The table is wide, The paper is thick*, and *The plane is late*. Likewise for nonmeasure adjectives and their nominals: the genitive NPs in *Mary's beauty, John's perversity, the rock's whiteness, the table's flatness, the desert's dryness, the sky's cloudiness, that proposal's absurdity*, and *Sarah's diffidence* are restricted in precisely the same way as the subjects of the sentences *Mary is beautiful, John is perverse, The rock is white, The table is flat, The desert is dry, The sky is cloudy, That proposal is absurd*, and *Sarah is diffident*. Under a transformationalist theory these relations are expressed by a transformation which converts embedded predicative Ss into the corresponding NPs, the predicate adjective becoming a noun. Under the Lexicalist Hypothesis the adjectives and nouns are given separate but related lexical entries; the noun phrases are generated by the base rather than by the transformational component.

For purposes of comparison, we will again contrast the derived nominals with gerundive nominals. Gerundive nominals occur with degree adverbs, but derived

nominals occur with degree adjectives:

(2.29) the building's being extremely high
the table's being relatively flat
(2.30) the building's extreme height
the table's relative flatness

There is a range of degree phrases possible in gerundive nominals that is impossible in derived nominals.

(2.31) the building's being very high
the table's being too flat
(2.32) *the building's very height (wrong reading)
*the table's too flatness

The characteristic equative construction for gerundive nominals is *as . . . as*; for derived nominals *the same . . . as* is sometimes applicable.

(2.33) the building's being as high as that tree
the table's being as flat as the desert
(2.34) The building has the same height as the tree.
?The table has the same flatness as the desert.

Subjectless gerundive nominals cannot use the definite article; subjectless derived nominals may.

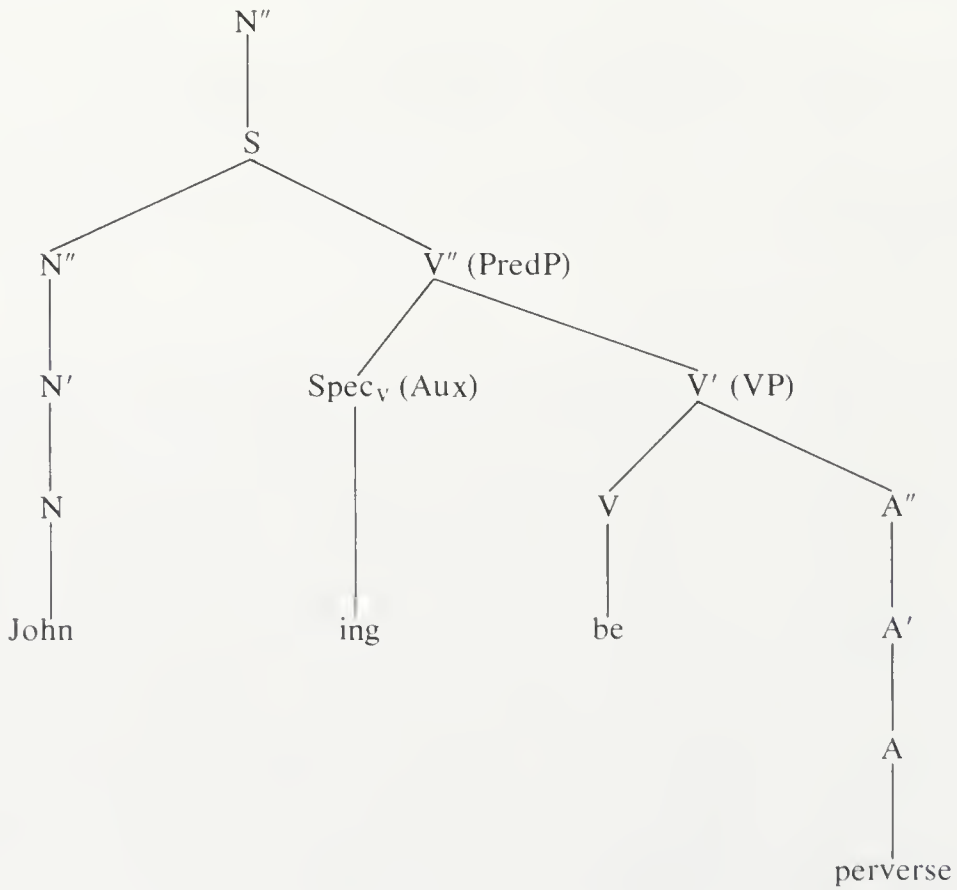
(2.35) *the being high
(2.36) the height

Thus the adjective phrases in adjectival gerundive nominals have all the properties one would expect of adjective phrases; the corresponding derived nominals have the internal structure of noun phrases, taking all the characteristic noun modifiers such as the definite article, *the same*, and prenominal adjectives. Such behavior is consistent with the Transformational Hypothesis, in that one can always state transformations to make all the necessary adjustments; but the existence of these transformations is taken to be an accident of English syntax—things could just as well be otherwise. Under the Lexicalist Hypothesis, however, only the gerundive nominals contain an underlying AP; the derived nominals are deep structure NPs. This theory thus predicts the difference between the two types of nominals and describes the behavior of derived nominals as an automatic consequence of the behavior of ordinary NPs. Since ordinary lexical properties determine syntactic behavior, the Lexicalist Hypothesis says that things would be less general if they were otherwise, for instance if (2.31) instead of (2.32) were ungrammatical and everything else were the same.

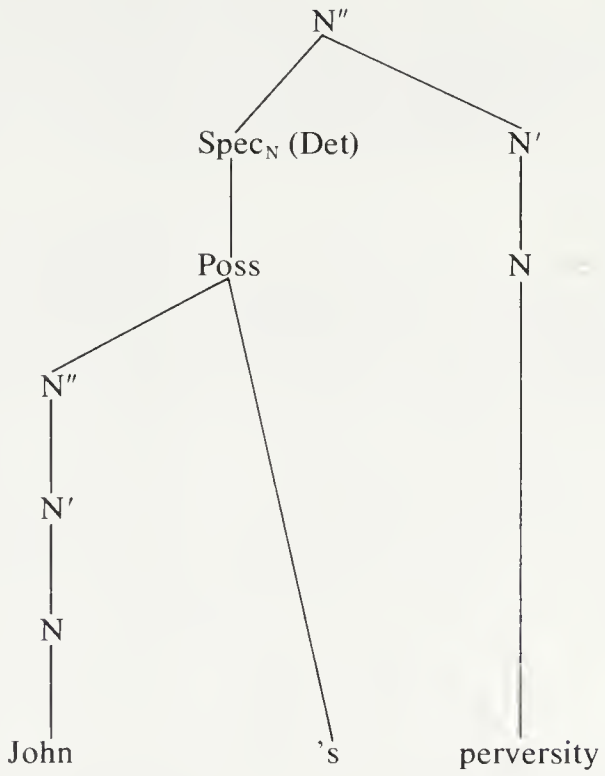
As with the deverbal nominals, there are irregularities in the relationship between adjectives and nouns. *Height* does double duty for *tall* and *high*, there being no **tallness*. In addition, the nouns *elevation* and *altitude* cover part of the same semantic area but are unrelated to adjectives. **Fastness* is supplanted by *speed* (which is in turn related to an adjective *speedy*, which is in turn related to a noun *speediness*), **oldness* is supplanted by *age*, and **afraidness* is supplanted by *fear* (which is related to an adjective *fearful*). *Ability* covers a much wider semantic range than its adjective root *able*; the semantic relation between *clean* and *cleanliness* is irregular. Adjectives which take infinitive complements (*She is pretty to look at; It is easy to please Bill*) do not have direct nominal counterparts **her prettiness to look at, *the ease/easiness to please Bill,* though the latter type is related to a gerundive complement construction *the ease/easiness of pleasing Bill.* In a transformational theory, each of these irregularities must be derived by a lexically specific transformation or constraint. In the Lexicalist Hypothesis, they simply represent deviations from the most ideal relationship; irregular pairs contain more independent information than regular ones, but do not otherwise complicate the grammar.

Under the Lexicalist Hypothesis, gerundive nominals and derived nominals will have structures (2.37) and (2.38), respectively.

(2.37)



(2.38)



As with the nominals of section 2.1, we must account for the generality of selectional restrictions, e.g. for the fact that *perverse* restricts the subject in (2.37) in precisely the same way that *perversity* restricts the genitive phrase in (2.38). The transformationalist theory of course explains this relationship as an automatic consequence of the deep structure ascribed to (2.38), which, like (2.37), will contain the S *John be perverse*.

In the lexicalist framework, we must again appeal to the use of redundancy rules in the lexicon to express the similarity of the selectional restrictions of *perverse* and *perversity*. As for the rules which enforce selectional restrictions, we can appeal to the extension to NPs of the notion "subject-of" proposed in section 2.1. Since the genitive NP in (2.38) is to be considered the subject of (2.38), the rules which impose selectional restrictions on grammatical subjects will apply equally in (2.37) and (2.38). Thus, as in section 2.1, the Lexicalist Hypothesis requires a cross-category generalization to be expressed, but this time in terms of adjectives and nouns.

Bowers points out that there is a similarity in the semantic functions of various subordinate clauses in APs and NPs.

(2.39) a. The table is so wide that the cloth won't cover it.
b. The table is as wide as I thought it was.
c. The table is wider than I thought it was.
d. The table is too wide to be useful.

(2.40) a. The table $\left\{\begin{matrix}\text{has}\\\text{is}\end{matrix}\right\}$ such width that the cloth won't cover it.

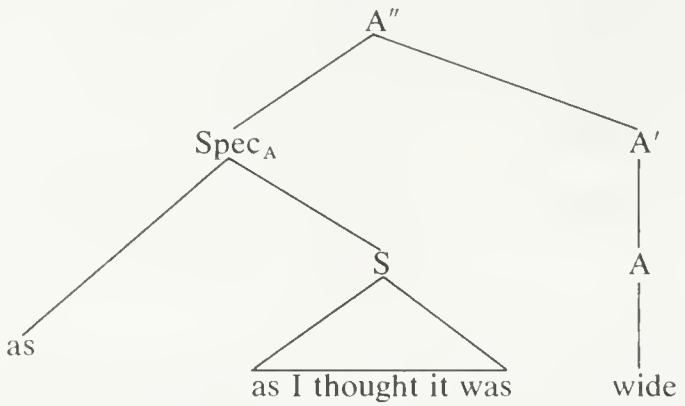
b. The table has the width I thought it had.

c. The table has more width than I thought it had.

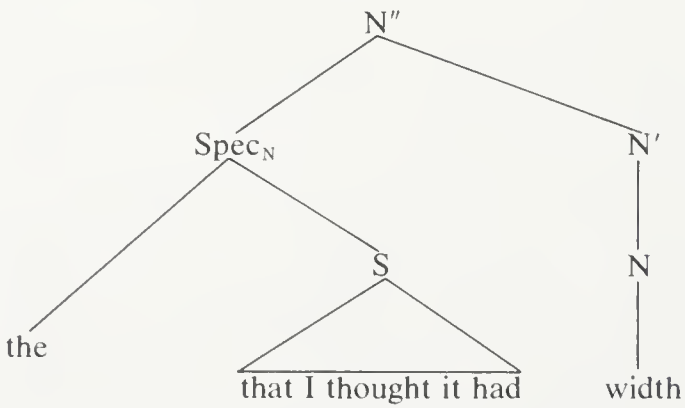
d. The table $\left\{\begin{matrix}\text{doesn't have the}\\\text{isn't of a}\end{matrix}\right\}$ width to be useful.

While there is not a clear one-to-one correspondence between these phrases, there is enough similarity to require explanation. In particular the relation between the equative (2.39b), a comparative clause, and (2.40b), a relative clause, is striking. Bowers suggests a similarity in the structures of comparative and relatives roughly like (2.41) and (2.42).

(2.41)



(2.42)



An extension of the grammatical relation obtaining between a noun and its relative clause to that between an adjective and its comparative clause is called for. The Lexicalist Hypothesis by now leads us to expect such extensions. We will deal with this particular relation in chapter 8.

Throughout this section we have ignored an alternative form for the derived nominals: *the height of the building, the width of the table,* etc. Bowers argues that this form is transformationally derived, and that the rule deriving it from the form with a genitive NP accounts for some of the difficulties in Chomsky's treatment of the passive. We will deal with this rule in connection with the passive, in section 4.7.

In sum, then, adjectives and their nominalizations present a picture quite similar to verbs and their nominalizations. The same difficulties accrue to the Transformationalist Hypothesis, and the same problems of lexical relations and cross-category generalizations must be solved by the Lexicalist Hypothesis.

2.3. Adjectives and Adverbs

The similarities between adjectives and adverbs in English are far more striking than those between other pairs of categories we have considered, and many people have considered it self-evident that adverbs should be derived transformationally from adjectives. Nevertheless, we will show here that a theory which does not distinguish between the two categories in deep structure leaves something to be desired.

First observe why a transformational theory is so tempting. Nearly all adverbs consist morphologically of an adjective plus the *-ly* suffix, and the meanings of adjective–adverb pairs are usually quite close. Furthermore, the modifier systems of the two categories are identical, as pointed out by Bowers (1968b), among others.

(2.43) a. John's car is $\left\{\begin{matrix}\text{rather}\\ \text{incredibly}\\ \text{very}\end{matrix}\right\}$ fast.

b. Its acceleration is so rapid that you'll never catch him.

c. That quartet is more beautiful than you can imagine.

d. The stop was too sudden to react to.

e. The sun was not as bright as the sunlamp.

(2.44) a. John's car can travel $\left\{\begin{matrix}\text{rather}\\ \text{incredibly}\\ \text{very}\end{matrix}\right\}$ fast.

b. It can accelerate so rapidly that you'll never catch him.

c. They played the quartet more beautifully than you can imagine.

d. It stopped too suddenly for us to react.

e. The sun didn't shine as brightly as the sunlamp.

But deriving adverbs from adjectives in any semantically enlightening way turns out to be a process conspicuously lacking in generality: observe the variety of sources needed (examples from Jackendoff (1972, 52–55)).

(2.45) John is careless at driving his car. → John drives his car carelessly.

The manner in which John disappeared was elegant. → John disappeared elegantly.

It was easy for Stanley to win the race. → Stanley won the race easily.

I am frank in saying there is no reason for it. → Frankly, there is no reason for it.

The time at which Harry was known as "The Red Death" was a former time. → Harry was formerly known as "The Red Death".

The extent to which this new development complicated matters is double the extent to which matters were complicated before. → This new development doubly complicates matters.

There are many cases where the related adjective exists but provides no convincing paraphrase for the adverb.

(2.46) a. The men were individually asked to leave.
*It was individual that the men were asked to leave.
*The manner in which the men were asked to leave was individual.

b. Ira readily accepted the offer.
Ira was ready to accept the offer. (wrong meaning)
*The manner in which Ira accepted the offer was ready.

c. Stanley completely ate his Wheaties.
?*Stanley's eating of his Wheaties was complete.
*The degree to which Stanley ate his Wheaties was complete.

d. Irving finally ran away.
*It was final that Irving ran away.
*The event in which Irving ran away was final.

e. Tom absolutely refuses to go.
*The degree to which Tom refuses to go is absolute.
*Tom is absolute in refusing to go.

f. Actually, John can't lose.
*It is actual that John can't lose.

g. I was merely trying to help.
*The $\left\{\begin{matrix}\text{manner}\\\text{degree}\end{matrix}\right\}$ in which I was trying to help was mere.

h. This data virtually shatters the transformational theory.
*The degree to which this data shatters the transformational theory is virtual.

Thus the transformational theory is again faced with the necessity for a large number of more or less lexically specific transformations. For cases like (2.46), a transformational derivation cannot even succeed in properly relating the adjective and the adverb forms, so the motivation for the derivation is vitiated. These are just the kinds of semantic and syntactic irregularities we found in the previous sections.

Because of the great similarity in the syntax of adjectives and adverbs, it is difficult to find arguments against the transformational analysis which are based on the contrast in structure between APs and AdvPs (like the arguments based on (2.1)–(2.10) and (2.29)–(2.36)). However, there is one major syntactic difference between the two kinds of phrases: many adjectives strictly subcategorize a PP or an S in their complements, but the related adverbs do not.

(2.47) a. Tired (of the noise), John left the room.
Tiredly (*of the noise), John left the room.

b. Fearful (of a revolt), the king ordered a purge.
Fearfully (*of a revolt), the king ordered a purge.

c. The manner in which John grimaced was expressive (of his needs).
John grimaced expressively (*of his needs).

d. Happy $\left\{\begin{matrix}\text{at their departure} \\ \text{that they were leaving}\end{matrix}\right\}$, John waved goodbye.

Happily $\left\{\begin{matrix}\text{*at their departure} \\ \text{*that they were leaving}\end{matrix}\right\}$, John waved goodbye.

e. Eager ($\left\{\begin{matrix}\text{with anticipation} \\ \text{to leave}\end{matrix}\right\}$), John chewed his nails.

Eagerly ($\left\{\begin{matrix}\text{*with anticipation} \\ \text{*to leave}\end{matrix}\right\}$), John chewed his nails.

In the transformational theory of adverbs, all the various adverb-creating transformations could be restricted so as to apply only to adjectives without complements; alternatively, in order to capture the generality of the phenomenon, a surface structure constraint could be imposed forbidding complements in adverb phrases. But this latter ploy, the more general of the two, is rather clearly a lame attempt to enforce at the surface a restriction which is easily stated in deep structure under the Lexicalist Hypothesis: the phrase structure rules do not generate complements in adverb phrases. Thus, as in the other two cases we have discussed, there are structural reasons as well as semantic arguments and considerations of irregularity which favor the Lexicalist Hypothesis.

This last consideration also argues against a mixed transformationalist–lexicalist account of adverbs, first proposed (to my knowledge) in Emonds (1970). Emonds observes that there are no deep-structure positions which allow both adjectives and adverbs: adjectives are dominated by NP and Predicate (i.e. the complement of *be, become,* etc.—cf. Chomsky (1965, 107)), and adverbs are dominated by VP, S, and Degree. Emonds suggests that there is a single category A which can appear in all of these positions but receives the *-ly* suffix transformationally or during lexical insertion when it is in the adverb positions.

This account successfully avoids the difficulties of stating adverb formation transformations, since adverbs are not derived from adjective paraphrases but simply from adjectives inserted in adverb positions. Presumably the irregularities in meaning are accounted for in the lexicon, as in the Lexicalist Hypothesis. But there is no way to account for the difference in complement structure; an AP is predicted to have the same structure no matter what position of the sentence it is generated in. Treating adjectives and adverbs instead as separate categories permits the base rules to assign them different complement structures.

In fact, Emonds's theory does not differ from Chomsky's theory of derived nominals, except in one crucial respect: Chomsky distinguishes the two categories *verb* and *noun*, for obvious reasons, but Emonds collapses adjectives and adverbs into a single category, their differences being less apparent. What Emonds's theory and the original transformationalist theory account for is the almost complete similarity of modifier structure and the morphological similarity.

The Lexicalist Hypothesis, on the other hand, is based on expressing the differences between categories. It thus must find a new way to express the similarities, since the use of a common category label is not possible. Again, we need a way to express cross-category generalizations, just as in the previous two cases.

2.4. Coming Attractions

This chapter has shown that the Transformationalist Hypothesis runs up against serious difficulties in dealing with a number of different problems, and that there is a potential solution to these difficulties called the Lexicalist Hypothesis. Chomsky's article "Remarks" sketches a great number of suggestions concerning the formulation of the Lexicalist Hypothesis and its consequences, but does not deal with any of the issues in much detail.

A great deal of subsequent work has verified the need for the theory to express cross-category generalizations, as predicted by the $\bar{X}$ Convention. For instance, Bowers (1968b), Selkirk (1970), Bresnan (1973), and Milner (1973) explore the generalizations among the determiner systems of NPs, APs, and AdvPs that we will address in chapter 6. Dougherty (1970–1971) shows that the generalizations about surface conjunction of various constituents are not to be accounted for by Conjunction Reduction from conjoined sentences, but must be treated as cross-category generalizations in the base. Van Riemsdijk (1973; 1976) and Jackendoff (1973) show that prepositions form a lexical category that participates in the $\bar{X}$ Convention.

Outside the categorial component, Selkirk (1974) shows that the $\bar{X}$ Convention plays a role in determining liaison in certain dialects of French. Akmajian (1975) argues that NP is a cyclic node, as part of an account of an extraposition rule that applies in both Ss and NPs. Jackendoff (1971) shows that the rules of Gapping and VP-Deletion have close NP analogues. Horvath (1976) demonstrates how a topicalization process in Hungarian appears within several different syntactic categories. Chomsky (1973) uses the cross-categorial notion of subject in defining his Specified Subject Constraint on rule application. Perhaps most striking, Bresnan (1976a) revises Chomsky's A-over-A Condition into a cross-categorial "X-over-X" Condition, showing that constraints on rule application depend in part on how cross-categorially the rules are stated.

The Lexicalist Hypothesis has also sparked an interest in the structure of the lexicon. Such works as Halle (1973), Aronoff (1974; 1976), Siegel (1974), Rardin (1975), and Jackendoff (1975a) have begun to open up the lexicon as a rich area of inquiry independent of syntax and phonology.

What has remained relatively neglected is the theory of phrase structure. Though all the works just cited make reference to the $\bar{X}$ Convention, none attempts to explore systematically its implications for questions such as these: How is the $\bar{X}$ Convention to be formulated so as to provide the best account of the syntactic structures of the language? What is the system of syntactic features relating lexical and syntactic categories? Can the notions "possible syntactic category" and "possible phrase structure rule" be characterized?

The rest of this study will be concerned with these questions, and how the grammar of English can be brought to bear on them. Chapter 3 restates the $\bar{X}$ Convention, proposes a theory of syntactic features and a highly restrictive phrase structure schema, and uses them to reanalyze the grammar of the subject and the auxiliary. Chapters 4 through 8 apply the theory to substantial fragments of English phrase structure, defending detailed solutions to problems raised by complements, NP specifiers, the degree system, relative clauses, and degree clauses. Chapter 9 investigates a second phrase structure rule schema, one which generates a previously unremarked class of "deverbalizing rules"; chapter 10 restates a general overview, lest it have been lost in the intervening furor.

3: A Theory of Phrase Structure

Chapter 2 presented arguments that the $\bar{X}$ Convention plays a role in the the theory of syntax. This chapter will formalize what I take to be a quite conservative version of the $\bar{X}$ Convention, the formally simplest and most restrictive statement of the theory that appears adequate for the description of English. In adopting this approach and thus requiring reformulation of many well-accepted analyses which would weaken the theory, I am consciously adhering to the research strategy recommended by Chomsky in the preface of *Syntactic Structures*:

> Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deep understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed.

In attempting to flesh out the $\bar{X}$ Convention into a full theory, I will separate Chomsky's proposal into a number of parts, some of which will be supported here, others of which will not. Essentially, I will retain what I take to be the conceptual core of Chomsky's proposal, but I will modify most of the details.

3.1. Reformulation of the $\bar{X}$ Convention

The $\bar{X}$ Convention can be taken as a theory of syntactic categories in universal grammar, making three principal claims. First, universal grammar includes a set of syntactic distinctive features which defines the possible lexical categories of human languages. A particular language chooses its repertoire of lexical categories from among those provided by universal grammar, just as it chooses a phonological repertoire from the possibilities provided by universal grammar. Presumably the choice must be made in a constrained way—one would hardly expect a language with adverbs but no nouns. But this too is parallel to phonological theory, where one does not expect a language with the sound *ü* but no *a*.

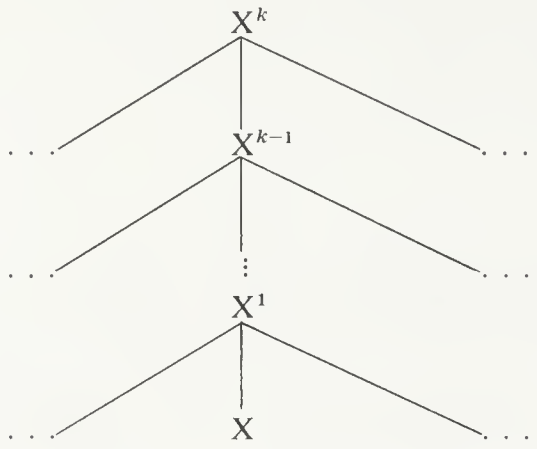
The second claim of the $\bar{X}$ Convention is that each lexical category X defines a set

of syntactic categories X′, X″, . . . , X^k, the *supercategories* of X, related by phrase structure rules of the form (3.1).

(3.1) $X^n \rightarrow \ldots X^{n-1} \ldots$

(3.1) is a phrase structure rule schema provided by universal grammar. It results in phrase structure configurations of the form (3.2).[1]

(3.2)



The *head* of a phrase of category X^n can be defined in two different ways, either as the X^{n-1} that it dominates or as the lexical category X at the bottom of the entire configuration. We make use of both senses in this study, making it clear when necessary which sense is meant. Both reflect traditional usages of the term. Refinements of (3.1) in section 3.3 will guarantee that the head is unique, that is, each X^n will dominate one and only one X^{n-1}. Section 3.6 will take up a number of exceptions to this generalization.

The third claim of the $\bar{X}$ Convention is that rules of grammar are stated in terms of syntactic feature complexes and the prime notation. To be more precise, each term of any rule of grammar must be either a specified lexical or grammatical formative or of the form $\begin{bmatrix} \alpha F_1 \\ \beta F_2 \\ \vdots \end{bmatrix}^i$. This third claim gives substance to the other two claims, since it is by attempting to write rules of grammar in these terms that we test hypotheses about the feature system and the hierarchical organization of categories.

To flesh out these claims, we must answer at least these questions: What are the syntactic distinctive features? Is there a maximum value for *n* in (3.1), and is it the same for each category? Are there any constraints on what may appear in place of the ellipses in (3.1)? Does (3.1) provide the only phrase structure rules of the language or

[1] Ivan Sag has pointed out (personal communication) that this particular aspect of the $\bar{X}$ Convention, the notion that syntactic categories are projections of lexical categories, is a fundamental assumption of Harris's approach to syntax (Harris (1946; 1951)). On the other hand, Harris does not (at least explicitly) use this notion to get at cross-category generalization, nor does he impute any significance to the number of supercategories associated with a lexical category.

are there other types? The rest of this chapter is devoted to formulating one possible set of answers to these questions; subsequent chapters defend these solutions on the basis of the grammar of English.

3.2. The Syntactic Distinctive Features

The choice among competing systems of distinctive features should be made on the basis of how easy it is to state actual rules of the language in terms of the proposed systems. One presumes that rules are more likely to generalize to "natural classes", those that take fewer features to pick out. This criterion is essentially the same as the one used in justifying phonological feature systems, and should therefore be altogether familiar.

In "Remarks on Nominalization" Chomsky analyzes the major lexical categories N, A, and V into the features $\pm N$ and $\pm V$ in the following way:

(3.3)

	+N	−N
+V	A	V
−V	N	

A fourth major syntactic category is P(reposition), about which we will have much to say in this study. Presumably Chomsky's system would analyze it as [−N, −V], completing the matrix in (3.3).

In distributing two distinctive features over four major categories, there are actually only two systems possible other than (3.3):

(3.4)

a.

N	V
A	P

b.

N	V
P	A

The choice among the three must be made on the basis of what categories go together in the rules of English (and other languages, of course). Anticipating the generalizations that appear in the rest of this study, we will adopt (3.4a) as our feature system. It will emerge that there are many rules which generalize across supercategories of N and V, and this is not expected in a feature system like (3.3). Similarly, there are many rules which generalize across supercategories of N and A, and many to V and P, and this is not expected in (3.4b). About the only rule that makes (3.3) and (3.4b) look more natural than (3.4a) is the formation of cleft sentences, in which the clefted item can be only an NP or a PP. Since the combination of N and P is so rare, and the combination of V and A at least equally rare, we will feel justified in provisionally accepting (3.4a) as the major division of lexical categories. As it would be too laborious to test the three systems continually throughout this study, I leave it to the interested reader to check the adequacy of this choice.

The classification of categories made by feature system (3.4a) corresponds to some

traditional divisions of categories: N and A are often called "substantives", and A and P are often thought of as "modifiers". (3.4a) also has the interesting property of corresponding to two significant divisions in the phrase structure rules of English. Though these divisions may not strictly obtain in other languages, we will for convenience name the features after the corresponding properties of English phrase structure.

The first division is between those categories which have a subject, N and V, and those for which no syntactic subject relation exists, A and P. We will accordingly call the feature ±*Subj*, and designate N and V as +*Subj* and A and P as −*Subj*.[2]

To emphasize the heuristic, nontheoretical significance of the names for the features, it is worth pointing out that in French, for example, nouns cannot take NP subjects (other than possessive pronouns, which may be articles). Nevertheless, we assume that French nouns and English nouns have the same syntactic features, and that it is only the way these features appear in the grammar that differs from one language to another.

The second major division in the phrase structure of English is between those categories whose complements may include a surface NP direct object after the head, i.e. V and P, and those categories whose complement cannot contain a surface NP, i.e. N and A. The feature will be named ±*Obj*; V and P are +*Obj* and N and A are −*Obj*. Again, the use of the term "direct object" is strictly syntactic: since there is no NP **his consideration the offer* and no AP **afraid Bill*, but only *his consideration of the offer* and *afraid of Bill*, these categories are not considered to have syntactic direct objects.

Next consider the "minor" lexical categories of English, starting with adverbs. They clearly must be related to adjectives, and we will designate them both as [−Subj, −Obj]. Since their major difference with respect to associated phrase structure is in the ability to take a complement, we will invent a feature to distinguish them and call it ±*Comp*: adjectives are +*Comp* and adverbs are −*Comp*.

We will also use the feature *Comp* to describe other minor lexical categories in the language. The clearest case is the category Particle: particles such as *up*, *on*, and *away* are morphologically identical to prepositions but permit no complements. To express this relationship, we designate both particles and prepositions as [−Subj, +Obj]; prepositions are +*Comp* and particles are −*Comp*.

The difference between modal verbs and ordinary verbs must be stated somewhere in the grammar, and ±*Comp* is a good place to localize it, calling verbs [+Subj, +Obj, +Comp], and modals [+Subj, +Obj, −Comp]. This analysis makes them separate

[2] One might question the claim that adjectives do not have a subject, since it is often assumed that *John* in *John is tall* is the subject of *tall*. This assumption is incorrect. Although *tall* imposes a selectional restriction on the NP in *NP is tall*, the NP bears the grammatical relation "subject-of" to the verb *be*, not to the adjective. This becomes clearer if other verbs are substituted for *be*, e.g. *John became tall*, *John made it tall*, etc. An NP bearing the "subject-of" relation to an adjective would have to be contained in the AP, as the subject of a noun is contained in the NP. Since there is no AP **John('s) fearful* corresponding to the NP *John's fear*, for example, we conclude that adjectives do not have syntactic subjects. For a discussion of how the selectional restriction is imposed without a grammatical relation, see Jackendoff (1974b), especially section 5.

categories, yet their close feature relationship makes plausible the historical creation of the category of modals in English (cf. Lightfoot (1974)): a particular class of verbs underwent change in one syntactic feature, moving into a category available in universal grammar but previously unrealized in English. We will deal with the English auxiliary in somewhat more detail in section 3.5.

Finally, there is a clear candidate for a *−Comp* category associated with nouns: the rather heterogeneous system of articles and quantifiers. Chapter 5 will further divide these into two classes, called Articles (Art) and Quantifiers (Q). Although the actual division will not be determined until chapter 5, we introduce the relevant feature *±Det* here for completeness: articles are [+Subj, −Obj, −Comp, +Det]; quantifiers are [+Subj, −Obj, −Comp, −Det].

We will use the feature *±Det* also to pick out a special class of adverbs, the "degree words" *so, too, as,* etc. We will call this category "Degree" (Deg) and designate it by the features [−Subj, −Obj, −Comp, +Det]; ordinary adverbs will be [−Det]. This choice will be justified in chapter 6; it depends primarily on parallelisms between Art and Deg, which we therefore wish to analyze as differing in only one feature.

There is no place in this system for coordinating conjunctions and complementizers. However, they do not participate in the grammar in the same way as the other categories, in that they do not strictly subcategorize complements and specifiers, so their exclusion is motivated. Section 3.6 will suggest places for them in the general theory. Subordinating conjunctions, on the other hand, can be described as prepositions with sentential complements, so they do participate in the feature system.

The feature system thus looks like this:

(3.5)

	Subj	Obj	Comp	Det
V	+	+	+	
M	+	+	−	
P	−	+	+	
Prt	−	+	−	
N	+	−	+	
Art	+	−	−	+
Q	+	−	−	−
A	−	−	+	
Deg	−	−	−	+
Adv	−	−	−	−

3.3. The Phrase Structure Rule Schema

As stated in section 3.1, the second claim of the $\bar{X}$ Convention is this: every lexical category X must be dominated in phrase structure by a hierarchy of categories X′, X″,

. . . , X^k; furthermore, the set of syntactic categories in a language is completely determined by the lexical categories plus the hierarchical categories projected from them. In formal terms, the claim is that all possible phrase structure rules are of the form (3.1), repeated here:

(3.1) $X^n \rightarrow . . . X^{n-1} . . .$

We will weaken this claim slightly in section 3.6. But aside from minor exceptions, we will claim that (3.1) represents the canonical form for all phrase structure rules.

Notice what sorts of rules (3.1) excludes. One kind of phrase structure rule that has been widely accepted in the literature is so-called Chomsky-adjunction,[3] which appears in a popular source for relative clauses (3.6a), one account of the verb–particle construction (3.6b), one theory of manner adverbs (3.6c), and many other places.

(3.6) a. NP → NP – S
b. V → V – Prt
c. VP → VP – Adv

Since these rules generate no category on the righthand side that is one level lower than the category on the lefthand side, they are impossible within schema (3.1).

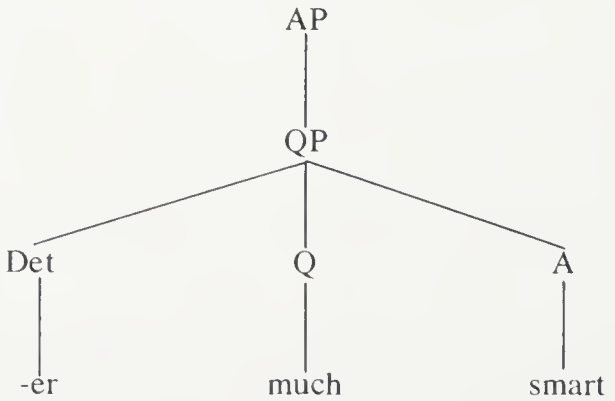
In Vergnaud's (1974) account of relative clauses, there is a phrase structure rule of the following form:

(3.7) N″ → ±Def – N‴ – S

Such a rule is impossible within schema (3.1), since (3.1) requires a rule expanding N″ to contain an N′ somewhere on its right-hand side.

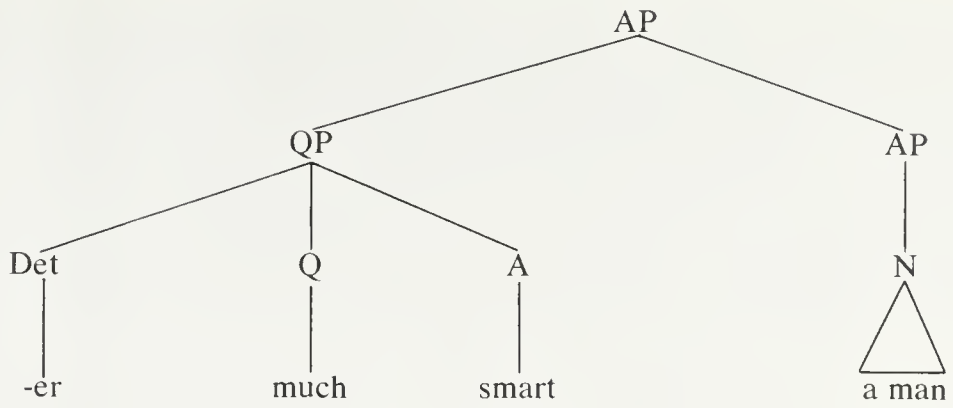
Berman (1974, 109) proposes a number of structures for comparative constructions which illustrate violations of (3.1) and which seem in general counterintuitive. There is nothing in the standard theory of phrase structure to rule them out, though:

(3.8) (= Berman's (127))
a.

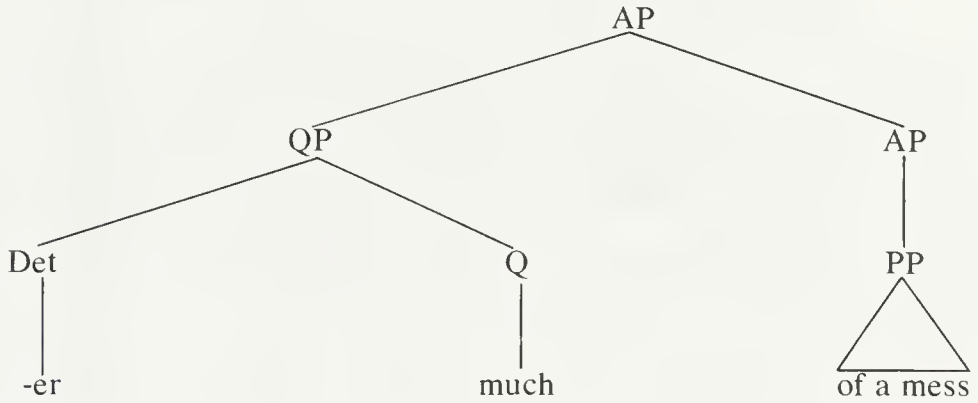


[3] This term has nothing to do with Noam Chomsky, and he regrets its existence (personal communication). However, since it seems firmly established in common linguistic usage, I will retain it, sparingly, with apologies.

b.



c.



In (3.8a), AP dominates QP, but immediately dominates no category A^i. In (3.8b), in addition to the Chomsky-adjunction, the lower AP dominates N rather than an A^i; in (3.8c), the lower AP dominates only PP and no A^i.

Thus (3.1) is a relatively restrictive theory of phrase structure rules, ruling out many counterintuitive structures as well as a number that have been taken rather seriously in the literature. In particular, the Chomsky-adjoined source of relative clauses (3.6a) enjoys a wide following. Chapters 4 through 8 will be concerned with showing that (3.1) permits well-motivated descriptions for a wide range of these constructions, and thus that (3.1) is not too restrictive a theory of phrase structure.

Schema (3.1) leaves open two issues: first, what is the maximum value of n and does it differ from category to category? Second, what is permitted on the righthand side besides X^{n-1}, filling in the ellipses in (3.1)?

Various answers to the first question appear in the literature. In Chomsky's original formulation, n equals 2 for nouns and 3 for verbs (assuming the verb is the head of the sentence). Vergnaud (1974) and Siegel (1974) have n equal to 4, at least for nouns; Dougherty (1968) has n equal to 3 for nouns and 6 for verbs; Jackendoff (1971; 1974a) has n equal to 2 for all categories. The best theory, of course, provides just enough structure to make the relevant structural differences and no more. It now appears to me that n must equal 3 for verbs and nouns. Chapter 6 will show that for reasons of structural parallelism with the syntax of NPs, n must equal 3 for several

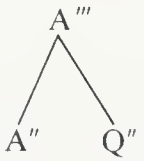
other categories as well, including the minor lexical categories Q and Adv. Thus, at the expense of some otherwise superfluous structure, the most uniform hypothesis is that n equals 3 for all categories. We will call an X''' a *major phrasal category*.

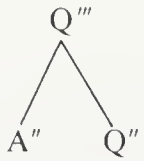
This hypothesis makes possible an answer to the second question above, close to the proposal of Emonds (1976, chapter 1). We will claim that every category to the left or right of X^{n-1} in (3.1) is either a major phrasal category or a specified grammatical formative such as *have*, *number*, *case*, or *tense*, and probably that it is optional. Thus the canonical form (3.1) for phrase structure rules can be refined to (3.9).

(3.9) $X^n \rightarrow (C_1) \ldots (C_j) - X^{n-1} - (C_{j+1}) \ldots (C_k)$, where $1 \leq n \leq 3$, and for all C_i, either $C_i = Y'''$ for some lexical category Y, or C_i is a specified grammatical formative.

We will refer to rule schema (3.9) as the Uniform Three-Level Hypothesis.

(3.9) rules out many further phrase structure combinations, for example the situation in (3.8a,b) in which QP dominates two lexical categories, Q and A. More generally, (3.9) guarantees that the category of a phrase and its head are unambiguously determined by the head: no situation such as (3.10a,b) can arise, in which two different categories dominate precisely the same constituents.

(3.10) a. 

b. 

We combine the feature notation of section 3.2 with the prime notation to arrive at distinctive feature analyses of all syntactic categories. For example, NP (now N''') is designated as [+Subj, −Obj, +Comp]'''; a rule applying either to P' or V' will mention the feature complex [−Obj, +Comp]'. The combination of the two notations makes the desired claim about possible cross-category syntactic generalizations: generalizations must be across categories of the same level, so it is impossible, for example, to pick out both N'' and Q' with a single term of a structural description. This hypothesis, as will be seen, is often crucial in deciding which of a number of possible structures to assign to a construction, in that parallelisms with the structures associated with other lexical categories constrain the choices in interesting ways.

Before going on, let us clarify our terminology. We will continue to use the traditional terms S, NP, AP, AdvP, PP, and QP informally for the major phrasal categories associated with V, N, A, Adv, P, and Q, respectively; the Three-Level Hypothesis claims that these traditional symbols for the major phrasal categories are equivalent to what rule schema (3.9) defines as V''', N''', A''', Adv''', P''', and Q'''. A glossary of all category names and their feature analyses appears at the end of this chapter.

Chomsky uses the term *specifier* to refer to the material in a phrase to the left of the head, and *complement* to refer to the material to the right of the head. According to the tree structures in "Remarks", Chomsky considers *specifier* to represent a syntactic category, but *complement* is simply an abbreviatory term for some concatenation of ordinary syntactic categories. However, there is to my knowledge no evidence that either complements *or* specifiers function as constituents—they do not move or delete as units, and unlike normal constituents, no part can be designated as a head. Consequently, I will use the terms *specifier* and *complement* for expository convenience only, with no theoretical significance implied.[4] In (3.9), the terms $C_1, \ldots, C_j$ will be referred to as X^n *specifiers,* and $C_{j+1}, \ldots, C_k$ as X^n *complements.*

This terminology is suitable for discussion of English, though not perhaps for other languages, since it calls the direct object in an SOV language part of a specifier, not of a complement. In fact, we will see even in English that there are a few cases where a particular grammatical relation is defined as part of an X^i specifier for some categories and part of an X^i complement for others. It must be understood, therefore, that the distinction between specifier and complement is to be regarded here as of no theoretical significance, but only as a convenience.

The next two sections work out two major aspects of English syntax, the subject and the auxiliary, within the Uniform Three-Level Hypothesis. Besides exemplifying the hypothesis in the most immediate way possible, they set the stage for the detailed analyses of complements, specifiers, and relative and degree clauses in subsequent chapters. After the discussion of the subject and the auxiliary, section 3.6 discusses certain phrase structure expansions not covered by the Uniform Three-Level Hypothesis.

3.4. The Generalized Subject Relation

The grammatical relation "subject-of (an S)" is defined by Chomsky (1965, 71) as [NP,S]; that is, in a string exhaustively dominated by S, the subject is a substring exhaustively dominated by NP, and that occurrence of NP is directly dominated by the S. We would like to define the relation "subject of an NP" for derived nominals in such a way that it generalizes structurally with the relation "subject of an S", so that the various rules involving subjects need not be stated separately for NPs and Ss.

This is the first of many analyses to be presented here in which structural parallelism across categories is a crucial consideration. The general principle entailed by the $\bar{X}$ Convention is that if parallel grammatical relations exist in two different categories, the categories must be syntactically parallel with respect to that grammati-

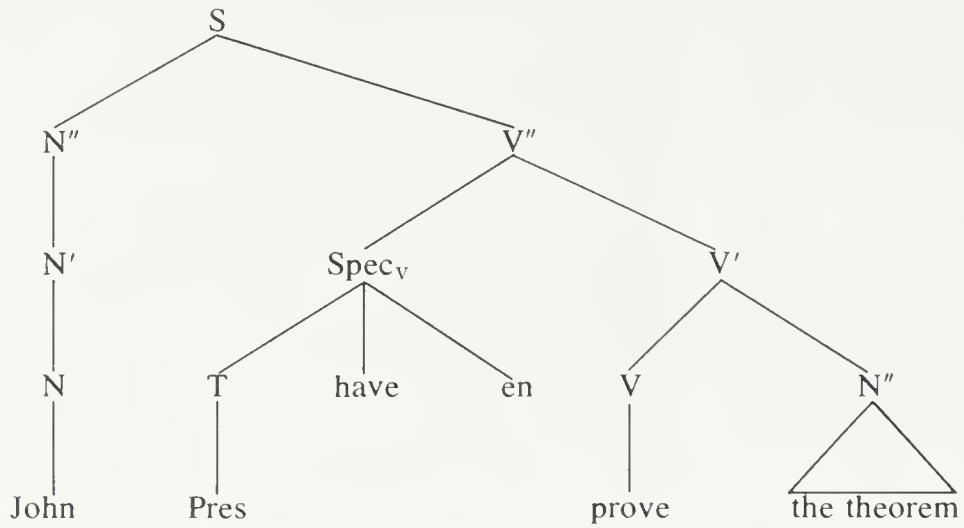
[4] Hornstein (1975) claims that calling the specifier not a category but a concatenation of nodes is unjustified, and that a treatment in which specifier is a category makes a stronger claim. The latter point is correct, in that my view of the specifier ascribes less highly differentiated structure to sentences. However, since neither he nor anyone else has given arguments that this extra structure is necessary (e.g. in terms of movement or deletion), the weaker position seems to be the more supportable.

cal relation. In this way rules involving that grammatical relation can be stated so as to apply to both categories, by appropriate use of syntactic distinctive features.[5]

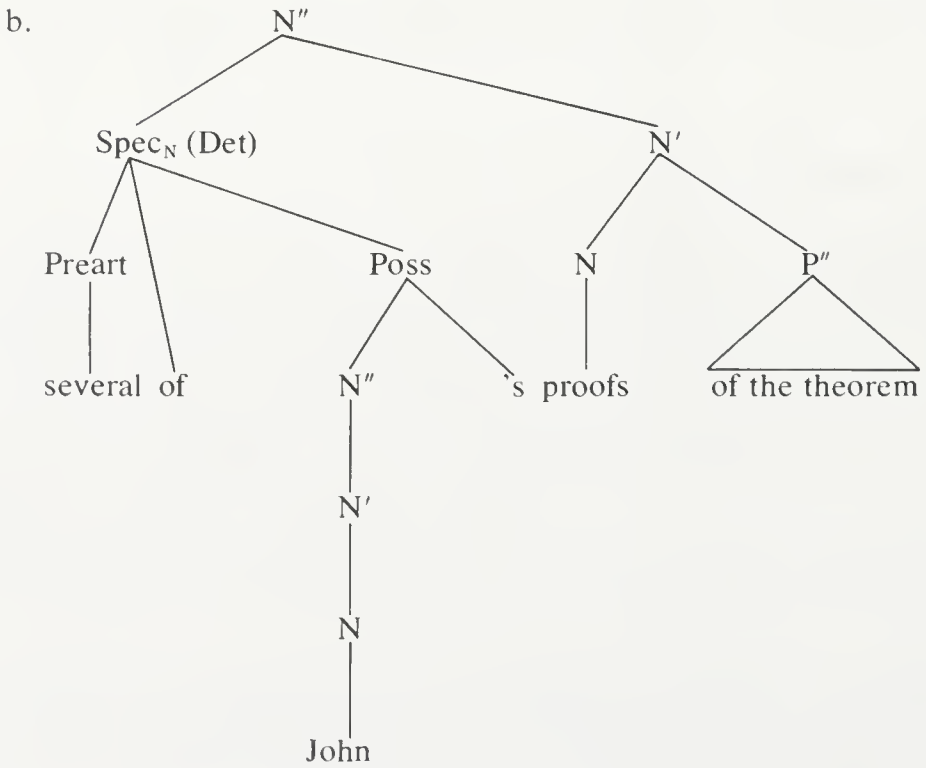
3.4.1. Modifying Chomsky's Analysis

Let us compare Chomsky's proposed deep structures for *John has proved the theorem* and *several of John's proofs of the theorem.*

(3.11) a.



b.

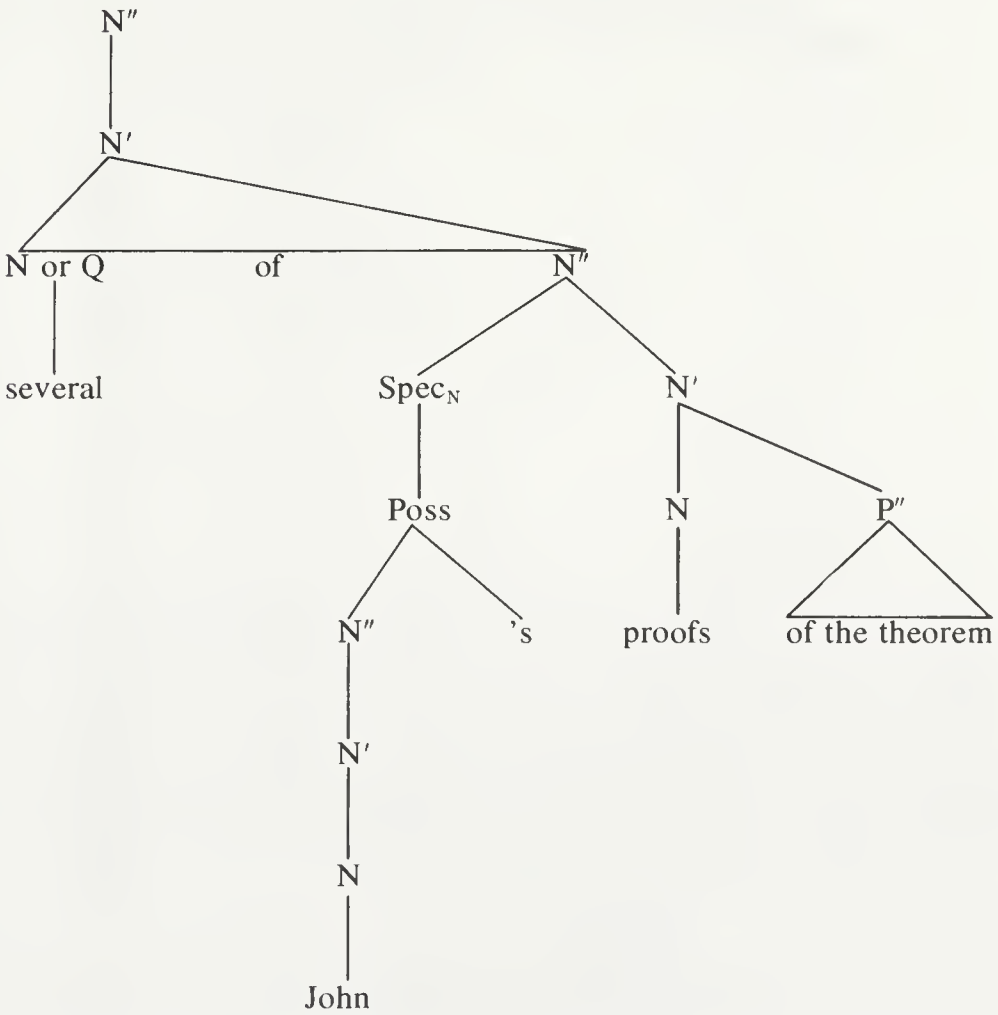


[5] Note that the $\bar{X}$ Convention says nothing about what to do with nonparallel structures. Hornstein (1975) objects to the generalization of the subject relation on the grounds that many other aspects of Ss and

(The details of $Spec_N$ are from Chomsky's (30), p. 37/200.) We would like to say that *John* is subject in both cases. Observe, however, that the subjects are not in structurally parallel positions. In particular, the subject of the S (3.11a) is dominated by the third node above the head verb, but the subject of the NP (3.11b) is dominated by the second node above the head noun. Furthermore, the subject of the NP is embedded in the Specifier node, whereas the subject of the S is directly dominated by S.

Let us first see what can be done about the structure of the derived nominals. We can make some improvement by taking *several of* not as part of the determiner of this NP but as part of a higher NP, thus:

(3.12)



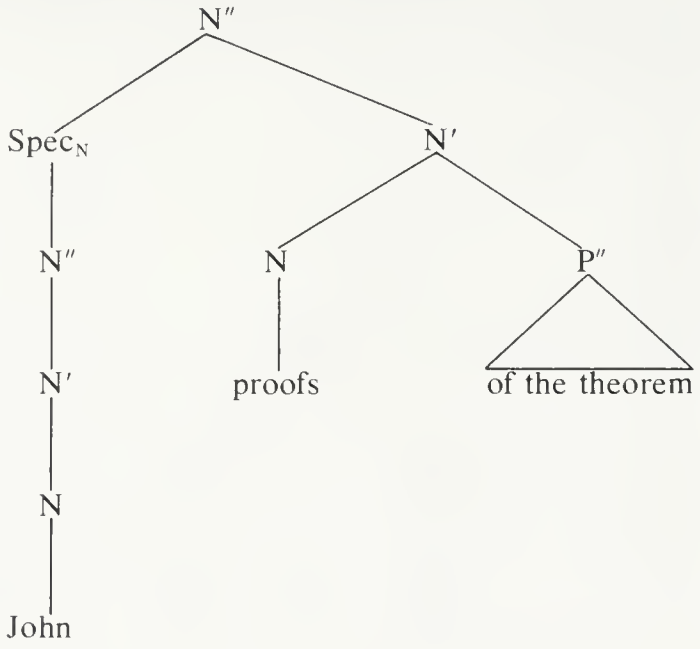
Chomsky suggests this possibility in note 26, p. 38/note 27, p. 219. Essentially this structure for quantifier constructions is argued for in Jackendoff (1968); we will develop it further in chapter 5.

Next, observe that the possessive affix *'s* occurs with all NPs in subject position in nominals, including those that arrive in that position through a transformation (e.g. *the*

NPs are not parallel—for example, Ss have auxiliaries and complementizers, and NPs have determiners. But these differences are irrelevant: the $\bar{X}$ Convention says simply that when parallelisms exist, they must be expressed.

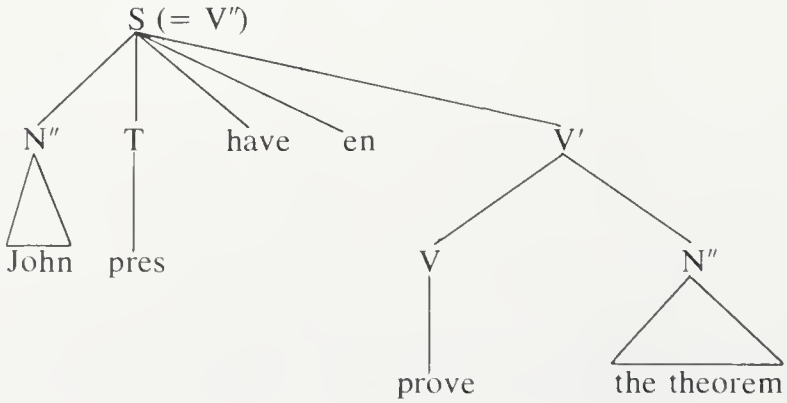
city's destruction by the enemy). This suggests that a relatively late obligatory transformation adds the affix to whatever NP occupies surface subject position. (In section 3.6 gerundive nominals will be analyzed in such a way that they too receive their *'s* by this transformation.) With such a rule, we can eliminate the node Poss in (3.12) and place the NP dominating *John* directly under $Spec_N$, yielding (3.13) instead of the lower N″ in (3.12).

(3.13)

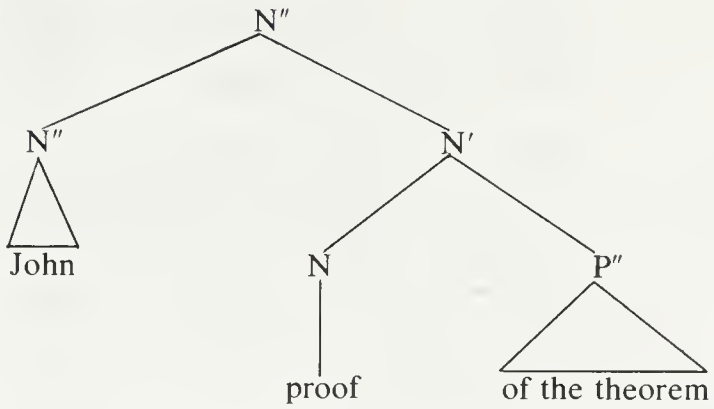


Next suppose we take the view of specifiers proposed in the last section: there is no *category* Specifier; rather "specifier" is simply an abbreviation for a concatenation of nodes. Then $Spec_N$ and $Spec_V$ can be eliminated in (3.11a) and (3.13), and *John* is in nearly parallel positions in the two structures. It remains to modify one or the other of the structures so that *John* is attached the same number of nodes up from the head. By dropping a node in the S (3.11a) we get a parallelism of the form (3.14); by adding an extra node in the NP (3.13) we get the pair (3.15).

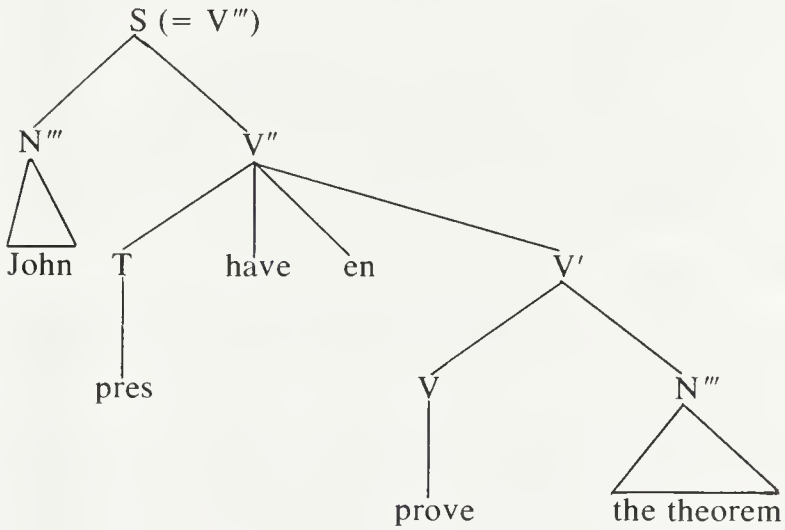
(3.14) a.



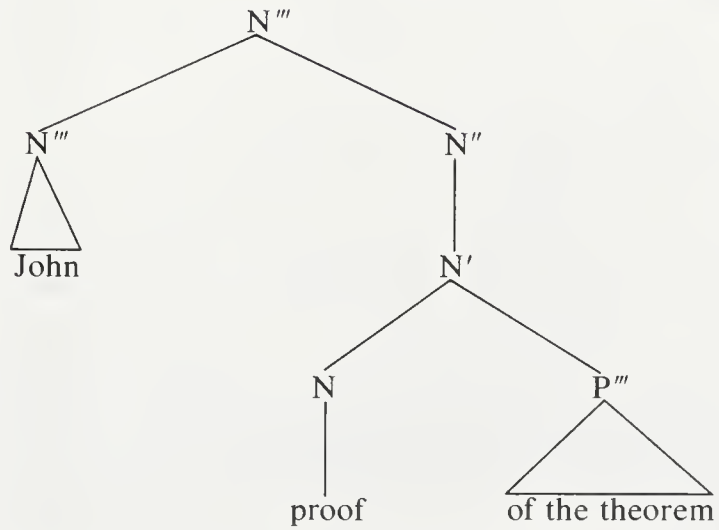
b.



(3.15) a.



b.



(3.14) is the two-level hypothesis of Jackendoff (1974a); (3.15) is the three-level hypothesis proposed in section 3.3. In either pair, *John* is in precisely parallel positions in NP and S; we can therefore define the generalized grammatical relation "subject-of" as [N″, [+Subj]″] in (3.14) and as [N‴, [+Subj]‴] in (3.15). The choice of (3.15) over

(3.14) will be dictated by the analyses of auxiliaries in section 3.5, of complements in section 4.1, and of NP specifiers in sections 5.2 and 6.2.

3.4.2. Arguments for the Uniform-Level Subject Relation

Here are some arguments which favor the uniform-level theories (3.14) or (3.15) over Chomsky's mixed-level theory. The arguments all show that where rules generalize across categories, it is invariably Ss, not VPs, that behave like NPs, APs, and PPs. Hence, by the assumptions of the $\bar{X}$ Convention, Ss must have the same number of primes as the other major phrasal categories.

First, observe that the rule of Topicalization preposes the major syntactic categories NP, AP, and PP:

(3.16) a. My brother(,) everyone expects Bill to like.
b. Taller than Marvin(,) no one ever expected you to be. (in "Yiddish" English)
c. Into the bucket(,) we asked you to put the bananas.

Topicalization also applies to sentences, not to VPs:

(3.17) a. That you were coming(,) no one ever expected Bill to find out.
b. *Coming tomorrow(,) no one ever expected Bill to find out that you were.

If the rule of Topicalization is to generalize across categories, the appropriate generalization thus is to sentence, arguing that it is of the same level as the other major syntactic categories.

The pronoun *it* can have as its antecedent an NP, an AP, or an S:

(3.18) a. A car drove up and I looked at it.
b. She is heavy, but she doesn't look it.
c. Bill came, but I didn't know it.

The anaphoric expression for VP is *do so* or *do it,* which does not generalize directly with any other category. Hence the generalization calls for S rather than VP being similar to NP and AP.

Similarly, NPs, APs, PPs, and Ss serve as antecedents of appositives.

(3.19) a. Charlie talked to Wendy, who was carrying the groceries.
b. Karl is famous, which you'll never be.
c. The tree was in the clock tower, which was certainly an odd place for it to be.
d. Harold left abruptly, which surprised no one.

If the rule which determines antecedents of appositives is to be stated in the simplest fashion, it should generalize over all four categories. In the $\bar{X}$ Convention this is

possible only if all are of the same level. (See sections 4.1 and 7.2–7.3 for further discussion of appositives.)

Certain predicates permit a PP to replace the subject *it*.

(3.20) a. It's a long way to Tipperary. ⇒ To Tipperary is a long way.
b. It's not too long from Groundhog Day to Purim. ⇒ From Groundhog Day to Purim is not too long.

This PP movement appears to generalize with the movement of S into subject, called variously Subject Replacement or Intraposition.[6]

(3.21) a. It's no wonder that Harry's so brilliant. ⇒ That Harry's so brilliant is no wonder.
b. It's a marvel that they ever get along. ⇒ That they ever get along is a marvel.

Again there is no generalization of this sort between PP and VP. Thus the phrase structure rule schema demands that PP and S be of the same level if this generalization is to be captured.

PP and S generalize again in a rule which extraposes certain PPs from within the subject and also extraposes relative clauses.[7]

(3.22) a. A review of that book appeared yesterday. ⇒ A review appeared yesterday of that book.
b. A man who was from Philadelphia came in. ⇒ A man came in who was from Philadelphia.

The rule of Gapping applies in conjoined Ss and NPs, with precisely parallel conditions (see Jackendoff (1971) for details):

(3.23) a. Max plays saxophone and Medusa(,) sarrussophone.
b. Max's recording of Klemperer and Medusa's of Bernstein

There is no corresponding generalization with VP. Again, the way in which the rule generalizes argues that S, not VP, should be of the same level as NP.

Thus there are numerous reasons to consider S structurally parallel to the other major syntactic categories, arguing for the uniform-level theory.

One difference between Ss and NPs, pointed out by Emonds (1976), is that subjects are obligatory in Ss but not in NPs. This fact is an apparent counterexample to the claim in section 3.3 that only heads are obligatory constituents. Since Emonds too accepts this claim, he concludes that the subjects of Ss are outside of the constraints of

[6] I assume here the account of Emonds (1970), not that of Emonds (1976). Some discussion appears in sections 4.3 and 4.8.

[7] Or, if there is not an extraposition transformation, the generalization is in the projection rule. See Akmajian (1975) for evidence that this rule applies also in NP, a further generalization of NP and S.

the $\bar{X}$ Convention, being generated by Chomsky's special phrase structure rule $S \rightarrow N'' - V''$.

However, let us examine the constraints on S subjects a little more closely. It is generally considered to be the case that those sentences which lack surface subjects, namely *to*-infinitives, do so because the underlying subject has been deleted or is the phonologically null form PRO. However, at least two recent articles, Lasnik and Fiengo (1974) and Brame (1975), have argued that certain of these constructions contain surface infinitives with no deep subject. Their claim is that these infinitives are VP complements in deep structure, not Ss. Such a solution violates the phrase structure rule schema (3.9), since it involves generating a V″ to the right of a head. But an alternative is to generate them as subjectless Ss, i.e. as Ss which do not even contain a PRO in subject position. This alternative is equally consistent with Lasnik and Fiengo's and Brame's arguments. Then the phrase structure rules could generate an optional subject in S, its obligatoriness under most conditions being due to conditions extrinsic to the phrase structure rules. Whether such conditions can be independently motivated must be left for future research; but such a solution simultaneously meets Emonds's objection and provides a way of generating so-called VP complements within the present framework. Should this solution not prove viable, we can, of course, always accept a weakening of the optionality condition to allow obligatory subjects.

A second difference between S subjects and NP subjects pointed out by Emonds concerns a general constraint in English against embedding sentences in specifiers, including subjects of NPs.

(3.24)
a. *[a man that Bill knows]'s brother
b. *[six more than Bill saw] boys
c. *a [proud to be an American] man
} —in Spec_N

d. *[more than Mildred is] beautiful
e. *[six hours that we regretted] late
} —in Spec_A

f. *[two bags of flour that we needed] more—in Spec_Q

g. *[the man that you hated]'s destruction of the records—in Spec_N (NP subject)

Note that the bracketed phrases in (3.24) are grammatical in other contexts, and that if the subordinate Ss are eliminated from the bracketed phrases, (3.24a–g) become grammatical. Apparently the only exception in English to this restriction is subject position in Ss, which obviously may contain a wide range of embedded sentences. Emonds argues therefore that this embedding constraint would be more general if the subject were not considered a part of the specifier of V, but were rather outside V″, as in Chomsky's system. In a uniform level theory, on the other hand, the sentence must be an exception to the embedding constraint.

However, there are languages in which even subjects may not contain sentential complements. One is Walbiri, as pointed out by Hale (1975). Interestingly enough, another is Old English, as pointed out by O'Neil (1976), who studies how relative clauses in subjects came to be possible in English through a structural reanalysis of certain topicalized constructions. In order to maintain Emonds's generalization, one would have to make the rather dubious claim that the historical change in English involved altering phrase structure in a radical way, moving the subject from V″ into V‴. A much more plausible change is that the embedding constraint itself was relaxed to permit the observed exceptions; this accords with O'Neil's analysis. Thus, accepting the nongenerality of Emonds's embedding constraint does not seem at all to be a consequence damaging to the uniform-level theory.

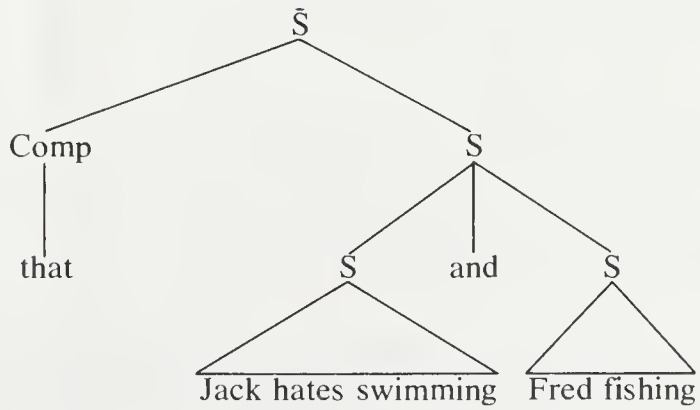
3.4.3. Treatment of the Complementizer

One problem that remains in justifying the Uniform-Level Hypothesis is the treatment of the complementizer. The question is where the complementizer is attached to the sentence. All the movement rules cited in section 3.4.2 move the sentence with its complementizer, and no rule ever leaves the complementizer behind. Thus, the simplest hypothesis is that the complementizer is a left sister of the subject, attached to V″ in (3.14a) or to V‴ in (3.15a). Unfortunately, this hypothesis appears to run afoul of the analysis of Gapping, for Gapping applies only if the complementizer is absent from the second clause:

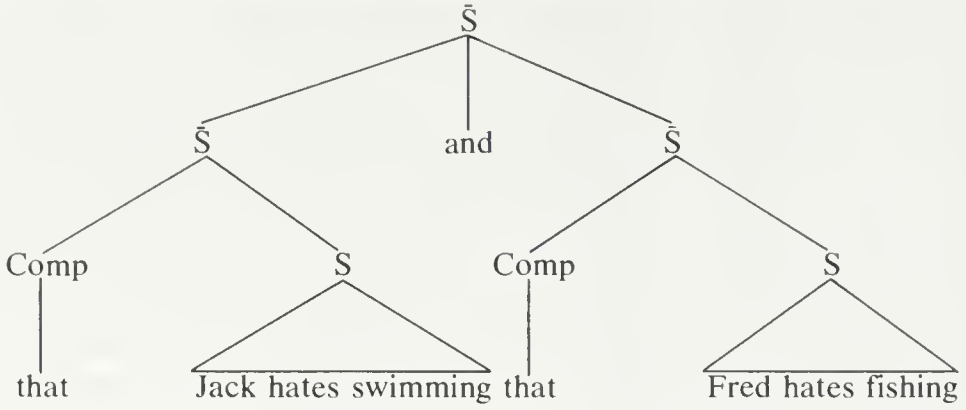
(3.25) a. It is hard to believe that Jack hates swimming and (*that) Fred fishing.
b. For Jack to hate swimming and (*for) Fred fishing would be a tremendous surprise.

One explanation of this is that the difference in structures is as shown in (3.26a) versus (3.26b).

(3.26) a.



b.



If Gapping were defined over S conjunction but not over S̄ conjunction, the difference would be immediately accounted for.

Other evidence against the theory that the complementizer is the sister of the subject is provided by Bresnan (1974), who shows that Right Node Raising applies only to single constituents, but that it can apply to sentences with or without their complementizers:

(3.27) a. Mike wouldn't tell us, but Randy readily volunteered, that Jenny was drinking again.
b. I've been wondering whether, but wouldn't positively want to state that, your theory is correct.

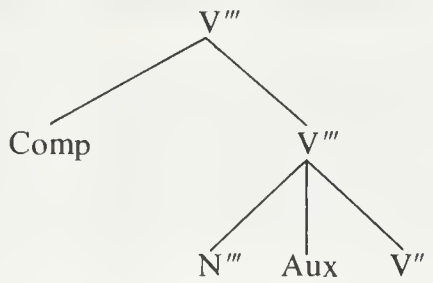
Again this argues that the complementizer is a sister of the entire S, not of the subject.

This raises a difficulty for the generality of movement rules under the $\bar{X}$ Convention: since S̄ is one level higher than NP, i.e. apparently V^4, it cannot be moved by the same rule that moves an N''', a P''', or an A'''. There are at least two possible ways to deal with this problem. One would be to claim that S̄ is indeed V^4 (assuming a three-level S and NP), but that the A-over-A Convention actually generalizes over category levels: the structural description of a rule would be met by the highest-level constituent of the category mentioned by the rule. For instance, by mentioning V in a rule, the structural description would be met by V^4; by mentioning N, the structural description would be met by N'''. In order to pick out the lexical category verb, a rule would have to mention specifically V^0. This is a not unattractive solution, although some care would be necessary in checking its feasibility. It has the interesting property of vitiating all the above arguments for parallelism based on movement rules, since a rule mentioning, say, +*Subj* will apply to either N''' or V^4.[8]

[8] Emonds has pointed out (personal communication) that this solution is in some sense a return to one attractive aspect of Harrisian transformations: one would state, say, an NP movement rule as "move N"; all the modifiers would be moved along automatically. This emphasizes the centrality of the notion "head" to syntax in a highly suggestive fashion.

A less drastic solution is to claim that both S and $\bar{S}$ are instances of V‴:

(3.28)



Under this proposal, the A-over-A Convention guarantees that any movement rule moves $\bar{S}$ rather than S, while preserving the generalization with the movement of other X‴. Such a structure represents a weakening of the theory of phrase structure, since it contains the forbidden configuration of a category dominating itself. Section 3.6 and chapter 9 will show, however, that this exception can be accommodated within a highly constrained class of rules, generated by a second phrase structure rule schema of considerable interest.

Which of these two solutions for the complementizer turns out to be correct, however, does not seem to be of major importance to the main hypothesis. For simplicity, we will assume the second solution, in which the complementizer is Chomsky-adjoined to S. This problem disposed of, the generalization of subjects provides considerable evidence for the Uniform-Level Hypothesis.

3.5. The Auxiliary

The problem of this section is where the various parts of the auxiliary are attached and whether they provide evidence for the general structure of the S. The position adopted here is a modification of the analysis in Jackendoff (1972, section 3.8), based in part on 1966 lectures by Klima and on Emonds (1970); other revisions have been proposed by Akmajian and Wasow (1975). In turn, these are all variants of the original analysis of Chomsky (1957).

The verb and its complements apparently form a constituent that does not include the modal, since there is a rule (pointed out by Ross (1967)) which fronts a VP after modals under certain conditions:

(3.29) They said she may attempt to leave, and attempt to leave she will.

The rule of VP-Deletion provides evidence that the verb and its complements form a constituent which does not include the aspectual verbs *have* and *be*, since a VP may delete, leaving aspect intact.[9]

[9] Hornstein (1975) presents this as evidence against the two-level theory of S, in which aspect must be a daughter of V′. As will be seen, a three-level theory accommodates it.

(3.30) a. We asked them to stay, and they could have.
b. Michael said he had been practicing, but I don't really believe he had been.

Jackendoff (1972) gives evidence that the first auxiliary is a daughter of S, but that subsequent auxiliaries are not daughters of S. The evidence is that sentence adverbs such as *frankly*, *probably*, and *evidently* occur in all possible positions as daughter of S—initial, final with comma intonation, and before the auxiliary. They also occur after the first auxiliary, but not after subsequent ones.

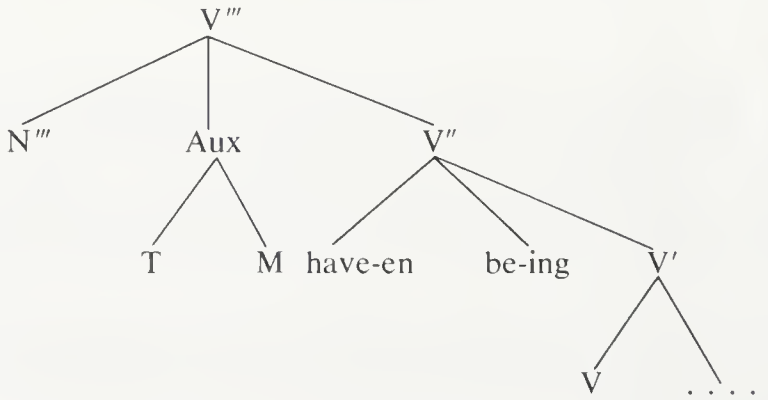
(3.31) a. George will $\left\{\begin{matrix}\text{frankly}\\\text{probably}\\\text{evidently}\end{matrix}\right\}\left\{\begin{matrix}\text{have amused}\\\text{be amusing}\end{matrix}\right\}$ the children by the time we get there.

b. ?*George will $\left\{\begin{matrix}\text{have}\\\text{be}\end{matrix}\right\}\left\{\begin{matrix}\text{frankly}\\\text{probably}\\\text{evidently}\end{matrix}\right\}\left\{\begin{matrix}\text{amused}\\\text{amusing}\end{matrix}\right\}$ them when we get there.

A simple way to account for this difference is to claim that the first auxiliary is a daughter of S but that the second auxiliary forms a constituent with the verb and its complements; then an adverb following the first auxiliary can be a daughter of S, but an adverb following the second auxiliary cannot be.

These three arguments for constituency lead to a three-level theory of the S:

(3.32)



One of the problems in Chomsky's original analysis of the English auxiliary is the recurrence of the configuration $\left\{\begin{matrix}\text{M}\\\textit{have}\\\textit{be}\end{matrix}\right\}$ in transformational rules as the first verbal element, including all uses of *be* and the aspectual use (and in some dialects the possessional use) of *have*. The solution of Klima, Emonds (1970), and Jackendoff (1972) is an obligatory transformation *Have–Be* Raising which moves the appropriate uses of *have* and *be* to a position under Aux, just in case there is no modal present.

After this transformation, any transformation can refer to the first auxiliary by mentioning the node Aux in its structural description, and the need for the repetitive braces is eliminated.

McCawley (1975) objects to the rule of *Have-Be* Raising on the grounds that it is entirely arbitrary to treat *have* and *be* as unlike other verbs in undergoing this rule; and that if modals are a category separate from verbs, there is no reason for modals, *have*, and *be* to behave alike. The theory of categories of sections 3.2 and 3.3, however, makes possible a refinement of the analysis which meets McCawley's objections in a rather interesting way.

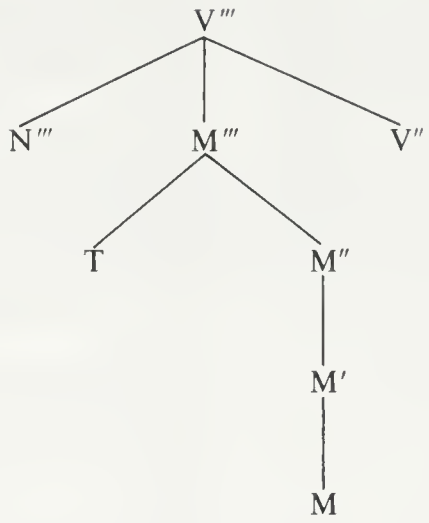
Recall that the syntactic distinctive features of Modal are [+Subj, +Obj, −Comp], contrasting with verbs only in the last feature. Now suppose that the lexical entries for the appropriate uses of *have* and *be* are assigned the syntactic features [+Subj, +Obj], being unmarked for *Comp*. Then, by the usual conventions for applying rules, any rule which applies to the feature complex for modals, [+Subj, +Obj, −Comp] will also apply to *have* and *be* but to no other verbs. In other words, [+Subj, +Obj] corresponds to Chomsky's (1957) category *v*; [+Subj, +Obj, +Comp] to *V*; [+Subj, +Obj, −Comp] to $\left\{\begin{matrix}\text{M}\\ \textit{have}\\ \textit{be}\end{matrix}\right\}$. The exceptionality of *have* and *be* and their falling together with modals is expressed naturally in terms of such a lexical feature analysis; and the description of the auxiliary thus looks considerably less arbitrary than before.[10]

For a further refinement, observe that the Uniform Three-Level Hypothesis claims that Modal is the progenitor of a series of larger categories M′, M″, and M‴. The category Aux, which so far has not played any role in the $\bar{\text{X}}$ Convention, can therefore be analyzed as M‴. Although there is no empirical evidence that as many as three levels are necessary, the superfluous structure is harmless and makes a more highly constrained theory possible; furthermore, the extra rules are part of universal grammar, and thus add no cost to particular grammars. We will arbitrarily attach Tense as a daughter of M‴ rather than at any of the lower levels, yielding the following phrase structure rules:[11]

[10] Wendy Wilkins has pointed out (personal communication) that this analysis brings to mind the arguments in Chomsky and Halle (1968, 382–385) which show a possible illegitimacy in the use of partially specified phonological feature matrices. However, examination of those arguments reveals that they depend on the fact that phonological rules may alter distinctive features. Since, under the assumptions of the Lexicalist Hypothesis, syntactic category features are never changed by transformations, the use of partially specified syntactic features is not susceptible to Chomsky and Halle's criticisms.

[11] Akmajian (personal communication) has objected to this approach to the auxiliary on the grounds that it is a distortion of the actual facts of English, and that there are auxiliary constituents in other languages which show no evidence of a modal. The approach taken here, however, is forced on me by the highly restrictive phrase structure rule schema (3.9). The quotation from Chomsky in the introduction to this chapter is relevant here: either this analysis will prove that the theory must be weakened in a highly specific way, or else it will provide a new analysis of what is otherwise an unprincipled exception to the general theory.

(3.33) a. $V''' \rightarrow N''' - M''' - V''$
b. $M''' \rightarrow T - M''$
c. $M'' \rightarrow M'$
d. $M' \rightarrow M$



Now since the phrase structure rule schema also predicts that a head is obligatory in M‴, the only way a modal can be omitted from a tensed sentence is to generate the empty node Δ under M. Thus both *Have–Be* Raising and *Do* Support can be stated in such a way as to fill this empty node with a category nondistinct from M, making one further peculiarity of the English auxiliary somewhat less peculiar.[12]

One further minor point: the phrase structure rule schema (3.9) claims that all complements and specifiers are optional, at least from the point of view of phrase structure rules. This is certainly true of M‴ in rule (3.33a), since in infinitive clauses there is no evidence that there has ever been a tense or modal present. We can thus feel justified in regarding the obligatoriness of Aux in main clauses as a semantic condition, or at least as a condition extrinsic to the phrase structure rules proper.

It is a little more difficult to justify making Tense optional in (3.33b), since the accepted syntactic treatment has always required it. Its obligatory presence seems in any event to be a minor exception to the overall generalization that only heads are obligatory in phrase structure expansions.[13]

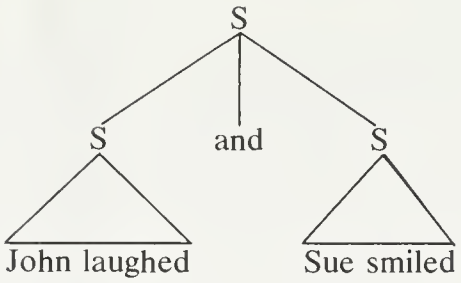
3.6. Two Exceptions to the Uniform Three-Level Hypothesis

One obvious exception to the theory of phrase structure presented so far is coordination. It is generally agreed that the node dominating conjoined Ss is an S, that the node dominating conjoined NPs is an NP, and so forth.

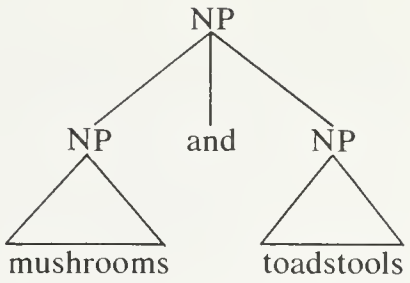
[12] If there are no auxiliaries, and Affix Hopping attaches Tense to the main verb, presumably the entire M‴ is deleted, so as to meet Emonds's (1976) condition that no empty nodes appear in surface structure.

[13] Emonds (1976) suggests that only phrasal categories are always optional; grammatical formatives are optional or obligatory. This weakening accounts for at least the present cases properly.

(3.34) a.



b.



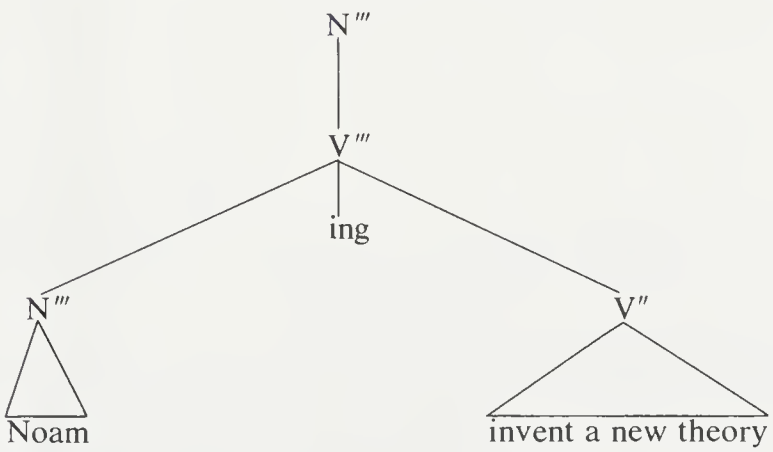
This violates even the hierarchic arrangement of categories predicted by (3.1), the primitive version of the phrase structure rule schema. Furthermore, no constituent of a coordinate construction can be identified as its head. Clearly a separate phrase structure rule schema is necessary. (3.35) is one possible form.

(3.35) $X^i \rightarrow X^i - (\text{conj} - X^i)^*$

This permits coordination of any syntactic category. Whether it can be refined to a more restrictive form is left for future research.

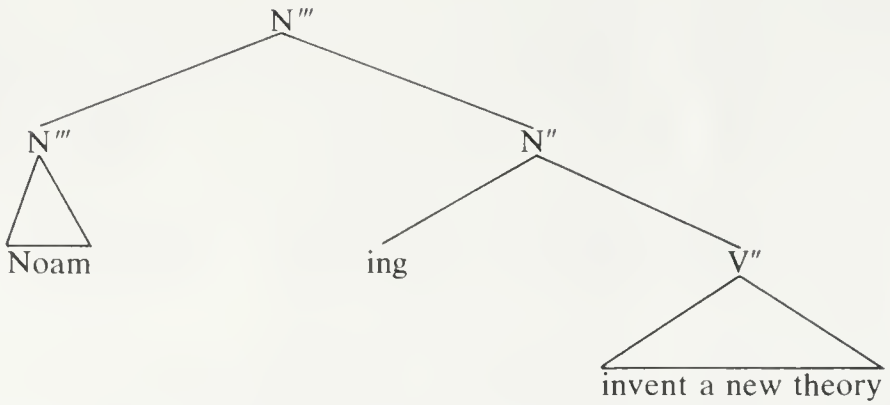
A different sort of exception to the phrase structure rule schema emerges from considering the structure of gerundive nominals such as *Noam's inventing a new theory*. The arguments of chapter 2 showed that they are closely related to sentences, having the same possibilities for modification as sentences and having interpretations with direct and productive relationships to the corresponding sentences. Emonds (1970) shows, furthermore, that (unlike *that* and *for-to* complements) they have precisely the distribution of NPs. Hence a first approximation to their structure might be (3.36).

(3.36)



Schachter (1976) and Horn (1975) independently suggest an alternative structure, in our terms represented as (3.37).

(3.37)



(3.37) has at least three immediate advantages over (3.36). First, it enables the rule inserting *'s* (POSS) on the subject of an N‴ to be generalized to the subject of gerunds without further ado. Second, it makes possible a certain limited class of gerunds such as (3.38) (quoted from Schachter).

(3.38) a. There is no enjoying this world without thee.
b. This telling tales out of school has got to stop.

The replacement of the subject by a normal NP determiner is inexplicable in structure (3.36), but is an expected possibility in (3.37). Third, it explains why gerunds allow aspect (a V″ specifier) but not tense and modal (V‴ specifiers), since gerunds contain a V″ but not a V‴. Horn presents further arguments that the subject of gerundive nominals behaves like the subject of an NP, not an S, where these differ.

The differences between gerundive and derived nominals observed in chapter 2 can still be explained in terms of (3.37), since they all depend on the difference between the internal structure of V″ and N″, a difference which is still present under the theory of (3.37). Since the Lexicalist Hypothesis generalizes the subject relation to NP, *Noam* can still be interpreted as the subject of the verb *invent*, at worst by ignoring the feature *+Obj* at some point in the derivation. Hence (3.37) appears to be a viable alternative theory of gerundive nominals, with a certain amount of evidence in its favor.

Observe, however, that neither (3.36) nor (3.37) conforms to the phrase structure rule schema, since the main N‴ has no noun head. Hence it is necessary in either case to treat gerundives as an exception to the schema, by adding phrase structure rule (3.39a) to generate (3.36) or (3.39b) to generate (3.37).

(3.39) a. $N''' \rightarrow V'''$
b. $N''' \rightarrow$ ing - V''

Either of these rules provides a way to use a verb phrase as a noun phrase, i.e. not as a verb phrase with respect to the context in which it is embedded. I believe that these rules are not an unprincipled exception to the phrase structure rule schema of section 3.3, but are part of a class of "deverbalizing" phrase structure rules described by the schema (3.40).

(3.40) $X^i \rightarrow af - V^i$

That is, a deverbalizing rule must expand a category as the verbal supercategory of the same level, and there may be no complements or specifiers other than an affix or other grammatical formative indicating the change of category.

This class of rules has up to now not been distinguished, perhaps because there has been no theory of phrase structure sufficiently constrained that rules such as (3.39) emerged as exceptional. Chapter 9 presents a number of candidates for deverbalizing rules, some of which are not at all obscure.

One particularly relevant example should be mentioned now. Recall that in section 3.4.3 we proposed that the complementizer be generated with rule (3.41).

(3.41) $V''' \rightarrow Comp - V'''$

As Emonds has pointed out (personal communication), this rule is an instance of schema (3.40), where X happens to be V, and the complementizer is a particular kind of affix or grammatical formative. Thus the apparently exceptional treatment of the complementizer takes its place as one of a principled set of violations of the main phrase structure rule schema.

3.7. Summary and Generalization of Rules

The $\bar{X}$ Convention as formulated in section 3.1 makes three claims: the class of possible lexical categories is determined by a set of distinctive features; the class of syntactic categories is determined by elaborating the lexical categories in terms of the prime notation; rules of grammar are to be stated in terms of these features and primes. We have argued that this theory can be strengthened into the Uniform Three-Level Hypothesis: for every lexical category X, there are syntactic categories X′, X″, and X‴, and no more, and the major phrase structure rules elaborating these categories are of the form given by rule schema (3.9), repeated here:

(3.9) $X^n \rightarrow (C_1) \ldots (C_j) - X^{n-1} - (C_{j+1}) \ldots (C_k)$, where $1 \leq n \leq 3$, and for all C_i, either $C_i = Y'''$ for some lexical category Y, or C_i is a specified grammatical formative.

One argument has appeared which bears on the need for three levels: the constituent structure of the auxiliary requires a level for aspectual verbs between the S level and the V′ level (section 3.5). Further arguments appear in subsequent chapters.

Rule schema (3.9) provides the bulk of the phrase structure rules of the language, but there are at least two other schemata which generate possible configurations ruled out by (3.9): the schemata for coordination (3.35) and for deverbalizing rules (3.40).

(3.35) $X^i \rightarrow X^i - (conj - X^i)^*$
(3.40) $X^i \rightarrow af - V^i$

These provide places in universal grammar for the parts of speech which do not

participate in the $\bar{X}$ Convention, namely coordinating conjunctions, which appear in (3.35), and complementizers, which appear as one realization of *af* in (3.40).

The theory leaves a number of questions in my mind. The first group concerns the feature system. It is clear that a theory of markedness for syntactic categories is necessary, which will predict the relative probability of parts of speech across languages: nouns and verbs are presumably universal, adjectives and prepositions less common, adverbs rarer, modals rarer still. Also, one of the drawbacks of the present feature system is that it incorrectly predicts that generalizations across *−Comp* categories should be as common as those across *+Comp* categories. Is this a question of markedness? Or are there substantive universals outside the feature system (such as the presence of Tense in sentences and deictic elements in NPs) which simply override possible generalizations?

A second group of questions concerns the adequacy of the constraint on the form of C_i in schema (3.9). Is it necessary to weaken the claim that all C_i are optional, and if so, how? Is the restriction to Y''' viable? If not, what is a properly restrictive way of weakening it?

Continuing with the summary of the chapter, here is the collection of phrase structure rules generated by (3.9) that have been mentioned so far.

(3.42) *X‴ Rules*

	Traditional Notation	$\bar{X}$ Notation
a.	$S \rightarrow (NP) - (Aux) - PredP$	$V''' \rightarrow (N''') - (M''') - V''$
b.	$NP \rightarrow (\begin{Bmatrix} NP \\ ArtP \end{Bmatrix}) - Nom$	$N''' \rightarrow (\begin{Bmatrix} N''' \\ Art''' \end{Bmatrix}) - N''$
c.	$Aux \rightarrow T - M''$	$M''' \rightarrow T - M''$

(3.43) *X″ Rules*

a.	$PredP \rightarrow (have - en)\ (be - ing) - VP$	$V'' \rightarrow (have - en)\ (be - ing) - V'$
b.	$Nom \rightarrow N'$	$N'' \rightarrow N'$
c.	$M'' \rightarrow M'$	$M'' \rightarrow M'$

(3.44) *X′ Rules*

a.	$VP \rightarrow V - (NP)$	$V' \rightarrow V - (N''')$
b.	$N' \rightarrow N - (PP)$	$N' \rightarrow N - (P''')$
c.	$M' \rightarrow M$	$M' \rightarrow M$

The main generalization worth noting at this point is that between (3.42a) and (3.42b), involving the position of the subject. These two rules can be generalized via feature notation into (3.45).

(3.45) $$\begin{bmatrix} X \\ +Subj \\ \langle +Obj \rangle \\ +Comp \end{bmatrix}''' \rightarrow \left(\begin{bmatrix} +Subj \\ -Obj \\ \langle +Comp \rangle \\ +Det \end{bmatrix}''' \right) - \langle (M''') \rangle - X''$$

As in phonology, the angle brackets indicate simultaneous occurrence: if *+Obj* is

chosen on the left, *+Comp* must be and M‴ may be chosen on the right; if *−Obj* is chosen on the left, *Comp* is unspecified on the right, so either N‴ or Art‴ may be chosen, and M‴ may not be chosen.

Unfortunately, the clumsiness of this notation will turn out to be proportional to the complexity of the rules being collapsed. (3.45) is a relatively simple case, and far more unsightly cases will appear in the next chapter. We will eventually conclude that the traditional abbreviatory conventions are not sufficient to express the linguistically significant generalizations of English phrase structure. Insofar as the choice of abbreviatory conventions is directly linked to the evaluation measure, this conclusion raises deeper questions for linguistic theory than we will be able to answer here.

On the following page is a glossary of lexical and syntactic categories.

3.8. Glossary of Lexical and Syntactic Categories

Major Category Grouping	X	Complete Feature Analysis	Traditional Category Name for X′	for X″	for X‴
[+Subj, +Obj]	V	[+Subj, +Obj, +Comp]	VP	PredP	S, S̄
	M	[+Subj, +Obj, −Comp]			Aux
[+Subj, −Obj]	N	[+Subj, −Obj, +Comp]	Nom	Nom	NP
	Art	[+Subj, −Obj, −Comp, +Det]			Det
	Q	[+Subj, −Obj, −Comp, −Det]			QP
[−Subj, +Obj]	P	[−Subj, +Obj, +Comp]			PP
	Prt	[−Subj, +Obj, −Comp]			
[−Subj, −Obj]	A	[−Subj, −Obj, +Comp]			AP
	Deg	[−Subj, −Obj, −Comp, +Det]			Det (?)
	Adv	[−Subj, −Obj, −Comp, −Det]			AdvP

4: Complements

The Uniform Three-Level Hypothesis of chapter 3 predicts that complements to a phrase can be attached in three possible places: to X′, X″, or X‴. Since any constituent following the head of a phrase could in principle be attached in any of the three complement positions, we must be concerned with how to pick out the correct position for various sorts of complement phrases. This chapter will explore the complement systems of each of the major categories V, N, A, and P. We will show that there are principled distinctions among the three levels of complements, and that the complement system of English reveals substantial cross-category generalization.

4.1. The Three Levels of Complements

If we classify complements on semantic grounds, we find that there are three distinct ways in which a complement may be integrated into a semantic interpretation: as a functional argument, as a restrictive modifier, and as a nonrestrictive modifier. We will attempt to identify these respectively with X′, X″, and X‴ complements. Thus we will be able to claim that there is a strong correlation between syntax and semantics in the complement system.[1]

Let us begin with functional arguments. Those lexical items which strictly subcategorize phrases in their environment can be thought of as semantic functions which take as their arguments the interpretations of the strictly subcategorized phrases. For example, the verb *give* strictly subcategorizes a subject, an object, and an indirect object, and can be thought of as a semantic function $f(x,y,z)$ which maps ordered triples of terms into propositions. Such an approach is developed in Katz (1966; 1972) and Jackendoff (1972; 1976), among other works, and has been used implicitly in almost every approach to semantic interpretation in generative grammar. Similarly, the noun *part* (*of*) strictly subcategorizes an NP, and can be treated semantically as a function $g(x)$ which maps terms into terms; the adjective *proud* (*of*) maps terms into properties

[1] Williams (1975) comes to similar conclusions about the complement system, on the basis of somewhat different evidence. Readers familiar with his work will notice that I have only three complement categories instead of his four. His types I, II, and IV correspond to X′, X″, and X‴ complements, respectively. His type III, the most problematic for him, appears to divide into X″ and X‴ complements in a way not fully clear to me.

(or predicates, or whatever kind of object the intension of adjectives is); the preposition *to* maps terms into (intensions which pick out) directions. Aside from subjects, all functional arguments in English come immediately after the head, typically preceding all other modifiers, and many grammarians have had the intuition that they are the most "tightly bound" to the head of all the complements. This intuition can be expressed by assigning all and only the strictly subcategorized phrases to the X′ complement in deep structure.

In many cases a phrase can be identified as strictly subcategorized by the fact that it cannot be omitted from the sentence without incurring ungrammaticality. For example, the PP in *Joe put the book on the table* must be strictly subcategorized, since **Joe put the book* is ungrammatical. However, this is only a sufficient condition for strict subcategorization, not a necessary condition, since many words optionally subcategorize phrases. For example, *tell* occurs in the contexts *John told Bill a lie* and *John told Bill*, but *a lie* appears to operate semantically as a functional argument in the former sentence and hence must be strictly subcategorized. Similarly, when adjectives and nouns strictly subcategorize, their arguments are usually optional.

A second criterion for X′ complements in Ss and NPs is provided by certain anaphoric processes. The phrase *do so* appears to be a pro-V′, and may be followed only by material which is outside the V′ complement (usually part of the V″ complement).[2] Consider (4.1).

(4.1) a. Joe bought a book on Tuesday, but Sim did so on Friday.
b. *Joe put a book on the table, but Sim did so on the chair.

The ability of *on Tuesday* to follow *did so* indicates that it is a V″ or V‴ complement in this sentence; the inability of *on the chair* to follow *do so* indicates that it must be inside of V′ in the antecedent sentence.

A parallel anaphoric process in NPs is the use of the pro-N′ *one*. Lakoff (1970a) observes that there is a contrast in the applicability of *one* which depends on the nature of the complement.

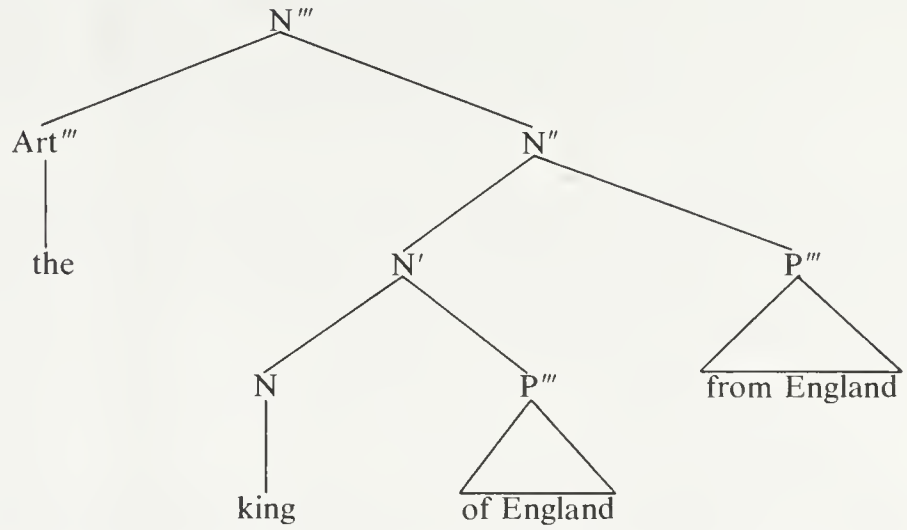
(4.2) a. Jack met the king from England, and I met the one from France.
b. *Jack met the king of England, and I met the one of France.

Notice that the PPs *of England* and *from England* play different semantic roles, since there may be a king of England from France. *Of England* appears to be a functional argument, since it specifies part of the function of the king; *from England*, on the other hand, specifies a somewhat inessential part of kinghood. Notice further that the two PPs may appear together only in one of the two possible orders: we cannot get **the king from France of England*. These facts can be described simultaneously if we suppose that *of England* is an N′ complement and *from England* is an N″ complement.

[2] This distinction in the use of *do so* was first observed by Lakoff and Ross (1966). I assume that the interpretation of *do so* will be carried out by focus-dependent anaphoric processes such as Akmajian (1973) describes.

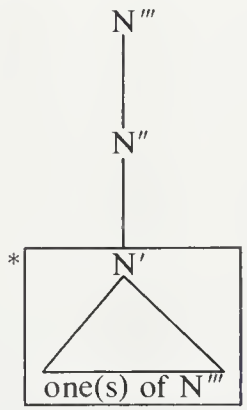
An N″ complement would automatically follow an N′ complement; the different positions would correspond to different semantic roles.

(4.3)



Based on this reasoning, the proper formulation of the difference between (4.2a,b) appears to be this: the pronoun *one* cannot be followed by the phrase *of NP* within the N′ complement. It can, however, be followed by other N′ complements, as will be seen shortly.[3] Thus, the inability of a particular *of NP* phrase to follow the pronoun *one* (where the antecedent is count) is a sufficient test for its being an N′ complement. (4.4) is the forbidden configuration.

(4.4)



This test requires some care. At first glance it would appear that *of water* must be an N″ complement in (4.5), since it follows *one*.

(4.5) Bill has two quarts of wine and one of water.

However, we notice that *one* can be pluralized in (4.2a) but not in (4.5).

(4.6) a. I met the ones from France.
b. *The quarts of wine and the ones of water were left behind.

[3] I am grateful to Noam Chomsky for pointing this out to me. See footnote 4.

The difference is explicated if we realize that there are two morphemes *one* that can function anaphorically in an NP, a numeral (a kind of Q, alternating with *two, three,* etc.—see section 5.5) and a pronoun. Only the latter, which takes the place of the head of the NP, can be pluralized to *ones*. (4.6b) shows that the *one* in (4.5) is the numeral one_Q, and that in fact the pro-N′ one_N cannot be followed by the complement *of water* here. We conclude therefore that *of water* is indeed an N′ complement in (4.5), and so, by parallelism, is *of wine*.

There are certain cases that appear to be ambiguous between N′ and N″ complements without appreciable difference in meaning. For example, because of the parallel between the direct object in (4.7a) and the PP in (4.7b), one would want to claim that the PP is an N′ complement.

(4.7) a. Bill pictured Fred.
b. Bill's picture of Fred

Yet *the pictures of Fred and the ones of Harry* is also acceptable, arguing that *of Fred* must be an N″ complement, similar to *the pictures which are of Fred.* The simplest solution is to accept both sources for such a case.

Another test to distinguish N′ from N″ complements is based on a distinction noticed by Lakoff (1970b):

(4.8) a. Fathers of few children have any fun.
b. *Fathers with few children have any fun.

Notice that *of few children* is, by our previous criterion, an N′ complement, since **ones of few children* is ungrammatical. But *with few children* can be paraphrased by the relative clause *who have few children,* suggesting that it is an N″ complement. This conjecture correctly predicts that only one order of the two complements is possible, i.e. *fathers of few sons with many daughters* but not **fathers with many daughters of few sons*. To describe the difference in grammaticality between (4.8a,b), then, one is led to the hypothesis that a quantifier may extend its scope out of an NP dominating it if it is in the N′ complement but not if it is in the N″ complement. Thus *any* could be in the scope of *few* in (4.8a) but not in (4.8b); since *any* must be within the scope of negation, (4.8b) is unacceptable. I can offer no reason why this distinction should exist, but it is confirmed by considering the possibility of *wh*-questions and relatives: according to the previous criterion, only N′ complements can be *wh*-ed.

(4.9) a. Fathers of which children had fun?
I met some children the fathers of whom like to drink.
b. *Fathers with which children had fun?
*I met some children the fathers with whom like to drink.

Another pair with this contrast is (4.10).

(4.10) a. Arguments with few people yield any satisfaction.
Arguments with which people satisfy you?
He is a person arguments with whom are fruitless.
b. ?*Arguments with few premises yield any satisfaction.
?*Arguments with which premises satisfy you?
?*This is a premise arguments with which are useless.

(4.10b) is all right if it is read *arguments* ***against*** *which premises,* parallel to *arguments with which people,* but impossible if read *arguments* ***employing*** *which premises.* It is plausible to assume that a person or premise against which an argument is directed is strictly subcategorized, i.e. an N′ complement, but that the instrument of argument, the premise employed, is an N″ complement like other instrument phrases. This is supported by the grammaticality of *arguments with Bill with few premises* and the ungrammaticality of **arguments with few premises with Bill.* Thus, this distinction between complement types supports the proposed interaction between level of complement and scope of quantifiers and *wh*; we see that inability of quantifiers or *wh* to extend scope outside the NP seems to be a sufficient condition for N″ rather than N′ complements.

Since there has been little study of a possible distinction among different kinds of NP complements, I am not familiar with enough strong cases to completely verify the validity of the two tests described above, much less to explain why these particular distinctions should obtain. However, the proposed distinction between N′ and N″ complements does seem to be fairly well borne out by the few examples cited here.[4]

The distinction between X″ complements and X‴ complements is somewhat more straightforward. In sentences, the V″ complements are the expressions of manner, means, accompaniment, instrument, purpose, and other so-called VP adverbials. Semantically, they map predicates into predicates of the same number of arguments, and they contribute to the main assertion of the sentence. As such, they can be focused, clefted, and affected by sentence negation:

(4.11) a. John hit the nail *softly*.
b. It was with a hammer that John hit the nail.
c. We didn't buy this for *your* benefit.[5]

Because they add extra truth conditions to the assertion of the sentence, restricting the extension of the sentence, V″ complements can be called *restrictive modifiers*.

V‴ complements, by contrast, add no conditions to the assertion of the sentence,

[4] Notice, by the way, that (4.10) provides evidence that the *ones* test is valid only for *of-NP* complements, not for other PPs in N′, since both of the following examples are acceptable.

(i) Arguments with Bill are less fruitful than ones with Harry. (N′ complement)
(ii) Arguments with many premises are less impressive than ones with few premises. (N″ complement)

[5] Just because they are in V″ and focused does not mean they are automatically affected by negation. Much depends on intonation. Cf. Jackendoff (1972, sections 6.7 and 8.6).

but rather add some sort of auxiliary assertion (one of whose arguments is usually the main assertion). They include sentence adverbials of all sorts, sentential appositives, parentheticals (e.g. *John is a fink,* I THINK), and various other subordinate clauses. When they occur at the end of a sentence, they are set off by comma intonation. Since they are not part of the main assertion of the sentence, they cannot be focused, clefted, or affected by sentence negation:

(4.12) a. *John hit the nail, *of course*.
b. *It was {probably / in my opinion} that John hit the nail.
c. *John didn't hit the nail, I *think*. (cf. John didn't hit the nail, I don't think.)

The geometry of the sentence predicts that V‴ complements must follow V″ complements, and this prediction is obviously borne out (except for some stylistic inversions).

(4.13) John hit the nail softly, of course.
?*John hit the nail, of course, softly.

?John ran away quickly, probably.
*John ran away, probably, quickly.

John hit the nail with a hammer, which surprised no one.
*John hit the nail, which surprised no one, with a hammer.

John hit the nail with a hammer, I think.
?John hit the nail, I think, with a hammer.[6]

Similar contrasts can be found in NPs. The clearest example is the difference between restrictive and nonrestrictive (appositive) relative clauses; the former follow previous modifiers without a break, may contain foci, and may be affected by sentence negation; the latter are separated by comma intonation, may not contain foci, and may not be affected by sentence negation.

(4.14) a. I didn't see the man who brought the *strawberries*.
b. *I didn't see the man, who brought the *strawberries*.

We would like to claim, therefore, that restrictive relative clauses are deep structure N″ complements and that appositives are deep structure N‴ complements. But since many proposals exist for deriving relative clauses from other sources, we will delay the defense of this claim to chapter 7, where more space can be devoted to it.

[6] The insertion of parentheticals into V″ may be what Banfield (1973) calls a *stylistic* transformation, one of a class of rules with somewhat different properties from ordinary transformations. Alternatively, Emonds (1976) gives an extensive treatment of parentheticals in which the complement is moved to the right of the parenthetical.

However, other constructions besides relative clauses can appear in N″ and N‴ complements. For example, the PPs in *the king from England, the weather at 6:00, the man with a big nose,* etc., are not in the N′ complement, follow the head without a break, and can be focused and affected by sentence negation.

(4.15) We didn't mention { the king from *England* / the weather at *6:00* / the man with a big *nose* }.

They are thus candidates for N″ complements. Descriptive adjectives, despite their prenominal position, have similar semantic properties to these PPs, arguing that they are attached to N″ (see section 7.4 for more discussion). This of course makes them parallel in structure and function to preverbal VP adverbs, which are in V″, consistent with the predictions of the $\bar{X}$ Convention.

By contrast, there are certain nonsentential appositives which can be characterized as N‴ complements:

(4.16) a. I will sell you these bagpipes, the finest in all Poland, for only 4000 zloty.
b. Perhaps you have heard of my brother, known the world over as a notorious womanizer.
c. She presented Picasso, then in his blue period, with a blueberry pie.

One might want to argue that all these are reduced nonrestrictive relative clauses, and this is certainly conceivable. In any event, they have the characteristic comma intonation and inability to be affected by sentence negation which we identify with an X‴ complement.

Again, the N‴ complements must follow the N″ complements, as predicted by the geometry of the NP:

(4.17) the man that brought the strawberries, who was dangerous
*the man, who was dangerous, that brought the strawberries

these bagpipes from Poland, the finest known
*these bagpipes, the finest known, from Poland

the man with a big nose, then in his blue period
*the man, then in his blue period, with a big nose (only ok as two appositives)

We see therefore that there are three distinct kinds of complements in both sentences and NPs, corresponding to the three levels predicted by the phrase structure rule schema. Furthermore, the parallels in structure and function across the two categories S and NP are clear and confirm the view of grammatical parallelism advocated in chapter 3.

In the remaining major categories, AP and PP, the complements are less produc-

tive, but the three types are distinguishable. Both adjectives and prepositions strictly subcategorize arguments, as was pointed out above. Most of the modifiers at the A″ and P″ level are part of the degree system, which precedes the head; they will be treated separately in chapter 8. At the A‴ and P‴ level, there are appositive clauses again, as usual marked by comma intonation.

(4.18) a. Martha is proud of her height, which you'll never be.
b. We went from Aspen to Denver, which seems like a long way, in less than four weeks.

As usual, the order of the appositive and the strictly subcategorized arguments cannot be reversed:

(4.19) a. *proud, which you'll never be, of her height
b. *from, which seems like a long way, Aspen to Denver
*from, Aspen to, which seems like a long way, Denver

Thus the facts of APs and PPs are consistent with the Three-Level Hypothesis, although they do not push it to its limits as do Ss and NPs.

This section has shown that the three levels of complements made possible by the theory of phrase structure of chapter 3 can be rather clearly distinguished, both syntactically, through ordering and anaphoric behavior, and semantically, through the means of integration into the interpretation. We have seen that there is strong evidence for all three levels in S and NP, and for at least two levels in AP and PP.

The rest of this chapter will be devoted to the details of X′ and X″ complements, concerning itself in particular with what cross-category generalizations appear in their phrase structure. We begin with V′ and N′, the most elaborate and best-known structures, then move to V″ and N″ complements in section 4.3. Section 4.4 deals with the complements of A, Adv, and P, and section 4.5 discusses how the rules generalize. The appendices deal with further problems of NP and S complements.

4.2. The V′ and N′ Complements

This section will be concerned with showing how a number of important syntactic differences between Ss and NPs stem from differences in the phrase structure rules for V′ and N′.

4.2.1. Easy Parts of V′

A first approximation to the phrase structure rule for V′ is (4.20).

(4.20) $V' \rightarrow V - (NP) - (Prt) - (\left\{\begin{matrix} NP \\ AP \end{matrix}\right\}) - (\left\{\begin{matrix} AdvP \\ QP \end{matrix}\right\}) - (PP) - (PP) - (\bar{S})$

The verb and direct object are obvious. Skipping the next three positions for a moment,

the two PPs are to provide for the strictly subcategorized arguments in constructions like (4.21a). To see the need for both PPs, note that neither PP can follow *do so* (4.21b,c), although (4.21d) shows that the verb *talk* can in other cases serve as the antecedent of *do so*. We conclude that both PPs in (4.21a) are in V′.

(4.21) a. John talked to Bill about Harry.
b. *John talked to Bill about Harry, but he didn't do so about Fred.
c. *John talked about Harry to Bill, but he didn't do so to Fred.
d. John talked to Bill about Harry on Sunday, but he didn't do so on Thursday.

Return now to the $\begin{Bmatrix}\text{AdvP}\\\text{QP}\end{Bmatrix}$ position. It is needed for the strictly subcategorized phrases in constructions like these:

(4.22) a. The job paid (us) $\begin{Bmatrix}\text{handsomely}\\\text{too little}\end{Bmatrix}$.
b. Bill worded the letter carefully.

Note that the AdvPs and QP in (4.22) cannot be omitted without incurring ungrammaticality, and that, unlike most such phrases (cf. (4.23c)), they cannot follow *do so*.

(4.23) a. ?*The job on Tuesday paid us $\begin{Bmatrix}\text{handsomely}\\\text{too little}\end{Bmatrix}$, but the one on Wednesday did so $\begin{Bmatrix}\text{very poorly}\\\text{enough}\end{Bmatrix}$.
b. ?*John worded his speech carefully, but Fred did so carelessly.
c. John talked about Martha $\begin{Bmatrix}\text{thoughtfully}\\\text{quite enough}\end{Bmatrix}$, but Fred did so $\begin{Bmatrix}\text{inconsiderately}\\\text{far too much}\end{Bmatrix}$.

Thus the adverbs in (4.22) pass the tests for V′ complements.

The $\bar{S}$ in (4.20) represents the subordinate clause position in sentences like *No one would claim that Ss are NPs*. There is no doubt that such an $\bar{S}$ is strictly subcategorized; it cannot be omitted, and it cannot follow *do so*. Furthermore, it is little remarked that although the $\bar{S}$ must follow all other strictly subcategorized phrases, it may be followed by manner and degree phrases.

(4.24) a. He claimed that Ss are NPs in a loud strident voice.
b. He believes that Ss are NPs quite fervently.

We will suppose, therefore, that these Ss are daughters of V′ in deep structure, and

that their more usual surface position at the end of V″ is the result of an extraposition transformation, perhaps the same one that extraposes heavy NP direct objects.[7]

4.2.2. Particles and Predicates

Returning to the two positions after the direct object, we first find Particle.[8] Emonds (1972a) argues that in verb–particle constructions such as (4.25), the separated form (4.25a) is underlying and the form with the particle next to the verb (4.25b) is derived.

(4.25) a. Howard looked the answer up.
b. Howard looked up the answer.

The principal objection to Emonds's position is that *look up* is a lexical item and thus, in the standard theory, must be inserted under a single constituent in deep structure; (4.25b) but not (4.25a) can fulfill this requirement on deep structure. However, Emonds points out idioms like *take NP to task* and *take NP for a ride* ('dupe NP'), which (like (4.25a)) are discontinuous and which therefore provide independent evidence against monocategorial lexical insertion. This vitiates the argument against (4.25a) being an underlying form.

Following the particle is what Emonds has called the Predicate position. This constituent appears in sentences like (4.26).

(4.26) We elected Harrison chairman.
We named the baby Hortensia.
We painted the barn red.
You make me sick.

It has been argued from time to time that these sentences have a deleted complement *to be*. However, such sources as **We named the baby to be Hortensia, *You make me (to) be sick,* the alleged deep structures, are implausible, since *elect, name,* and *paint* (at least) never appear at all with an infinitive complement. Thus, given the choice of adding a transformation or extending the base to account for these verbs, the latter solution appears better motivated. In fact, we can then claim that the copula involves the Predicate position in *John is a boy* and *John is sick,* capturing the semantic generalization between (4.26) and the copula directly in the formulation of the base and

[7] As ever so slightly suggestive evidence for this extraposition, notice that no intonation breaks are possible in (4.24), but pauses may be inserted in the extraposed forms:

(i) He claimed(,) in a loud strident voice(,) that Ss are NPs.
(ii) He believes(,) quite fervently(,) that Ss are NPs.

And even when the pauses are not inserted, the intonation in these forms is somewhat more "leisurely". Since such pauses are characteristic of deletion sites or traces (cf. Selkirk (1972)), (i)–(ii) must be the derived forms, not (4.24).

[8] According to the phrase structure rule schema, this should probably be Prt‴, with no complements or specifiers in the expansion of Prt‴ down to Prt. Since universal grammar provides this much of the phrase structure rules, using Prt‴ instead of Prt in (4.20) incurs no extra cost to the particular grammar of English. We will make the revision explicit in section 4.4.

projection rules, without recourse to a transformation. This makes *be* a verb like any other in its predicative uses, avoiding the extra phrase structure rule of *Syntactic Structures* and *Aspects*.

Emonds (1970) claims that the Predicate position is dominated by a node Pred which is in turn dominated by VP. Such an analysis would violate our theory of categories, of course. I believe, however, that postulating this node is as much a mistake as postulating a node Agent. Rather, Predicate, like Agent, is a semantic relation which may be assigned by the projection rule for the NP/AP position in the VP.[9] As syntactic evidence that the second NP node is not distinguished by a dominating node Pred, we offer a few expressions which use the second NP position for other than a predicate, and which do not appear to be derived by any sort of dative shift (these examples were suggested to me by Bowers (1971)):

(4.27) a. Margaret struck { Joe a heavy blow / *a heavy blow to Joe }.

b. I envy { you your beauty / *your beauty to you }.

c. He called { her names / names to her (wrong meaning) }.

We need some justification for placing the Prt position before the Predicate. One piece of evidence comes from sentences like (4.28), with a predicate AP.

(4.28) a. ?*They painted up the barn red.
b. They painted the barn up red.
c. *They painted the barn red up.

Clearly the best position for *up* is between the direct object and the predicate. I know of no cases which contain a direct object, a particle, and an NP predicate; but *John grew up a Catholic* is suggestive, in that the particle cannot be interchanged with the predicate as it can with direct objects.[10]

[9] Jackendoff (1976), following Gruber (1965), characterizes this semantic relation as Identificational Location, showing how it fits into a general account of functional semantic interpretation.

[10] In Jackendoff (1974a) it was claimed that *a Catholic* is not strictly subcategorized in this example. However, it cannot follow *do so* as can the similar construction to which it was compared there:

(i) *John grew up a Catholic and Bill did so a Moslem.

(ii) John walked into the meeting { drunk / a saint / having forgotten who he was }, but Bill did so { sober / a sinner / in full knowledge of the hazards }.

Though the NP case in (ii) is somewhat worse than the others, it seems better than (i). The evidence that Prt precedes predicate NP is admittedly inconclusive; but there is no evidence at all against it. This is a case where we can let the theory decide.

Another bit of evidence for the base configuration V–NP–Prt–NP is that it permits a simple structure-preserving analysis of Particle Movement and Dative Shift. Notice first that if there is a single object, it can be generated in either NP position, i.e. on either side of the particle. Hence the difference between (4.25a) and (4.25b) need not be the application of a transformation at all: it may be due to the deep structure option as to object position.

Next, observe that when there is a direct object, a particle, and a PP in the complement, the object may as usual be on either side of the particle, as in (4.29). But, as Emonds (1972a) points out, in a double-object complement, the particle must fall between the two objects, as seen in (4.30).

(4.29) a. The secretary sent a schedule out to the stockholders.
b. The secretary sent out a schedule to the stockholders.

(4.30) a. The secretary sent the stockholders out a schedule.
b. ?*The secretary sent out the stockholders a schedule.
c. *The secretary sent the stockholders a schedule out.

This constraint can be accounted for if Dative Shift is a structure-preserving rule which moves the object of *to* into the NP position directly after the verb (lower case denotes an empty node):

(4.31) *Dative Shift*[11]
$X - V_a - np - (Prt) - NP - [_{PP}to - NP] - Y \Rightarrow$
$1 - 2 - 7 - 4 - 5 - \phi - \phi - 8$
OPTIONAL

According to this statement of Dative Shift, the structural description of the rule can be met only if the first NP position in V′ is empty, that is, if the direct object is generated in the position following Prt. Thus, the configuration (4.30a) is the only possible ordering of the two objects and the particle under this analysis. In turn, the analysis justifies the base configuration V – NP – Prt – NP in (4.20).

Finally, there are some possible improvements in the statement of (4.20), making

[11] In this rule V_a denotes the class of verbs that govern Dative Shift.

Emonds (1972a) describes these facts with a Dative Shift that actually interchanges the two objects. The present analysis is better in that it does not violate the generalization that transformations move only one constituent at a time.

(4.31) is however sufficient to describe only Emonds's Dialect B, which happens to be mine. For the other two more complex dialects he describes, a movement of Prt to postverbal position will still be necessary. Since Prt is not a phrase node, this rule does not fall under the structure-preserving constraint, but rather under Emonds's notion of a "local transformation".

An interesting feature, perhaps an advantage, of this analysis is that it permits Dative Shift to be governed by deep structure strict subcategorization rather than the traditional rule feature. Verbs which permit Dative Shift strictly subcategorize an optional empty second NP, permitting the structural description of (4.31) to be met. Verbs which do not allow Dative Shift will omit the Predicate NP in their strict subcategorization. Thus, in fact, V_a can be replaced simply by V in (4.31), if it is the case that all verbs which allow Dative Shift also allow Particle Movement in the presence of a particle.

use of feature notation. The predicate position, $\begin{Bmatrix}NP\\AP\end{Bmatrix}$, can be rewritten as [−Obj, +Comp]‴; the $\begin{Bmatrix}AdvP\\QP\end{Bmatrix}$ position can be rewritten as [−Obj, −Comp, −Det]‴. Since it appears that strictly subcategorized adverbs and QPs as in *The job paid handsomely/enough* never cooccur with the predicate position, these two positions can be collapsed into one (syntactically, though not semantically) with the feature complex [−Obj, −Det]‴. Similarly, there never seems to be a need for two PPs and an $\bar{S}$ all at once. If this is the case, the last two terms in (4.20) can be collapsed into [+Obj, +Comp]‴. Thus, the final form of the base rule, using feature notation where necessary, is (4.32).

$$(4.32)\quad V' \rightarrow V - (NP) - (Prt) - \left(\begin{bmatrix}-Obj\\-Det\end{bmatrix}'''\right) - (PP) - \left(\begin{bmatrix}+Obj\\+Comp\end{bmatrix}'''\right)$$

4.2.3. The N′ Complement

This subsection will show that certain differences between the phrase structure of V′ and N′ account for a number of further divergences in their syntax. The expansion of N′ is at most (4.33).

$$(4.33)\quad N' \rightarrow N - (Prt) - (NP) - (PP) - \left(\begin{Bmatrix}PP\\\bar{S}\end{Bmatrix}\right)$$

Again skipping the Prt and NP for just a moment, we see that two PP positions are necessary for phrases like *our talk about the war with Bill,* paralleling the two PPs in the V′ of *We talked about the war with Bill.* The $\bar{S}$ in (4.33) appears in such constructions as *Frank's claim (to Tony) that Sim was the culprit,* where the $\bar{S}$ corresponds to the one in the parallel V′ of *Frank claimed (to Tony) that Sim was the culprit.*[12] There never seems to be a need for two PPs and an $\bar{S}$ at once, so I have collapsed the second PP with the $\bar{S}$; in feature notation the combination would be identical to the one at the end of (4.32). Returning to the Prt, in the rare cases when there is a particle in N′, its position is always immediately after the head, e.g. *Bill's looking up of the information,* not **Bill's looking of the information up.*

There remains the NP position. In surface structure, of course, N cannot be followed by an NP object or predicate. An N′ corresponding to a V′ with a direct object generally contains *of* (although there are exceptions such as *doubt*).

[12] Note, by the way, that the $\bar{S}$ at the end of N′ confirms Emonds's (1970) hypothesis that *that* and *for-to* complements may be generated directly under V′, rather than being extraposed there from under NP, as proposed by Rosenbaum (1967). The Lexicalist Hypothesis claims that *Frank claimed that Sim was the culprit* and *Frank's claim that Sim was the culprit* must have parallel structures. In Rosenbaum's complement system, the deep structures are [$_S$Frank past [$_{VP}$claim [$_{NP}$it [$_{\bar{S}}$that Sim was the culprit]]]] and [$_{NP}$Frank claim [$_{\bar{S}}$that Sim was the culprit]], clearly nonparallel structures. To solve this difficulty, one must assume either that both Ss are in [$_{NP}$it $\bar{S}$] constructions or that they are both attached to X′. As Emonds's arguments motivate the latter configuration, it is the structure of V′, not of N′, that must be changed from Rosenbaum's system.

(4.34) destroy the city/destruction of the city
explain the result/explanation of the result
establish a claim/establishment of a claim
etc.

Verbs which take PP complements generally preserve the preposition in the derived nominal:

(4.35) persist in being stubborn/persistence in being stubborn
pray for war/prayer for war
escape from despair/escape from despair
remind him of his appointment/reminder of his appointment
lust for power/lust for power
approve of your behavior/approval of your behavior

Since in these cases the form of the complement is preserved in derived nominals, Chomsky (1970a) suggests (p. 41–42/204 and note 27, p. 42/note 28, p. 219) that the same is true of direct objects in underlying form: they occur in both V′ and N′ without the preposition, and an obligatory transformation inserts *of* to produce the surface form.[13] Note that this *of* cannot be the preposition *of*, since there is no PP structure into which it is inserted. Rather, it must be a specified grammatical formative like *poss*. Another instance of the specified grammatical formative *of* will appear in section 5.4.

The transformation inserting *of* is therefore to be stated as (4.36).

(4.36) *Of*-Insertion
$[_{N'}\text{N} - (\text{Prt}) - \text{NP} - \text{X}] \Rightarrow 1 - 2 - \textit{of} + 3 - 4$
OBLIGATORY

[13] Note that *of* is the only formative inserted. In *his doubts about the proposal, about* will be present in deep structure, and *about the proposal* is a PP. The fact that the verb and noun do not agree in their complement structure must be entered in the lexicon as a complication of their normal verb–noun pairing. Chomsky's footnote cited above defends the *of*-insertion as opposed to an *of*-deletion in VP. A further argument appears in the discussion of the passive, section 4.7.

Jackendoff (1969c) argues that possessive constructions like *a picture of John's* are a kind of partitive, analogous to *one of John's pictures*. If this is the case, the *of* is inserted by (4.36). Since *Of*-Insertion applies only to the NP immediately after the head noun, the genitive phrase must be the leftmost element. This explains Chomsky's judgment (p. 46/206) of the "quite clumsy phrase *the portrait of Aristotle by Rembrandt of the Metropolitan Museum's*", since there is no way for the grammar to generate it. If *by Rembrandt* is omitted, the ungrammaticality is even clearer. Also, note the contrast of *a portrait of the Met's by Rembrandt* and ?**a portrait by Rembrandt of the Met's*.

Jackendoff (1974a) takes a somewhat different view of the inserted *of*: it is not a specified grammatical formative, but a real preposition which fills an empty head of a PP complement. This analysis was necessary in part because the notion of specified grammatical formative was not included in the theory of phrase structure assumed there. The present analysis considerably simplifies the generalization of the grammatical relation *object*, and is much closer to Chomsky's account.

To distinguish between these two theories, one would ideally like to use extraction to test the constituency of *of*-NP. But the results are inconsistent. *Wh*-Preposing cannot carry the *of* along: **Of which boys did you meet the fathers*? But, as will be seen in section 5.3.1, Extraposition from NP does move *of*-NP, arguing against the present view. On the other hand, Akmajian and Lehrer (1976) show that certain apparent extractions of *of*-NP must be base-generated rather than extracted, so perhaps Extraposition from NP is such a case too. I leave the question open.

Thus, the underlying form of the nominals in (4.34) will be *destruction the city, explanation the result, establishment the claim.*

The absence of an NP position preceding Prt in (4.33) explains why Dative Shift and Particle Movement do not occur in N′. We have analyzed Particle Movement in V′ as the choice between two underlying phrase structure positions for the direct object, one on either side of the particle. Since N′ contains only the NP position after the particle, the other ordering is impossible and the object can only follow Prt. Similarly, Dative Shift in V′ requires a structure-preserving movement of the indirect object into the NP position before Prt. Since N′ contains no such position, Dative Shift cannot apply.

Furthermore, the absence of Predicate complements like **the election of John (of) President* follows from the fact that there is only one bare NP position in N′: a construction parallel to the V′ *elect John President* would require two NP positions in N′ and there is only one. The closest parallel uses a PP with *as*: *the election of John as President.* Thus, our analysis of the phrase structure of N′ gives a principled explanation of three apparently unrelated differences between V′ and N′ on the basis of the N′ lacking the first NP position.

Having justified base rule (4.33), we now address ourselves to how it generalizes with V′.

4.2.4. Generalization of V′ and N′ and the Definition of Object

Here are the two phrase structure rules we wish to relate.

$$(4.32)\quad V' \rightarrow V - (NP) - (Prt) - \left(\begin{bmatrix} -Obj \\ -Det \end{bmatrix}'''\right) - (PP) - \left(\begin{bmatrix} +Obj \\ +Comp \end{bmatrix}'''\right)$$

$$(4.33)\quad N' \rightarrow N - \qquad (Prt) - \quad (NP) \qquad - (PP) - \left(\begin{bmatrix} +Obj \\ +Comp \end{bmatrix}'''\right)$$

All positions are parallel, except for two: N′ lacks the first NP position, and allows only an NP where V′ allows NP, AP, QP, or AdvP. Using subscripted angle brackets, the two rules can be collapsed as (4.37), which takes only three features more than the rule for V′ alone.[14]

$$(4.37)\quad \begin{bmatrix} X \\ +Subj \\ \langle +Obj \rangle_1 \\ \langle -Obj \rangle_2 \\ +Comp \end{bmatrix}' \rightarrow X - \langle (NP) \rangle_1 - (Prt''') - \left(\begin{bmatrix} \langle +Subj \rangle \\ \langle +Comp \rangle_2 \\ -Obj \\ -Det \end{bmatrix}'''\right) - (PP) - \left(\begin{bmatrix} +Obj \\ +Comp \end{bmatrix}'''\right)$$

We would also like to develop a definition of the grammatical relation "object-of". In terms of the *Aspects* notation for grammatical relations, the generalized object

[14] We change Prt to Prt‴ as suggested in note 8, page 66; this will be relevant in section 4.4.

relationship should be stated as [N‴, [+Subj]′], that is, in a string exhaustively dominated by [+Subj]′, the object is a substring exhaustively dominated by N‴, and that occurrence of N‴ is directly dominated by the [+Subj]′. While this would give the correct results in N′, it would incorrectly pick out both the object and the Predicate in the V′ *elect John President*.

It is not clear to me whether this is a serious problem. But if it is, the definition of object can be sharpened to include order as well as domination relations:

(4.38) In $\left[\begin{matrix} X \\ +\text{Subj} \end{matrix}\right]'$, the object is the N‴ immediately dominated by X′ in the configuration [$_{X'}$X – (Prt) – N‴ – Y].

Note that this definition allows an NP on either side of Prt to be designated as object, as long as there is no nonempty NP between it and the head.

For completeness, we should also define the grammatical relation Predicate.

(4.39) In V′, the Predicate is the italicized phrase immediately dominated by V′ in the configuration [$_{V'}$V – (N‴) – (Prt) – $\left\{\begin{matrix} N''' \\ A''' \end{matrix}\right\}$ – Y].

Some Predicates will of course be structurally indistinguishable from some direct objects, for example, *Bill became a butterfly* (predicate) vs. *Bill saw a butterfly* (object). But since the choice is determined by the verb, no structural difference is needed. The same two interpretations are possible within a PP, as in such a minimal pair as *The coach turned into a driveway* vs. *The coach turned into a pumpkin*. Again the possible interpretations of *turn* can determine which way to read the NP without a structural difference. Hence, it should not be disturbing that (4.39) defines a set of positions which overlaps partially with the set of positions defined by (4.38). In a particular reading of a particular sentence, only one grammatical relation will be applicable to an NP in one of the ambiguous positions, and that grammatical relation will be established by the verb.

This completes our discussion of the phrase structure of V′ and N′. We have seen that they are parallel, except for one crucial position which N′ lacks. The missing NP position has been used to explain a number of further differences. Section 4.7 is an appendix which discusses some movement rules in and out of V′ and N′. We now go up one level to V″ and N″.

4.3. The V″ and N″ Complements

As pointed out in section 4.1, the X″ complement contains restrictive modifiers. In V″ these are AdvPs and PPs of manner, means, time, instrument, and so forth, as well as various adverbial clauses. In N″, the complements include PPs of time, place, accompaniment, and so forth, relative clauses, and (for semantic reasons, their prehead

position notwithstanding) APs. The phrase structure of X″ is on the whole rather simple; this section will discuss three intricacies involved in the proper formulation of X″ phrase structure rules and in the use of X″ positions.

First let us deal with the differences between AdvPs and PPs in V″. The most salient difference is that AdvPs may appear preverbally as well as postverbally, whereas PPs may only be postverbal.

(4.40) a. Bill dropped the bananas {quickly / with a crash}.

b. Bill {quickly / *with a crash} dropped the bananas.

Furthermore, although PPs can appear one after another, there is a stylistic preference for avoiding two consecutive adverbs, wherever they are attached (see Jackendoff (1972, chapter 3) for discussion of adverb attachment to V″ vs. V‴).

(4.41) a. ?Bill frankly cleverly left. (two S adverbs)
b. ?Bill probably suddenly left. (S and VP adverbs)
c. ?Bill quickly completely finished. (two VP adverbs)
d. ?Frankly, probably Bill left. (two S adverbs)
e. ?Bill finished completely, probably. (VP and S adverbs)
f. ?Bill finished completely quickly. (two VP adverbs)

Note that if the adverbs are separated the examples in (4.41) become acceptable.

(4.42) a. Frankly, Bill cleverly left.
b. Probably, Bill suddenly left.
Bill suddenly left, probably.
Bill probably left suddenly.
c. Bill quickly finished completely.
d. Frankly, Bill probably left.
Frankly, Bill left, probably.
e. Probably, Bill finished completely.
Bill completely finished, probably.
f. Bill quickly finished completely.

Since this condition can involve any type of adverb, and in particular can cross constituent boundaries, it cannot be stated as part of a set of context-free phrase structure rules, but must rather be stated as some sort of an output string condition. Thus the phrase structure component may be left free to generate an indefinite number of adverbs as well as an indefinite number of PPs, and the number actually possible will be constrained by the output condition and the necessity that they all receive interpretations which can be integrated into the interpretation of the sentence.

These considerations suggest, for a first approximation, the phrase structure rule (4.43), where X* denotes a concatenation of an indefinite number of Xs.

(4.43) $V'' \rightarrow$ (have – en) – (be – ing) – (AdvP) – V′ – $(\begin{Bmatrix}PP\\AdvP\end{Bmatrix})^*$

There are two things wrong with (4.43). First, the final term, expressed in distinctive features, is [−Subj, αObj, αComp], not a wonderfully natural combination. Furthermore, in the N″ complement there is no AdvP, and no juggling of features will be able to combine [αObj, αComp] with [+Obj, +Comp] in a particularly felicitous fashion. This suggests that the combination $\begin{Bmatrix}PP\\AdvP\end{Bmatrix}$ is incorrect, though it appears necessary because PPs and AdvPs may be freely interspersed in V″.

The second difficulty is the repetition of AdvP in two positions in V″, though it has essentially the same semantic functions in either position. Within the usual framework of phrase structure grammar, this cannot be avoided, since the V′ gets in the way of collapsing the two occurrences. Moreover, the manner adverb may occur before an aspect that has not been moved up into the Aux, as in *John will completely have finished.* Thus in fact *three* AdvP positions appear necessary.

The usual solution to this problem is to generate the desired constituent in a single position in underlying structure, then to move it to its various surface positions by means of a transformation. Here, however, the nature of the generalization lends itself to a different treatment: the generalization is that AdvP may appear in any position where it is a daughter of V″. To express this generalization, we will follow the suggestion of Keyser (1968) and allow the phrase structure component to designate certain constituents as "transportable": a transportable node will have free word order with respect to all of its sisters.[15] We will indicate transportability with a feature [+Trans], and assume that [−Trans] is unmarked, that is, that if the feature is not mentioned in a rule, [−Trans] is intended (or filled in by marking conventions). We can now revise (4.43) to (4.44).

(4.44) $V'' \rightarrow$ (have – en) – (be – ing) – $(\begin{bmatrix}Adv\\+Trans\end{bmatrix}''')^*$ – V′ – (PP)* – ($\bar{S}$)

The position of the Adv in (4.44) allows collapsing with the rule for N″, (4.45), to the rule (4.46).

(4.45) $N'' \rightarrow (A''')^*$ – N′ – (PP)* – ($\bar{S}$)

(4.46) $\begin{bmatrix}X\\+Subj\\\langle+Obj\rangle_1\\\langle-Obj\rangle_2\\+Comp\end{bmatrix}'' \rightarrow \langle$(have – en) – (be – ing) $\rangle_1$ – $(\begin{bmatrix}-Subj\\-Obj\\\langle\{-Comp\}\rangle_1\\\{+Trans\}\\\langle+Comp\rangle_2\end{bmatrix}''')^*$ – X′ – (PP)* – ($\bar{S}$)

[15] Jackendoff (1972, chapter 3) argues that free adverb order in VP is impossible because strictly subcategorized adverbs must be postverbal. In the present treatment, the placement of strictly subcategorized adverbs in V′ separates them syntactically from other adverbs, which can now be free to wander throughout V″.

Note that the feature *+Trans* appears only in V″, so that adjectives will appear only prenominally and will not be free within N″.

A second problem in the structure of V″ is the ability of a strictly subcategorized PP (i.e. a PP in V′) to appear after a manner adverb, as in *John gave the beans quickly to Bill*. To account for this, either the adverb must be lowered into the AdvP position in V′, or the PP must be raised into the first PP position in V″. Both operations are structure preserving and in that respect harmless.

One piece of evidence that might be thought relevant to the choice is that a strictly subcategorized PP cannot follow *do so* even when it follows a manner adverb:

(4.47) *John gave the beans quickly to Bill, and Fred did so (slowly) to Susan.

This apparently argues that the adverb is moved down into V′. But in fact such an argument is incorrect; it is also the case that strictly subcategorized clauses cannot follow *do so* even when extraposed around various PPs in V″.

(4.48) *John said {in a loud voice / suddenly / at 6:00} that smoking was fun, but Susan did so ({softly / in jest / at 5:00}) that it was bad for you.

Thus there must be an independent mechanism to rule out (4.48), possibly having to do with the trace left behind in V′ by the extraposed constituent; whatever this mechanism is, it will also rule out (4.47) under the theory that the PP *to Susan* has been extraposed. Thus there are no strong considerations to decide which constituent moves in inverting a manner adverb with a strictly subcategorized PP. The spirit of the Extended Standard Theory does include a general uneasiness with lowering rules, so perhaps the raising of the PP is a better solution.

Finally, there is an $\bar{S}$ at the end of both V″ and N″. This is the position of restrictive relative clauses in N″ and of comparative clauses in both V″ and N″; chapters 7 and 8 will discuss these uses at some length. It is also the position for various restrictive subordinate clauses in V″ such as in *It only hurts when I laugh*, and the position to which extraposed V′ sentential complements move by a structure-preserving rule, for example in (4.48). A more controversial use is as the position for sentence-final subject complements, as in *It bothers Bill that she came*.

The controversy arises over whether the V″ complement is only a derived position for these complements, as claimed by Rosenbaum (1967), or whether it is an underlying position, as claimed by Emonds (1970). Emonds shows that there are a vast number of differences between *that* and *for-to* complements on the one hand and gerundive (*poss-ing*) complements on the other, concluding that only gerundives behave like NPs. Rosenbaum's theory, in which all three types of complements are dominated by NP in

underlying structure, is at a loss to explain this fact without a great number of unusual conditions on the Extraposition transformation, which derives (4.49b) from (4.49a).

(4.49) a. [$_{NP}$it [$_{S}$that she came]] bothers Bill
b. It bothers Bill that she came.
c. That she came bothers Bill.

Emonds, on the other hand, shows that the syntactic distribution of complement types is much better accounted for if (4.49b) is underlying and (4.49c), which occurs only in a rather restricted class of contexts, is derived from it by a simply stated root transformation of Subject Replacement (or Intraposition).

On general syntactic grounds, the Intraposition theory has much to recommend it. However, a number of difficulties in this theory are pointed out in Higgins (1973), and in response to them Emonds (1976) abandons the Intraposition theory in favor of a variant of Rosenbaum's analysis. The advantageous properties of Emonds's earlier analysis are accounted for in the new analysis, but only in a somewhat artificial and suspect fashion. The issue is not crucial to the present theory, but Emonds's original view of these complements does fit attractively into our analysis here. Section 4.8 presents some new considerations which favor the Intraposition analysis, and which show that V″ is the correct position for subject complements to be attached.

The V″ position of subject complements has a parallel in N″. Chomsky (1970a) points out (p. 44/203) that sentences such as (4.50a) are paralleled by nominals like (4.50b).

(4.50) a. It is necessary for John to leave.
b. the necessity for John to leave

Because Chomsky assumes the complement clause in the sentence is an extraposed subject, the $\bar{X}$ Convention forces him to generate the complement of the nominal before the head too. He then suggests that Agent Postposing (the first part of Passive) accomplishes the extraposition of the clause in (4.50b) to its surface position. The assumption that (4.50a) is in its underlying order, as Emonds (1970) suggests, eliminates the dubious extraposition in the nominal, since the $\bar{X}$ Convention then requires (4.50b) to be in its underlying order as well. Thus if the $\bar{S}$ in V″ is a subject complement position, so is the $\bar{S}$ in N″.

4.4. Complements of [−Subj] Categories

4.4.1. Adjectives

The syntax of A′ is similar to that of N′, but more limited. Adjectives can take PP complements, as in *good at chess, yellow with age, eager for help.* They can also take $\bar{S}$ complements, as in *afraid that Bill is fierce, eager to please, proud to be a frog,* and *happy that he won.* There are a few examples such as *dependent on Bill for help* which

have two PPs, and *dependent on Bill to help him* with PP followed by $\bar{S}$. Thus the phrase structure rule for A′ includes at least (PP) – $(\left\{\begin{matrix}\text{PP}\\ \bar{S}\end{matrix}\right\})$.

A significant number of adjectives related to transitive verbs take *of*-NP in their complements: *fearful/considerate/desirous/solicitous of NP,* for example, are related to *fear/consider/desire/solicit NP.* To simplify the statement of lexical relations, we can consider these particular *of*-NP complements to be simple direct objects in deep structure, treating the *of* as a specified grammatical formative which is inserted transformationally. Thus, we can generalize the definition (4.38) of the grammatical relation *object* to adjectives; and rule (4.36), *Of*-Insertion, generalizes to the domain [−Obj]′, saving a feature in its statement and simultaneously simplifying lexical relations. Thus the expansion of A′ must be something like (4.51).

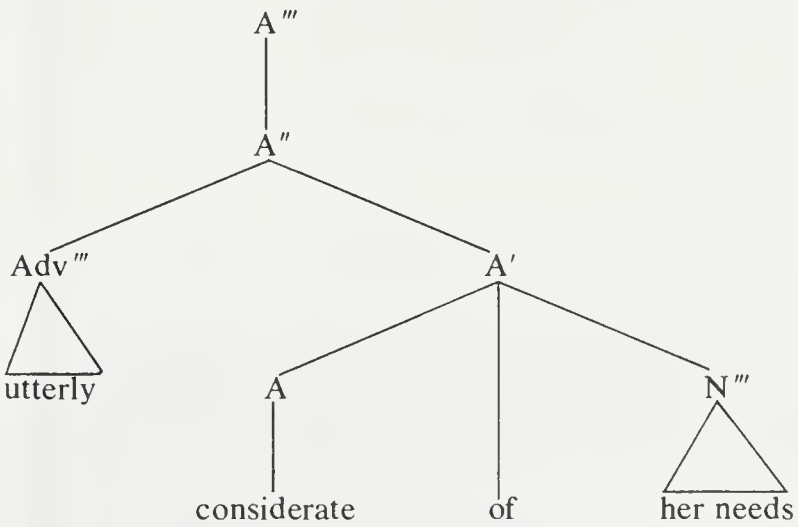
(4.51) $\quad A' \rightarrow A - (NP) - (PP) - (\left[\begin{matrix}+\text{Obj}\\ +\text{Comp}\end{matrix}\right]''')$

In the A″ complement there is a class of preadjectival modifiers, for example *completely/utterly/partially corrupt.* These are structurally parallel to the adverbial degree modifiers in Ss, as in *John has been completely/utterly/partially corrupting the government,* and to prenominal adjectives as well, so base rule (4.52) will be a special case of the more general rule for this position. Since the adverb can in turn take a few modifiers (particularly *not*, as in *a not utterly divine concert*), the adverb position must be Adv‴ and not just Adv. Further, A″ contains an $\bar{S}$ for degree clauses, to be discussed in chapter 8.

(4.52) $\quad A'' \rightarrow (Adv''') - A' - (\bar{S})$

(4.53) is an AP which illustrates some details of the structures just proposed.

(4.53)



4.4.2. Adverbs

On the whole, adverbs take no complement: there is no **fearfully of Bill* parallel to *fearful of Bill*, for example. However, there are a few PP complements to adverbs, for example *unfortunately for our hero*. The question arises as to whether they are attached to Adv′ or to Adv″. We would like to maintain the generalization that [−Comp] categories do not strictly subcategorize anything, i.e. that they have no X′ complements. This would require that *for our hero* is an Adv″ complement.

Evidence within the adverb phrase itself is hard to come by, but the parallel AP constructions are suggestive. Notice that in (4.54a), the PP *for our hero* can prepose (or alternate with a V″ complement which preposes) to form (4.54b). But a PP strictly subcategorized by AP, as in (4.54c), cannot prepose.

(4.54) a. It was unfortunate for our hero that Rome burned.
b. For our hero, it was unfortunate that Rome burned.
c. Bill is dependent on Fred.
d. ?*On Fred(,) Bill is dependent. (ok only with "Yiddish" topicalization)

This difference may be ascribed to attachment of the two PPs in different places,[16] arguing that *for our hero* in (4.54a) and in the parallel AdvP is an X″ complement.

As for $\bar{S}$ complements, we will take the position that subordinating conjunctions are prepositions rather than adverbs, so "adverbial clauses" introduced by *because, although*, etc. are not Adv′ complements. There appear to be no *-ly* adverbs with $\bar{S}$ complements. Thus, the phrase structure rule for Adv′ is the minimal (4.55), and we can maintain the claim that [−Comp] categories take no X′ complements.

(4.55) $\text{Adv}' \rightarrow \text{Adv}$

Like adjectives, adverbs may be preceded by *utterly, completely*, etc.; if such degree adverbs are dominated by A″, they must also be within Adv″. Again the presence of a few modifiers such as *not* and *quite* (as in *The state is governed quite utterly corruptly*, infelicitous primarily because of the repetition of *-ly*) shows that the degree modifier is an Adv‴ rather than just an Adv. Adv″ also contains degree clauses, as well as the PPs just mentioned. Hence the base rule for Adv″ is (4.56); (4.52) can also be revised to include the PP.

(4.56) $\text{Adv}'' \rightarrow (\text{Adv}''') - \text{Adv}' - (\text{PP}) - (\bar{S})$

4.4.3. Prepositions

Although the usual uninformed view is that the complement of a preposition is invariably NP, it is shown in Jackendoff (1973), using in part evidence from Klima (1965) and Emonds (1972a), that the base rule for P′ is at least (4.57).

[16] Williams (1975) gives this sort of difference as criterial for preposing in Ss.

(4.57) $P' \rightarrow P - (NP) - (PP)$

If both optional constituents are omitted, we get a class of "adverbs" without *-ly* such as *here, there, outside, downstairs, beforehand,* and *afterward.* If NP is present, we get the familiar forms *in the house, down the hatch, after the ball,* etc. If only PP is present, we get phrases like *out of the barn, down into the darkness, from inside the barrel,* and *up to your bedroom.* If both NP and PP are present, we get phrases like *across the street from Bill's house, down the street toward Bill,* and *to Bill in New York.* The expansion P – PP may use an intransitive preposition in the lower PP, as in *over here, down there,* and *from within.* It can also be expanded with a P – PP in the lower PP, as in *from out of the darkness, down from above the altar,* etc. Thus (4.57) seems to be perfectly productive and free of restrictions; I find it surprising that it has gone unnoticed for so long. Note that the definition of the grammatical relation "object", (4.38), generalizes directly to P', so (4.38) can now be left altogether free of category features on X.

Klima (1965) claims that subordinating conjunctions can also be analyzed as prepositions which take an $\bar{S}$ complement (from which the complementizer *that* is deleted obligatorily in modern English, optionally in Middle English). This provides the simplest description of the relation between the prepositions in $\left\{\begin{matrix}before\\after\end{matrix}\right\}$ *the ball* and the "conjunctions" in $\left\{\begin{matrix}before\\after\end{matrix}\right\}$ *the ball is over*: *before* and *after*, like many verbs, nouns, and adjectives, allow either an object or a subordinate clause. The $\bar{S}$ does not cooccur in the complement with either NP or PP, to my knowledge, so it must be an alternative to them, yielding this phrase structure rule:

(4.58) $P' \rightarrow P - \left\{\begin{matrix}(NP) - (PP)\\(\bar{S})\end{matrix}\right\}$

PPs also permit the degree Adv''', as in *not completely up the tree, quite utterly out of his mind, partly into the room,* etc. As will be seen in chapter 8, they also allow an $\bar{S}$ for degree clauses. Thus the base rule for P'' is in part (4.59):

(4.59) $P'' \rightarrow (Adv''') - P' - (\bar{S})$

4.4.4. Particles

According to the feature system of chapter 3, the [−Comp] counterpart of preposition is the category Particle. Particles take no complement structure at all; but if not followed by an NP or AP in V', they can occur with a few of the prehead modifiers that are characteristic of prepositions, as we see from (4.60a). Before NP or AP, these modifiers are impossible, though, as shown by (4.60b).

(4.60) a. I'll look the answer right up, sir.
He ate his lunch all up.
You can turn the faucet completely off.

b. We painted the house (*completely) up red.
I'll look (*right) up the answer, sir.
He ate (*all) up his lunch.

Expressing this restriction is a problem. One's first impulse is to generate PrtP as a PP to the right of all NPs in V′, then to move PrtP leftward, optionally, if it consists only of Prt. However, such an account does not explain why, in double object constructions in V′, Prt can appear only between the two NPs but not after them, nor why Prt always precedes the object in N′; these were the advantages of the alternative analysis in section 4.2. I leave the resolution of this dilemma for future research.

In any event, the phrase structure expansions for Prt are something like (4.61).

(4.61) a. Prt′ → Prt
b. Prt″ → (Adv‴) - Prt′

These rules do not account for the specifiers *all* and *right* in (4.60a); the rules for PP will have to be emended in parallel fashion.

4.4.5. Case Markers and Prepositions

Before discussing generalization of the phrase structure rules for complements, we digress briefly to make a small point about case marking. It is often claimed that prepositions are not a lexical category, but rather that they are simply case markers on noun phrases, possibly even inserted by transformations. This approach is institutionalized in Fillmore's case grammar (Fillmore (1968)); Postal (1971) trivializes the role of prepositions still further. These analyses are based on the mistaken assumption that the only possible complement to a preposition is NP; if prepositions enforce strict subcategorization restrictions and occur with such bizarre complements as PP and even NP - PP, the analysis is obviously untenable. Nevertheless, there is some relationship between prepositions and cases which we would do well to clarify at this point.

The reason for considering prepositions to be case markers is semantic: prepositions in one language often translate as case markers in another language, and vice versa. And even within a single language, prepositions and case markers often perform very similar semantic functions, as alternative realizations of indirect objects, for example.

One might therefore suggest that case markers are transformationally reduced prepositions. Though diachronically the reduction of prepositions to case markers is undeniable, it is not clear that such a process plays a role in synchronic grammar.

Alternatively, it is possible to consider case markers and prepositions as distinct syntactic entities, the ranges of whose interpretations overlap to some extent. Such is the case with adverb phrases and prepositional phrases, for example (cf. Jackendoff (1972, chapter 3)). Prepositions then can be considered a lexical category on a par with nouns, verbs, and adjectives, while case markers are affixes that are attached to NPs.

Taking this view allows a more unified theory of case marking in languages such as German and Russian, which have a mixture of prepositional phrases and cases: prepositions, like every other lexical category, determine the cases of their complements. Thus the lexical and transformational machinery necessary to enforce case marking is very general, applying to the complements of all lexical categories. By contrast, in the prepositions-as-case-markers view, the case markings due to prepositions must be accounted for separately from those due to other categories. Thus, treating prepositions as a genuine lexical category, besides accounting for obvious facts about their distribution, permits a more unified treatment of case marking.

4.5. Generalization of Phrase Structure Rules

Having described a substantial number of phrase structure rules for individual categories, let us see what the prospects are for generalizing them.

Here are the rules for nontrivial expansions of X′, in the forms developed in preceding sections, without collapsing.

(4.32) $V' \rightarrow V - (NP) - (Prt''') - (\begin{bmatrix} -Obj \\ -Det \end{bmatrix}''') - (PP) - (\begin{bmatrix} +Obj \\ +Comp \end{bmatrix}''')$

(4.33) $N' \rightarrow N \quad - (Prt''') - (NP) - (PP) - (\begin{bmatrix} +Obj \\ +Comp \end{bmatrix}''')$

(4.51) $A' \rightarrow A \quad - (NP) - (PP) - (\begin{bmatrix} +Obj \\ +Comp \end{bmatrix}''')$

(4.58) $P' \rightarrow P \quad - \left\{ \begin{matrix} (NP) & - (PP) \\ & (\bar{S}) \end{matrix} \right\}$

Certain generalizations about word order are apparent from inspection of these rules. In X′, the head precedes all other phrases. There is a strong recurrence of N‴, P‴, and $\bar{S}$ in the rules for X′, and their relative order is the same in all cases where they cooccur: N‴ always precedes P‴ and $\bar{S}$, and P‴ always precedes $\bar{S}$. If these are linguistically significant generalizations (as I believe they are), the evaluation measure for the theory of grammar ought to be constructed in such a way as to favor this grammar over one (for example) in which V′ expands as V - N‴ - P‴ but P′ expands as P - P‴ - N‴. The usual way to construct an evaluation measure is to develop notational conventions in terms of which the more general case takes fewer symbols to state. This is the purpose of parentheses, braces, and distinctive feature notations (cf. Chomsky and Halle (1968, chapter 8)).

In this case, though, a number of instances arise in which the ordinary abbreviatory conventions are of no avail. First, suppose that we try to collapse just the V′ and P′ rules so that the appropriate parts of the complement coincide. In simplified form, showing only the relevant parts, we get rule (4.63) as an abbreviation for (4.62a,b).

(4.62) a. $V' \rightarrow V - (N''') - (P''') - (\bar{S})$

b. $P' \rightarrow P - \left\{ \begin{matrix} (N''') - (P''') \\ (\bar{S}) \end{matrix} \right\}$

(4.63) $\begin{bmatrix} X \\ \langle +\text{Subj} \rangle_1 \\ \langle -\text{Subj} \rangle_2 \\ +\text{Obj} \\ +\text{Comp} \end{bmatrix}' \rightarrow X - \left\{ \begin{matrix} (N''') - (P''') \\ \langle (\bar{S}) \rangle_2 \end{matrix} \right\} - \langle (\bar{S}) \rangle_1$

Capturing the generalization about the relative order of N''' and P''' misses the generalization about the presence of $\bar{S}$ in the complement. In (4.63) the two occurrences of $\bar{S}$ have nothing to do with each other, and either could be changed to something else without affecting the complexity of (4.63). This is clearly an incorrect result. The only way to improve on this is with a complex Boolean condition such as in (4.64), simulating the effect of the braces in (4.62b) within the constraint of a linear string.

(4.64) $\begin{bmatrix} X \\ +\text{Obj} \\ +\text{Comp} \end{bmatrix}' \rightarrow X - (N''')_1 - (P''')_2 - (\bar{S})_3$

Condition: If X= P, not (3 and (1 or 2)).

Apparently, then, all the X′ rules can be collapsed, while preserving the generalizations about order, only by means of negative Boolean conditions. The problem becomes worse if we try to refine all the X′ rules to prevent nonexistent expansions. For example, A′ appears not to allow NP to cooccur with any of the other complements; thus, a more refined rule would contain a set of braces different from that in P′, requiring more Boolean conditions still.

It is hard to judge how good such a solution is. The introduction of more or less arbitrary Boolean conditions increases the power of a phrase structure grammar considerably, and one would have to ask very seriously how they can be constrained. Furthermore, it is not clear how one would define an evaluation measure over such a set of conditions, so as to know exactly what is a generalization and what is not. Nonetheless, the numerous partial generalizations of English seem to require some such rather complex solution. I can offer no interesting suggestions for alternatives to Boolean conditions. A serious study of the phrase structure rules of a number of other languages would no doubt elucidate what problems must be solved by an explanatory theory of cross-category phrase structure rules.

To make it clear that this is not a pseudoproblem, let us just mention a few other cases in the phrase structure component of English where the traditional abbreviatory conventions are inadequate. First reconsider the rules for X‴ developed in chapter 3:

(3.42) a. $V''' \rightarrow (N''') - (M''') - V''$

b. $N''' \rightarrow (\left\{ \begin{matrix} N''' \\ Art''' \end{matrix} \right\}) \quad - N''$

Anticipating the results of chapter 6, we will also mention these rules for the A''' and Adv''' specifier for comparison.

(4.65) a. $A''' \rightarrow (Deg''') - A''$

b. $Adv''' \rightarrow (Deg''') - Adv''$

There is a clear relationship between the Degree phrase in A''' and Adv''' and the Article in N''', and in fact a few words such as *this* and *that* appear in both categories. What Deg''' and Art''' have in common is that they are the [−Comp, +Det] counterparts of their respective heads. Now observe that M''' is likewise the [−Comp] counterpart of its head, a generalization of possible interest, not pointed out in chapter 3. Again, however, we are faced with a dilemma: Deg''' in (4.65) can be collapsed either with M''' in (3.42a) or with Art''' in (3.42b), but not with both; there is no way to collapse all four rules at all without a Boolean condition that restates the braces.

(4.66) $\begin{bmatrix} X \\ \alpha Subj \\ \beta Obj \end{bmatrix}''' \rightarrow (N''')_1 - (\begin{bmatrix} \alpha Subj \\ \beta Obj \\ -Comp \\ +Det \end{bmatrix}''')_2 - X''$

Conditions: If X = [−Subj], not 1
If X = N, not (1 and 2)

This case is thus formally parallel to the V′ – P′ generalization pointed out before.

A different sort of problem arises in trying to collapse the rules for X″.

(4.44)	V″	→ (have – en) – (be – ing) – (	$\begin{bmatrix} Adv \\ +Trans \end{bmatrix}'''$	)*	– V′	– (P‴)*	– (S̄)
(4.45)	N″	→		(A‴)*	– N′	– (P‴)*	– (S̄)
(4.52)′	A″	→		(Adv‴)	– A′	– (P‴)	– (S̄)
(4.56)	Adv″	→		(Adv‴)	– Adv′	– (P‴)	– (S̄)
(4.59)	P″	→		(Adv‴)	– P′		– (S̄)
(4.61b)	Prt″	→		(Adv‴)	– Prt′		

The problem lies in capturing the generalization that all categories have a prehead modifier that is [−Subj, −Obj]. N″ and V″ have special wrinkles in the modifier: both allow more than one; in V″ it is transportable; and in N″, unlike all other categories, it is [+Comp]. Expressing these deviances with any of the usual notations would be extremely clumsy, if at all possible. Of course, unconstrained Boolean conditions can

handle the problem:

(4.67) $X'' \rightarrow (\text{have} - \text{en})_1 - (\text{be} - \text{ing})_1 - (\begin{bmatrix} -\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''')_2 - X' - \ldots$

Conditions: If X ≠ V, not 1
If X = [+Subj], 2 is *
If X = V, 2 is [+Trans]
If X = N, 2 is [+Comp]
If X ≠ N, 2 is [−Comp]

These conditions are still less constrained than those in (4.66), since they specify actual features and not just presence or absence of constituents. The problem is to find some device less constrained than traditional abbreviatory conventions but more constrained than such arbitrary conditions.

Perhaps the most striking case is one that will be motivated in more detail in chapter 6. All [+Comp] categories except verbs permit a measure phrase (a kind of NP) or a quantifier phrase in the X″ specifier:

(4.68)	a. two parts steel	(NP – N′)
	much steel	(QP – N′)
	b. two miles long	(NP – A′)
	little interested	(QP – A′)
	c. two miles down the road	(NP – P′)
	far down the road	(QP – P′)

In the verb phrase there is also an alternation of a measure phrase and a quantifier phrase in V″, but this time it is to the right of the head:

(4.69)	a. He jumped into the air two times.	(V′ – NP)
	b. He talked about sex too much.	(V′ – QP)

To make matters worse, the measure phrase and (under certain conditions) the QP can follow the head in PP:

(4.70)	a. down the road two miles	(P′ – NP)
	b. down the road quite far	(P′ – QP)

There is thus a generalization that [+Subj, −Obj, −Det]‴ can be generated as a daughter of X″. But traditional phrase structure rules cannot express this generalization at all, because of the difference in word order. In this case even Boolean conditions do not provide much help. The best one could do is something like (4.71) (relying on

semantic conditions to prohibit both at once in PP):

$$(4.71)\quad X'' \rightarrow \ldots \left(\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}^{'''}\right)_1 - \ldots - X' - \ldots - \left(\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}^{'''}\right)_2 \ldots$$

Conditions: If X = V, not 1
If X = [−Obj], not 2

But this does not capture the generalization that the same expression appears on both sides of X′; the grammar would be no more complex (in terms of number of symbols) if the subscripted terms in (4.71) were of entirely different categories but used the same number of features. One might suggest instead that all these modifiers are generated before the head, and a transformation permutes them obligatorily around a V′ and optionally around a P′. Such a rule, however, would be non-structure preserving, and besides would be supported by no independent syntactic evidence, at least in the case of V′.

Though this is the only such case I know of in English, there is reason to believe that it is not such a rare situation. It would arise in its most drastic form in a language with a verb-final V′ but a noun-initial N′; German, if it is SOV, as has been rather persuasively argued by many people, is such a language. Languages in which adjectives follow nouns but adverbs precede adjectives would be another such case; so would SOV languages with prepositions rather than postpositions. Such languages are cited by Greenberg (1963).

Greenberg points out, however, that such languages are relatively rare; his "universals" say that the opposite correlations occur with "overwhelmingly greater than chance frequency". This suggests that the evaluation measure for phrase structure rules counts parallelisms of word order across X^n as generalizations, as we have come to expect in the present study. But it appears also that parallel grammatical relations with differing orders across X^n also count as generalizations, though not as strongly; present notations have no way at all to express this. Again, detailed study of the phrase structure of languages other than English is necessary, in order to establish exactly what generalizations there are and how the formalism for phrase structure rules should be constrained so as to explain why these and no other generalizations appear.

4.6. Summary

Section 4.1 showed that complements divide semantically into three types: functional arguments, restrictive modifiers, and nonrestrictive modifiers, and that these can be identified with X′, X″, and X‴ complements, respectively. Thus, there is a systematic relationship between the syntax and the semantics of the complement system.

The rest of the chapter developed phrase structure rules for the X′ and X″ complements of most syntactic categories in English, showing the extent of cross-category generalization in the complement system. A close examination of the existing

generalizations revealed, however, that many of them cannot be expressed with the traditional notational conventions.

There follow two appendices. The first deals with the generalization of the Passive to NPs; the second presents some new arguments about Intraposition. Besides the phrase structure rules listed in the previous section, this chapter and its appendices discuss a substantial number of transformations; they are listed after the appendices.

4.7. Appendix: The Passive and Related Rules

This appendix describes how Chomsky (1970a) generalizes the Passive transformation to NPs, and proposes a number of improvements on his analysis.

4.7.1. NP-Preposing

Consider the relationship between the phrases in (4.72a) and the synonymous readings of the phrases in (4.72b).

(4.72) a. the destruction of the enemy
the murder of John
the picture of John
the proof of the theorem
b. the enemy's destruction
John's murder
John's picture
the theorem's proof

If we relate them with a transformation, taking (4.72a) as the underlying order, we can explain two things: the syntactic relationship of these phrases to the sentences *X destroyed the city, X murdered John,* etc., in which the NP is a deep object; and the ambiguity of (4.72b), where the genitive phrase can originate either as deep subject or deep object. Chomsky calls the transformation deriving (4.72b) from (4.72a) NP-Preposing (p. 41/203).

This rule applies only to NPs with a direct object (in the sense of section 4.2): there are no NPs **war's prayer, *my despair's escape, *power's lust* corresponding to *the prayer for war, the escape from my despair, the lust for power.* This provides some interesting evidence that the *of* of direct objects in NPs is not present at some earlier stage of derivation. The argument goes as follows: the sense of *approve* which takes a direct object, e.g. *approve the plan,* has the nominals *the approval of the plan* and *the plan's approval.* But the sense of *approve* which takes an *of*-phrase, e.g. *approve of your behavior,* has only the nominal *the approval of your behavior* and not **your behavior's approval.* We can explain this difference by claiming that the former sense of *approval,* like the verb, lacks *of* in deep structure, but the latter sense requires it. Then only the former sense will permit NP-Preposing, although the deep structure difference of the two senses is obscured by the later rule *Of*-Insertion.

Some apparent irregularities in the application of NP-Preposing are explained by Wasow (1975) as a consequence of the structure-preserving hypothesis. Wasow points out that for NP-Preposing to be structure preserving, the preposed NP must be moved into an independently generable empty NP position. If an empty subject NP were for some reason impossible, NP-Preposing could not take place. He then observes the contrasts (4.73) versus (4.74).

(4.73) a. The report of the invasion troubled John.
b. The reading of the honor roll embarrassed John.
c. The singing of the aria bored John.
d. John enjoyed the leveling of the city.

(4.74) a. The authorization of the invasion troubled John.
b. The publication of the honor roll embarrassed John.
c. The composition of the aria bored John.
d. John enjoyed the destruction of the city.

In (4.73) *John* may not be understood as having subject function in the nominal, but in (4.74) he may. Wasow claims that the difference is due to the ability of the nominals in (4.74) to take an empty (or PRO) NP specifier and the inability of the nominals in (4.73) to do so; the difference is justified independently in Wasow and Roeper (1972). Wasow then points out that this difference correctly predicts that only the nominals in (4.74) undergo NP-Preposing, since only they allow an empty NP into which the object can prepose.

(4.75) a. *The invasion's report troubled John.
b. *The honor roll's reading embarrassed John.
c. *The aria's singing bored John.
d. *John enjoyed the city's leveling.

(4.76) a. The invasion's authorization troubled John.
b. The honor roll's publication embarrassed John.
c. The aria's composition bored John.
d. John enjoyed the city's destruction.

Wasow's analysis does not exhaust all the situations where NP-Preposing fails, but it accounts for an interesting class of cases.

For the moment we will assume that *the* is inserted to fill an empty NP in the specifier position; we will return to this in section 4.7.4. Under this assumption, NP-Preposing can be stated as (4.77), where lower case denotes an empty node.

(4.77) *NP-Preposing*
np – Y – N – NP – Z
$\Rightarrow$ 4 – 2 – 3 – ϕ – 5
OPTIONAL

4.7.2. Agent-Postposing

Next consider the relationship of the phrases in (4.78a) to those in (4.78b).

(4.78) a. the enemy's destruction of the city
Lange's portrait of Mozart
Euler's proof of that theorem
b. the destruction of the city by the enemy
the portrait of Mozart by Lange
the proof of that theorem by Euler

Again, the synonymy of the pairs can be captured by a transformational relationship. The parallelism of (4.78a) to *The enemy destroyed the city, Lange portrayed Mozart,* and *Euler proved that theorem* suggests that (4.78a) is the underlying order and (4.78b) is derived. To state the transformation in a structure-preserving framework, we need to provide the complements in (4.78a) with a PP node whose NP is empty. We could also leave the preposition empty; but by filling it with *by* in underlying structure, we have some semantic material to associate with the agentive function evident in (4.78b).

(4.79) *Agent-Postposing*
$[_{NP}$ NP – X – N – Y – $[_{PP}$ by np] – Z]
$\Rightarrow$ np – 2 – 3 – 4 – 5 – 1 – 7
OPTIONAL

The underlying structures of (4.78b) will be *the enemy destruction the city by np*, etc.; (4.78a) will simply lack the final PP.

If we order NP-Preposing after Agent-Postposing, we can apply them in sequence:

(4.80) the enemy destruction the city by np $\overset{(4.79)}{\Rightarrow}$
np destruction the city by the enemy $\overset{(4.77)}{\Rightarrow}$
the city destruction ϕ by the enemy $\overset{(Poss\text{–Insertion})}{\Rightarrow}$
the city's destruction ϕ by the enemy

But the final form in (4.80) is precisely what we observed in chapter 2 as a parallel to the passive sentence *The city was destroyed by the enemy.*

The Passive transformation, although traditionally written as a single rule which moves two NPs, can clearly be broken into two components, the movement of the subject into the *by*-phrase and the movement of the object into subject position. The first component obviously generalizes with Agent-Postposing: the *by*-phrase is used in precisely the same way, and the NP is moved out of parallel positions. The rules combine as (4.81).

(4.81) *Agent-Postposing—Passive, Part 1*

$$[_{X'''}\text{NP} - \text{Y} - \begin{bmatrix} \text{X} \\ +\text{Subj} \\ +\text{Comp} \end{bmatrix} - \text{Z} - [_{\text{PP}}\text{by np}] - \text{W}]$$

$$\Rightarrow \text{np} - 2 - 3 - 4 - 5 - 1 - 7$$

The second component of the sentential Passive does not immediately generalize so clearly with NP-Preposing. This part of the Passive, unlike NP-Preposing, is obligatory, since there are no sentential forms like (4.78b) (e.g. **it destroyed the city by the enemy*). Furthermore, under certain conditions the passive construction permits a preposition to intervene between the NP to be preposed and the verb, as in *This bed has been slept in by an infamous violinist.* The second half of the Passive must also introduce the auxiliary *be-en*.

(4.82) *Passive, Part 2*

$$\text{np} - \text{Y} - \text{V} - (\text{P}) - \text{NP} - \text{Z}$$

$$\Rightarrow 5 - 2 - \text{be-en} - 3 - 4 - \phi - 6$$

OBLIGATORY

(4.82) could clearly be combined with (4.77) by the usual notations. However, the difference between optional and obligatory use casts a certain amount of suspicion on the generalization. Fiengo's (1974) analysis of the passive provides some hope of improvement. In Fiengo's theory, *be-en* is present in the underlying structure of passives, and semantic conditions on traces guarantee that only passive surface structures are associated with it. Furthermore, he generalizes the rule so as to use it to derive examples like *These books sell easily* from sources like *np sell these books easily*; thus *be-en* plays a role neither in the structural description nor in the structural change of the Passive. Because of the presence of various independent conditions, then, Fiengo's analysis eliminates the *be-en* and the obligatory application from (4.82). The only remaining discrepancy is then the optional preposition, so the generalization of NP-Preposing looks somewhat more plausible than in the analysis of "Remarks", interpreted strictly. Conceivably the generalized form could be at least as clean as (4.83).

(4.83) *Generalized NP-Preposing*

$$\text{np} - \text{Y} - \begin{bmatrix} +\text{Subj} \\ \langle +\text{Obj} \rangle \\ +\text{Comp} \end{bmatrix} - \langle (\text{P}) \rangle - \text{NP} - \text{Z}$$

$$\Rightarrow 5 - 2 - 3 - 4 - \phi - 6$$

OPTIONAL

Besides the generalization with NP passives and the elimination of a rule moving two constituents simultaneously, there is some independent reason to separate the two components of the Passive. As pointed out by Emonds (1970), sentences like (4.84)

could be generated as deep structures with empty subjects and derived by the use of (4.83) alone, avoiding the suspicious rule of *By-Someone* Deletion used traditionally in deriving examples of this sort.

(4.84) Frege was born in 1485.
John is $\left\{\begin{matrix}\text{said}\\\text{rumored}\end{matrix}\right\}$ to like tamales.

4.7.3. NP-Postposing

Continuing with the syntax of N‴, Chomsky appeals (p. 44/205) to

> a minor transformational rule which will replace *by* by *of* under certain conditions, permitting *the refusal to leave of those men* (or *the refusal of those men to leave*). Presumably it is this rule that applies in the case of the nominals *the growling of the lion*, etc.

However, Bowers (1968a) argues against the *by* → *of* rule, proposing instead that there is an independent rule of NP-Postposing with the following effect (stated in the present framework):

(4.85) *NP-Postposing*
NP – Y – N – np – Z
⇒ np – 2 – 3 – 1 – 5
OPTIONAL

Bowers's argument concerns the forms of nominals related to nonpassivizable sentences. Compare (4.86a) with (4.86b).

(4.86) a. the building's height
the railroad's length
the lake's depth
the desert's dryness
that proposal's absurdity
the rock's whiteness
b. the height of the building
the length of the railroad
the depth of the lake
the dryness of the desert
the absurdity of that proposal
the whiteness of the rock

In order to avoid duplicating the strict subcategorization restrictions, selectional restrictions, and projection rules for these two forms, we can relate them transformationally. If we choose (4.86b) as the underlying order, we can derive (4.86a) by NP-Preposing. But this choice has unfortunate consequences for the statement of lexical

relations. For these nouns and the adjectives *high, long,* etc., would be related in the simplest way if they took syntactically parallel arguments. Since there are no forms **high of the building, *long of the railroad,* etc., but rather only *The building is high, The railroad is long,* etc., the optimal relation between nouns and adjectives under the $\bar{X}$ Convention favors (4.86a) as the underlying order. This of course requires the existence of a rule like (4.85).

Similar considerations apply in the following nominals.

(4.87) a. John's resemblance to Mary
the parcel's weight
the book's cost
b. the resemblance of John to Mary
the weight of the parcel
the cost of the book

Again a transformational relationship is called for. If we choose (4.87b) as the underlying order, NP-Preposing can derive (4.87a), but we lose the parallelism of grammatical relations with *John resembles Mary, The parcel weighs a lot,* and *The book costs a lot.* If we take (4.87a) as the underlying order, so as to have parallelism with the verbal forms, we could conceivably use Agent-Postposing and Chomsky's *by* → *of* rule. But in these cases there is evidence against the use of Agent-Postposing, since the corresponding verbs do not undergo it: **Mary is resembled by John, *A lot is weighed by the parcel, *A lot is cost by the book.* Thus if we want to keep lexical relations as simple as possible, we need NP-Postposing to derive (4.87b) directly from (4.87a), with no intermediate step involving *by.* Bowers argues further that choosing (4.86a) as the underlying form of the adjectival nominalizations correctly predicts that (4.86b) does not alternate with an indefinite form **a height of the building,* etc., since NP-Postposing always leaves behind a definite article. We incorporate this argument into our account of determiners below.

NP-Postposing derives various examples of Chomsky's: it applies to "intransitive" nouns such as those in (4.86a) and (4.87a), yielding also *the growth of the corn* from *the corn's growth* (cf. Chomsky, p. 25/192), *the fall of Rome* from *Rome's fall, the growling of the lion* from *the lion's growling,* and *the refusal of those men to leave* from *those men's refusal to leave* (I happily find Chomsky's example *the refusal to leave of those men* ungrammatical). NP-Postposing thus has an effect exactly opposite to that of NP-Preposing. Since the underlying structures involved are generally well-motivated, the existence of two transformations with opposite effects should not be disturbing. In certain cases, though, it may as a result be difficult to determine which of two forms is underlying and which derived, for example *the friend of the princess* vs. *the princess's friend.* Some evidence will appear shortly.

The system of rules at this point thus includes Agent-Postposing in NPs and Ss, NP-Preposing in NPs and Ss, and NP-Postposing in NPs.

4.7.4. The Determiner

A major difficulty with the analysis so far, essentially a variant of Chomsky's analysis,[17] is that it deals only with NPs whose determiner is either a possessive or *the*. It therefore gives no account of NPs such as (4.88).

(4.88) a. a victory by the Vietnamese
b. three books by Harry
c. every effort by Bill to win
d. some proofs of that theorem by well-known mathematicians

The putative source of the *by*-phrases in these examples, Agent-Postposing, would leave behind an empty NP in the determiner, which by assumption would have to be realized as *the*, not as any of the other determiners above. Transformational insertion of an arbitrary determiner in place of the empty NP makes no sense; but the determiner and the agent cannot both be generated in the same position in the base either (unless we resort to Emonds's dubious doubly-filled nodes). Apparently the only possible alternative remaining is to allow the *by*-phrase to be generated in the base, adding a projection rule that can read it as an Agent. But this throws doubt on the existence of the Agent-Postposing transformation: why derive something by transformation that is already generable by the base?

Now if Agent-Postposing is eliminated in NPs and replaced by a projection rule for a complement *by*-phrase, it would be missing a generalization to retain the transformation in sentences. Rather, we must conclude that passive *by*-phrases in Ss are also generated by the base and interpreted by a generalized form of the projection rule for the NP *by*-phrase. According to this analysis, the underlying form of *The city was destroyed by the enemy* is *np past* (*be–en*) *destroy the city by the enemy,* and the only transformation that takes place is (4.83), the preposing of the object.

Alternatively, one might consider retaining Agent-Postposing in Ss as a second source of *by NP* in addition to base generation. The application of this rule could be used to account for some of the discrepancies between NP passives and S passives, such as **the sense of danger by John, *the fear of Harry by John,* and **the respect for Mary by John* vs. *Danger was sensed by John, Harry was feared by John,* and *Mary*

[17] Our account differs from Chomsky's analysis not only in the treatment of NP-Postposing in NPs, discussed above, but also in the treatment of the determiner. Instead of an empty NP determiner in *the destruction of the city by the enemy*, filled transformationally by *the*, Chomsky proposes that the determiner position have the syntactic feature analysis [+NP, +Def], with the latter feature spelled out as *the* if NP is moved away. Since at the time of "Remarks" empty nodes were not an available device, Chomsky's solution appeared much more necessary. Jackendoff (1974a) makes an attempt to work it out in some detail; the complexity of the resulting theory has been criticized with some justification by Hornstein (1975). Besides its complexity, there are two arguments against Chomsky's stratagem of combining category and definiteness features at a single node in the base. First, no one has found another use for this device since 1967, the time at which "Remarks" first appeared. Second, this device strikes me as utterly inadequate to cope with examples like (4.88). I thus feel little regret in abandoning it.

was respected by John (these cases pointed out by Hornstein (1975)). Note that these all are cases in which the *by*-phrase does not designate an agent, so they could not be interpreted correctly by the proposed projection rule. They are thus good candidates for a transformational derivation of the *by*-phrase from subjects. On the other hand, this solution might make it more difficult to give a semantic account of those verbs which do not undergo Passive at all, e.g. *resemble, weigh,* etc. I leave the issue open.

Note that even if Agent-Postposing is abandoned altogether, we are not led to something like case grammar, in which *all* agents are generated as *by*-phrases in underlying structure. Rather, NPs can be interpreted as agents either in subject position or in *by*-NP complements, and there is no transformational relationship between the two positions at all.[18]

Since NP-Preposing requires an empty NP in subject position, and since the determiner of an NP may be either an article or an empty NP, the following derivations may take place.

(4.89) a. {every / that / the} proof of {that / a} famous theorem

(no transformations apply)

b. np proof of {that / a} famous theorem $\xrightarrow{\text{NP-Preposing}}$

{that / a} famous theorem's proof

c. np proof of {that / a} famous theorem $\xrightarrow{\textit{The}\text{-Insertion}}$

the proof of {that / a} famous theorem

We are thus positing two sources of *the proof of that theorem,* one in which *the* is base-generated as an article, and one in which *the* is inserted transformationally to fill an empty NP.

One might think that the arguments for eliminating Agent-Postposing would apply equally against (4.85), NP-Postposing. But interestingly enough, a substantial number of the nominals which we proposed to derive by NP-Postposing appear only with the article *the*, so counterexamples like (4.88) do not arise.

[18] Emonds (1976) proposes that all agents originate in a *by*-phrase and are optionally preposed to subject position. Though I have not investigated the consequences of this alternative thoroughly, it strikes me that it will require at least equal complexity in the grammar, and that it is somewhat less plausible than the present theory, which simply invests an optional phrase structure position with its own interpretation, similar to instrument and accompaniment phrases.

(4.90) $\left\{\begin{matrix}\text{the}\\ *\text{a}\\ *\text{some}\\ *\text{every}\end{matrix}\right\}$ cost(s) of the book

$\left\{\begin{matrix}\text{the}\\ *\text{a}\\ *\text{some}\\ *\text{every}\end{matrix}\right\}$ resemblance(s) of John to Mary (compare $\left\{\begin{matrix}\textit{the}\\ \textit{a}\\ \textit{some}\end{matrix}\right\}$ *resemblance(s) between John and Mary*)

$\left\{\begin{matrix}\text{the}\\ *\text{a}\\ *\text{some}\\ *\text{every}\end{matrix}\right\}$ height of the building

Within the proposed analysis, a very simple account of this determiner restriction emerges: these nouns subcategorize only NPs in the determiner, not articles and quantifiers. Thus a configuration like (4.89a) cannot come about. *The* is introduced transformationally when the NP position either is empty in the base or becomes empty through the application of NP-Postposing.

A similar case is the action reading of *destruction,* which subcategorizes an object. Though *some destruction* is all right, **some destruction of the city* is not. We can describe this difference by claiming that the action reading subcategorizes not only an object but also an NP subject. Thus, the only article that can appear with it is *the*, inserted if the subject happens to be empty at the surface.

Suppose we find nominals which are like (4.90) in that what is the subject of the corresponding sentence follows the head of the nominal, but which differ from (4.90) in that other determiners besides *the* are possible. Our solution must be to generate these forms in the base instead as a result of NP-Postposing. In this case the verb and its nominalization will not be optimally related, since the verb's subject will correspond to the noun's object instead of to the noun's subject as in (4.90).

Some speakers may consider *growth* and *growling* relevant examples of this sort, finding grammatical *all growth of corn* (related nonoptimally to *Corn grows*) and *all growling of lions* (related to *Lions growl*). I find these examples marginal at best, and *all* is the only determiner other than *the* that sounds even faintly plausible. In any event, the existence of such nonoptimal examples is not a problem: I would guess that this is an area where individuals' lexicons will differ, since the evidence to decide between a subject which postposes and an object which preposes is relatively scanty in the corpus available to a language learner, and the choice has no serious consequences elsewhere in the grammar.

This completes our account of rules moving NPs in and out of the complement. There are still no doubt quite a number of anomalies and irregularities which we have not attempted to deal with, but we have added much precision and detail to Chomsky's

rather sketchy analysis, making several improvements along the way. In the end, we claim that NP-Preposing takes place in Ss and NPs, that Agent-Postposing may take place in Ss, and that NP-Postposing takes place in NPs. An empty subject position in an NP is filled by *the*; if articles and quantifiers other than *the* appear in an NP, there can have been no NP position in the determiner, and hence no NP to the right of the head can have been an underlying subject.

4.8. Appendix: Arguments for Intraposition

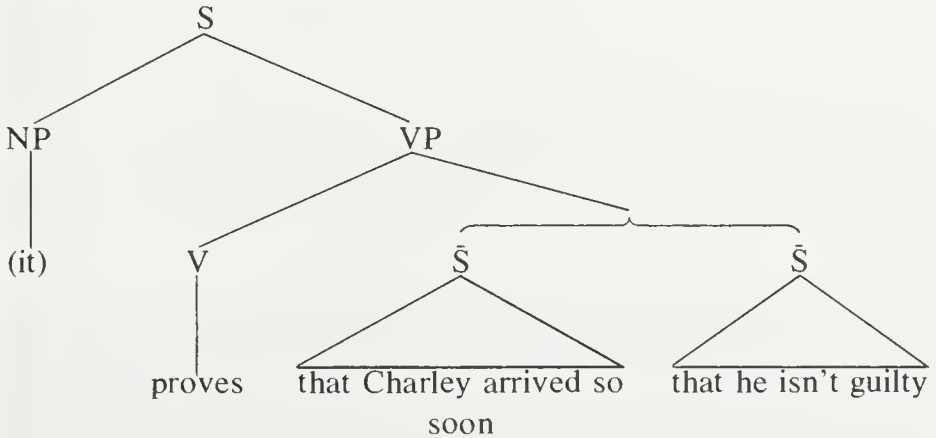
Section 4.3 claimed that the S at the end of V″ and N″ is an underlying position for so-called "subject complements" such as the $\bar{S}$ in *It is obvious that Bill left, It disturbs me that Bill left,* and *the necessity for Bill to leave.* This is the view of Emonds (1970) (though not Emonds (1976)), who claims that the synonymous sentences *That Bill left is obvious* and *That Bill left disturbs me* are derived by moving the complement leftward from underlying final position by a root transformation called Intraposition. Emonds opposes the traditional Rosenbaum (1967) Extraposition theory, which claims the derivations go in the opposite direction.

This section presents a number of arguments for the Intraposition theory, to meet certain objections that have arisen since the publication of Emonds's work. I do not, however, claim to meet all extant arguments against Intraposition. I presuppose the reader's familiarity with Emonds's arguments.

The most objectionable feature of the Intraposition theory was its account of double complement verbs such as *prove*. Since Emonds provided only one $\bar{S}$ position at the end of the sentence, he was forced to claim that (4.91a), for example, was derived from an underlying structure (4.91b) in which both complements occupied the same position, a situation Emonds called a "doubly filled node", which would have to be eliminated before surface structure.

(4.91) a. That Charley arrived so soon proves that he isn't guilty.

b.



Since the device of doubly filled nodes seems morally repugnant to most linguists, Emonds's account of (4.91a) seemed rather implausible, and many considered this

alone sufficient to dismiss the entire Intraposition analysis, or even the whole structure-preserving hypothesis.

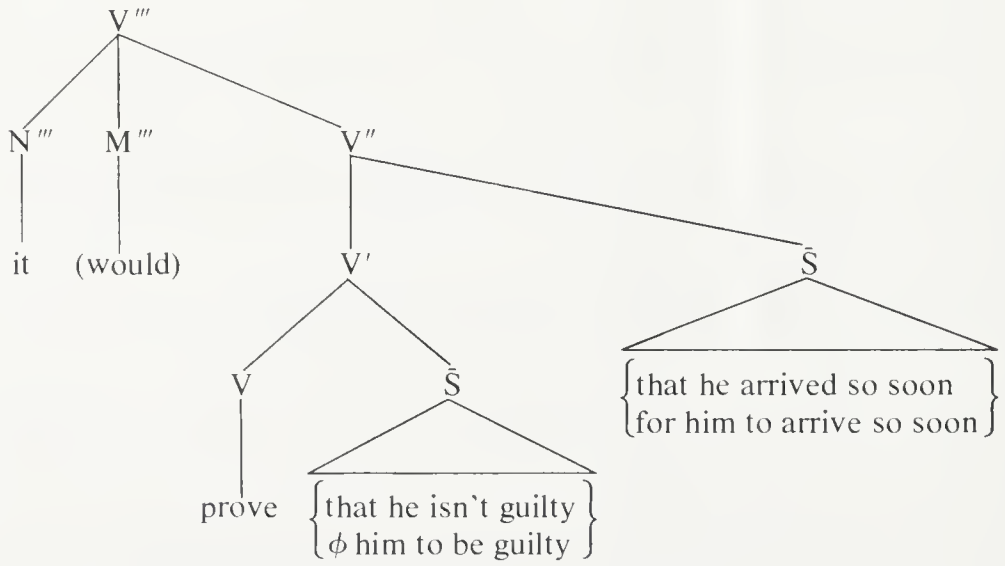
However, the problem of *prove* turns out to have a better solution within the Intraposition theory. There are some speakers who find (4.92a) and particularly (4.92b,c) acceptable.

(4.92) a. It proves that Charley isn't guilty that he arrived so soon.
b. It proves Charley not to be guilty that he arrived so soon.
c. It would prove that Charley isn't guilty for him to arrive so soon.

For these speakers, the doubly filled node solution is impossible, since the surface structure is not at all ill-formed. These examples show that Emonds was wrong in positing only one final $\bar{S}$ position.

In the present theory there are two distinct positions available for sentential complements, one in V′ and one in V″. It is possible to claim that the first clause in (4.92) is the V′ complement and the second the V″ complement, interpreted as a subject:

(4.93)



This would explain in particular why the clauses in (4.92) cannot be interchanged without changing meaning, a fact unexplained by a structure like (4.91b). It then remains to explain why (4.92) is unacceptable to many speakers (and notice that this problem exists in the Extraposition hypothesis as well). The most plausible solution in view of the dialect split is an output constraint similar to that for consecutive adverbs (see section 4.3), ruling out consecutive $\bar{S}$s within V″. Thus, doubly filled nodes can be eliminated in favor of a theoretical device of an independently necessary sort, itself

required in the competing theory. Once we realize this, the problem of *prove* simply ceases to decide between the two theories.

Another objection to the Intraposition theory concerns the relationship of the clause to the subject *it*. How does the grammar determine that the clause has all the semantic properties of a subject, without the defining syntactic property of preceding the auxiliary? Rosenbaum's solution is to claim that in the underlying form the clause *is* in subject position, accounting for its interpretation in the simplest possible way. Emonds, on the other hand, is forced to claim that there is an anaphoric relationship between *it* and the clause, established by an independent projection rule.[19] Thus, the Intraposition theory needs a rule not needed by the Extraposition theory.

However, there is some evidence that the interpretation rule is independently necessary. First, there are languages (e.g. French and Hungarian) in which a clause can never occupy subject position. There is thus no *syntactic* evidence for generating it there, and for these languages the anaphoric theory looks rather more attractive. Thus the rule of anaphora must be a possibility in linguistic theory. Second, there are sentences in English which cannot in good conscience be generated by an Extraposition theory but which have final clauses with the function of subject: compare (4.94a) and (4.94b).

(4.94) a. It bothers me that you're tickling me.
It doesn't surprise me that she came.
b. It bothers me when you tickle me.
It wouldn't surprise me if she came.

The (b) sentences have the preposed form (4.95a), not (4.95b).

(4.95) a. When you tickle me, it bothers me.
If she came, it wouldn't surprise me.
b. *When you tickle me bothers me.
*If she came wouldn't surprise me.

But since the clauses in (4.94b) have the same semantic function as those in (4.94a), their interpretation requires precisely the rule of anaphora proposed by Emonds to account for (4.94a), and in this case the alternative of an Extraposition analysis is far less plausible.

[19] Emonds says the relevant anaphoric relationship is coreference, which strikes me as somewhat incoherent, in that *it* and the complement S do not constitute two separate references to the same event. A more appropriate kind of anaphora might be termed "grammatical anaphora", by contrast with the usual "referential anaphora". When a relationship of grammatical anaphora is established, any projection rules which would normally apply to the proform by virtue of its syntactic position are directed to apply instead to the antecedent. Thus the antecedent receives an "indirect" grammatical relation. This kind of anaphora is probably the type involved in interpretation of traces, and perhaps with reflexives of certain types; it has an effect rather like a bound variable in logical notation.

Section 3.4.2 pointed out a different class of cases, in which a PP appears to take subject function:

(4.96) a. It's a long way to Tipperary.
b. It's a long time from Groundhog Day to Purim.

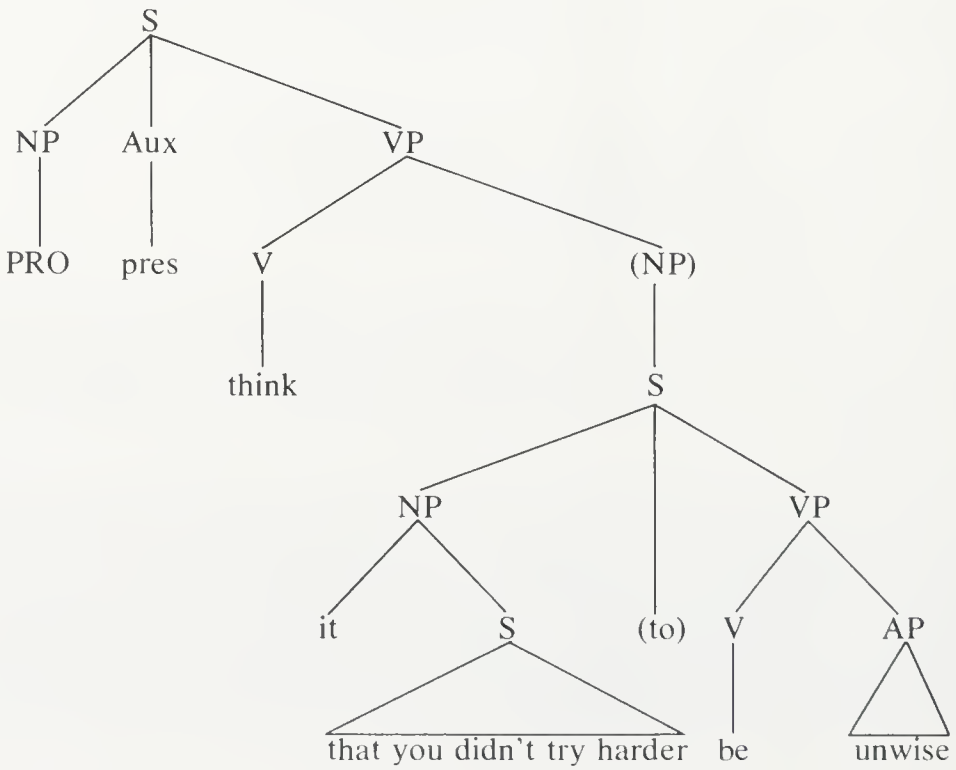
These PPs can be moved into the subject in the same environments as *that* and *for-to* complements can, and there is no problem in generating them under V″. To interpret (4.96), however, a rule of anaphora is needed which specifies the PP as what is being measured, i.e. the subject. This is the same rule needed for the Intraposition analysis. Thus, since the anaphora rule is needed anyway, it does not weigh against Emonds's theory, as originally thought.

The fact that the necessary rule of anaphora may be a *surface structure* rule of interpretation provides the key to a problem of derived structure raised by Higgins (1973). Higgins is concerned with the derivation of sentences like (4.97) (his (58)):

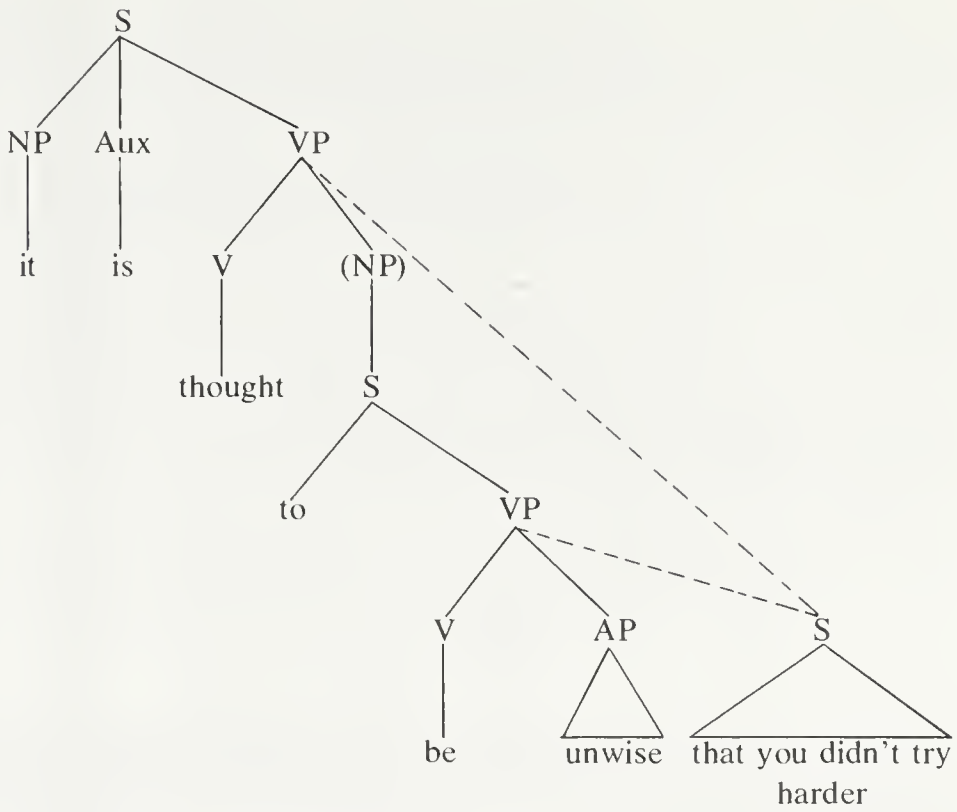
(4.97) a. It is thought to be unwise that you didn't try harder.
b. That you didn't try harder is thought to be unwise.

In the Extraposition theory, these are both derived from (4.98a), by Passive and Raising in the main clause. It is not clear on which cycle Extraposition takes place in (4.97a), so its derived structure could be either of the options in (4.98b). The fact that Extraposition is generally obligatory in subordinate clauses argues that the complement is probably attached to the lower VP.

(4.98) a.

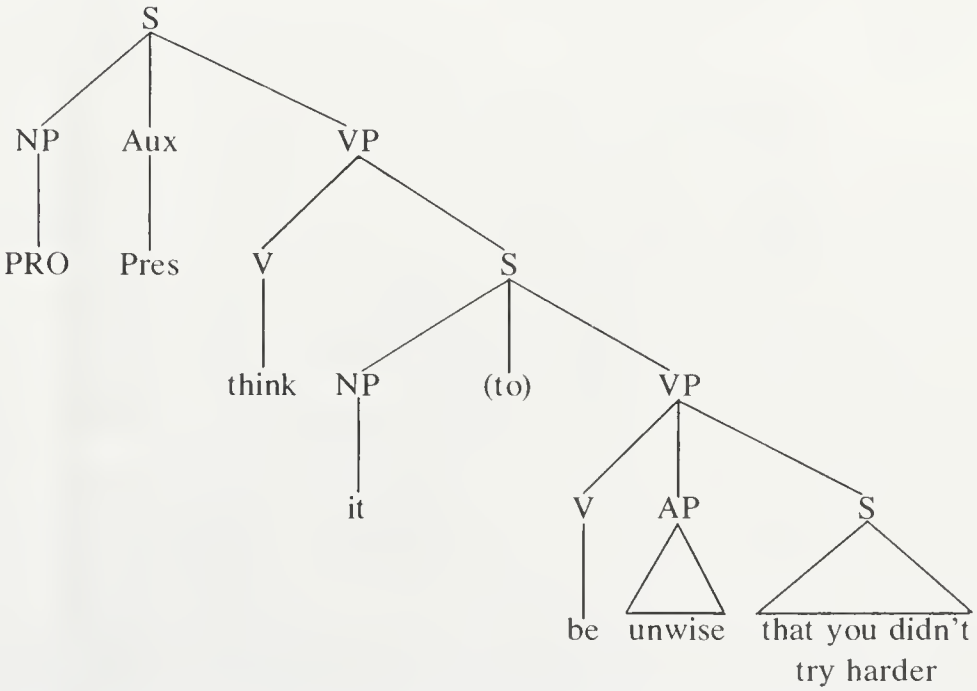


b.



Higgins assumes that in the Intraposition theory the deep structure of (4.97) is (4.99). Hence (4.97a) is derived by Passive and Raising of *it*; the derived structure is (4.98b), with the complement attached to the lower VP.

(4.99)



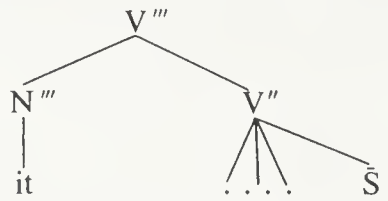
To produce (4.97b), Intraposition now has to raise the complement over a clause boundary, and Higgins correctly shows that this raising produces difficulties.

However, closer consideration of (4.97a) shows that the complement clause is attached in derived structure to the higher VP, not the lower one. For if the *by*-phrase of the passive is realized lexically, it precedes the complement rather than following it:

(4.100) a. It is considered to be unwise by many people not to try harder.
b. ?*It is considered to be unwise not to try harder by many people.

In the Extraposition theory this fact is difficult to account for, since the obligatoriness of Extraposition in subordinate clauses would force the *that*-clause to attach to the lower VP, on the earlier cycle. However, if the rule of anaphora in the Intraposition theory is like all other anaphora rules in applying to surface structures, it can be restricted to the environment (4.101), making only the higher attachment of the complement possible as a derived structure.

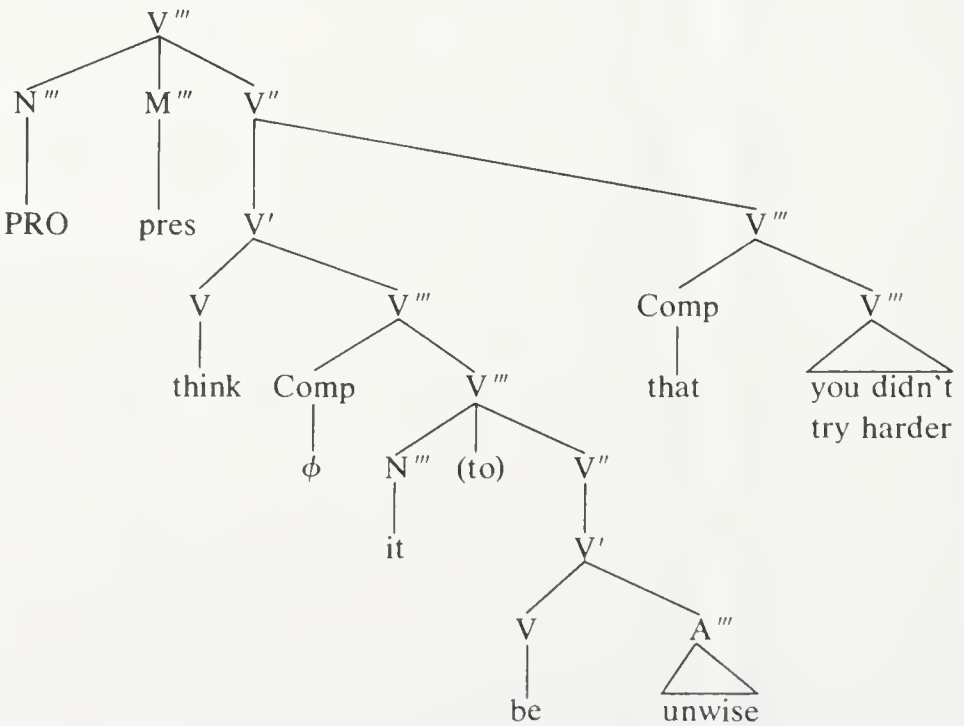
(4.101)



A movement over a clause boundary is thus unnecessary in deriving (4.97b) by Intraposition, and the problem of properly locating the *it* relative to the complement S does not arise.

This analysis of course implies that the deep structure of (4.97) also has the complement in the upper clause, thus:

(4.102)



This seems to violate our linguist's intuitions about grammatical relations, since the clause *that you didn't try harder* is entirely outside of the clause *it to be unwise* of which it is understood as the subject. This apparent anomaly disappears when we realize that this subject relation is not a normal grammatical relation, but is established only indirectly by a surface structure rule of anaphora. If *it* were not raised into the upper clause, the anaphora rule could not apply, and *that you didn't try harder* would be left without a semantic function.

This section has dealt with three objections to the Intraposition theory. First, we showed that the double-complement verbs such as *prove* should not be accounted for with doubly filled nodes, as Emonds suggested; rather, one complement is in V′ and one in V″. The constraint against concatenated complements of the same type applies equally in the Intraposition and Extraposition theories. Second, the rule of anaphora relating *it* to the V″ complement was shown to have independent motivation in both English and other languages. Third, the fact that the anaphora rule applies to surface structure enables the theory to meet Higgins's objection about derived structure.

4.9. List of Transformations

Dative Shift (section 4.2)

(4.31) $X - V_a - np - (Prt) - NP - [_{PP}to - NP] - Y$
$\Rightarrow 1 - 2 - 7 - 4 - 5 - \phi - \phi - 8$
OPTIONAL

Agent-Postposing in Ss (possibly) (section 4.7)
$NP - Y - V - Z - [_{PP}by\ np] - W$
$\Rightarrow np - 2 - 3 - 4 - 5 - 1 - 7$
OPTIONAL

Generalized NP-Preposing (section 4.7)

(4.83) $np - Y - \begin{bmatrix} +Subj \\ \langle +Obj \rangle \\ +Comp \end{bmatrix} - \langle (P) \rangle - NP - Z \Rightarrow 5 - 2 - 3 - 4 - \phi - 6$
OPTIONAL

NP-Postposing (section 4.7)

(4.85) $NP - Y - N - np - Z \Rightarrow np - 2 - 3 - 1 - 5$
OPTIONAL

Intraposition (sections 4.3, 4.8)
$[_{V'''}it - Y - [_{V''}Z - \bar{S}] - W] \Rightarrow 4 - 2 - 3 - \bar{s} - 5$
OPTIONAL—ROOT (may generalize to movement of [+Obj]‴)

Extraposition of $\bar{S}$ (section 4.2)

$[_{X'}Z - \bar{S}] - Y - \bar{s}]_{X''} \Rightarrow 1 - \bar{s} - 3 - 2$
OPTIONAL

Extraposition of PP (section 4.3)—generalizes with Extraposition of $\bar{S}$

$[_{X'}Z - PP - W] - Y - pp - U]_{X''} \Rightarrow 1 - pp - 3 - 4 - 2 - 6$
OPTIONAL

Of-*Insertion* (section 4.2, generalized in section 4.4)

$[_{X'}\begin{bmatrix} X \\ -Obj \end{bmatrix} - (Prt) - NP - Y] \Rightarrow 1 - 2 - of + 3 - 4$
OBLIGATORY

Poss-*Insertion* (section 3.4)

$[_{N'''}N''' - Y - N''] \Rightarrow 1 + poss - 2 - 3$
OBLIGATORY

The-*Insertion* (section 4.7)

$[_{N'''}np - Y - N''] \Rightarrow the - 2 - 3$
OBLIGATORY

That-*Deletion* (section 4.4)

$[_{P'}P - that - S] \Rightarrow 1 - \phi - 3$
OBLIGATORY

Output Constraints

? Y – Adv – Adv – Z (section 4.3)
? Y – $\bar{S}$ – $\bar{S}$ – Z (section 4.8)

5: NP Specifiers

5.1. Introduction

Chapter 3 dealt with aspects of the specifier that are central to the syntax of sentences: the subject and the auxiliary. This chapter and the next extend the study of specifiers to systems which play a minor role in sentences but which are important in NPs, APs, and PPs: determiners and expressions of quantification and degree.

There are problems in studying specifier systems that do not arise in studying complements. First, specifier systems involve very small numbers of lexical items and are riddled with idiosyncrasies. Thus general phrase structure rules must be supported on the basis of impoverished and skewed surface distributions. Only with judicious use of lexically determined transformations and interpretation rules can any order at all be brought into the system. One way to bring more data to bear on the problems is to exploit the hypothesis of cross-category generalization, and that will be a major source of evidence here.

A second problem with specifier systems, at least in English, is that it appears much less possible to correlate semantic regularities with syntactic positions, as we did in complements. This may of course be a function of our ignorance about the semantics of deixis, quantification, and measuring; but it may also be a brute fact about English or about specifiers in general. In the absence of a coherent semantic theory of specifiers, I will make the latter assumption, hoping that a better understanding of the syntax may make possible a more disciplined approach to the semantics.

This chapter deals with the syntax of the NP specifier. Chapter 6 extends this account to AP and PP, proposing a reanalysis of Bresnan's (1973) theory of the degree system.

5.2. Articles and Quantifiers

5.2.1. Semantic Roles of NP Specifiers

There are three different semantic roles that the elements of the NP specifier can play. Without being able to define these semantic roles precisely, we can still make the distinction among the three roles intuitively, at least for the clearest cases, and let the

most general analysis decide in the unclear ones, hopefully to be corroborated by evidence yet to be found. Among the *demonstratives* are the traditional definite articles *the, this, that, these, those,* the interrogatives *which* and *what*, and possibly the indefinite article *a* and the singular *some* (as in *Some man is at the door*). Among the *quantifiers* are *each, every, any, all, no, many, few, much, little,* and other uses of *some*. The third class is *numerals*, including all the usual cardinals plus *a dozen, a couple, a few,* and *a little*. Notice already the confusion that is cropping up, for *some* belongs to different categories depending on its stress, and negative *few* and *little* belong to a different category from nonnegative *a few* and *a little*. These irregularities will be made clearer in the course of the analysis.

A highly plausible semantic constraint can be based on distinguishing these three specifier roles:

(5.1) *Specifier Constraint*
An NP specifier may contain at most one demonstrative, one quantifier, and one numeral.

We will show that (5.1) interacts with the syntactic constraints on specifiers to produce some of the otherwise curious aspects of specifier distribution.

5.2.2. Phrase Structure Rules

Let us turn now to the syntax itself. If N′ is reserved for strictly subcategorized arguments, we are left with two NP specifier positions to account for: the N″ and the N‴ specifier. The N‴ specifier is the position where genitive NPs occur in complementary distribution with demonstratives. We have proposed two possible ways to generate this alternation, in sections 3.7 and 4.5. Assuming the former for simplicity, we have rule (5.2) generating demonstratives in the N‴ specifier.

(5.2) $N''' \rightarrow (\begin{Bmatrix} N''' \\ Art''' \end{Bmatrix}) - N''$

Genitive NPs and demonstratives may be followed by certain quantifiers, but not by others:

(5.3) a. * $\begin{Bmatrix} \text{Fred's} \\ \text{the} \\ \text{those} \\ \text{which} \end{Bmatrix} \begin{Bmatrix} \text{some} \\ \text{each} \\ \text{all} \\ \text{no} \\ \text{any} \end{Bmatrix}$ dwarf(s) b. $\begin{Bmatrix} \text{Fred's} \\ \text{the} \\ \text{those} \\ \text{which} \end{Bmatrix} \begin{Bmatrix} \text{many} \\ \text{few} \\ \text{several} \end{Bmatrix}$ dwarfs

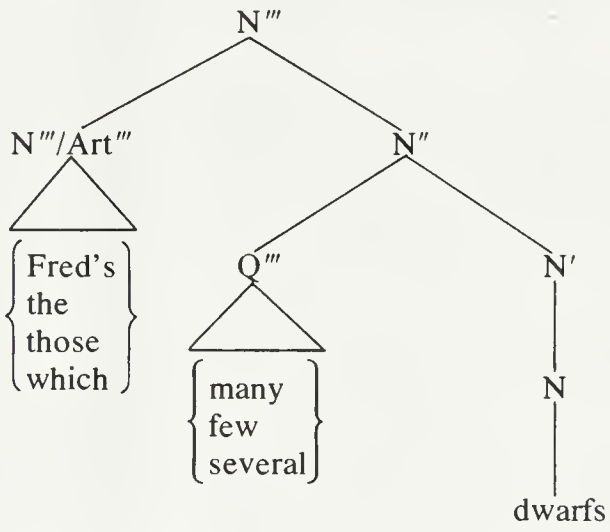
There is no apparent semantic reason for this division: why should **Fred's some apples* mean anything less innocuous than *Fred's many apples*? In order to account for this distribution, then, we will resort to dividing the class of semantic quantifiers into two syntactic categories. Those in (5.3a) will be assigned the category Art, which is generated only by phrase structure rule (5.2). Their inability to cooccur with demonstratives is a syntactic constraint, namely the availability of only one Art position in the

NP specifier. The quantifiers in (5.3b), however, will be assigned to the syntactic category Q, which we will generate in the N″ specifier:

(5.4) $N'' \rightarrow (Q''') - (A''')^* - N' - \ldots$

We choose the N″ specifier rather than the N‴ specifier for this QP for reasons to be discussed in section 6.3. Our analysis thus assigns (5.3b) this structure:

(5.5)



Since some quantifiers are now Arts and some are Qs, the phrase structure component will generate structures in which two quantifiers appear, one in each position, e.g. **no many men, *all several men, *any much wine*. These are ruled out semantically, however, by the Specifier Constraint (5.1), which forbids two (semantic) quantifiers in the same NP specifier. If the second quantifier is replaced by a numeral, so that the Specifier Constraint is satisfied, acceptable constructions like *no three men, all six men,* and *any dozen men* are generated.

For convenience in exposition, we will when necessary refer to the quantifiers in (5.3a) as N‴ quantifiers and those in (5.3b) as N″ quantifiers. The former have the syntactic features [+Subj, −Obj, −Comp, +Det] and the latter [+Subj, −Obj, −Comp, −Det].[1]

[1] A few doubtful cases should be mentioned here. *Little* behaves like an N″ quantifier but its positive *much* behaves like an N‴ quantifier:

(i) {John's / the} {little / *much} food

In most other relevant contexts, though, *much* patterns with *many, few,* and *little,* as will be seen throughout this chapter. It is difficult to know whether to call *much* an Art in this use or to call it a Q and introduce arbitrary restrictions that prohibit it from occurring with Art. More research into the semantics of English quantifiers may help decide which is the appropriate solution. I assume it is a Q.

A second apparent anomaly is that *every* (but not its near relative *each*) can occur after genitives but not after articles in the semi-idiomatic construction (ii).

(ii) We hung on {Fred's / *the} every {move / word}.

It is not clear whether *every* is functioning in its normal syntactic position here (indicating that it is actually a Q) or whether this use as a Q is simply an idiomatic exception. I assume the latter.

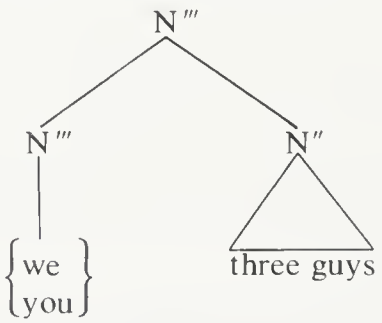
5.2.3. We *and* You

We should mention two little-remarked articles in English which provide evidence for the correctness of the feature system. Postal (1966) is concerned with the connection between the normal pronominal use of *we* and *you* and their use in expressions like *we three guys, you five rascals*. Plural but not singular *you* has this use productively: *I saw you guys/*guy yesterday*; though (inexplicably) singular *you* can be used this way in vocatives: *Come here, you bastard!* There are dialects in which *them guys* is also acceptable.

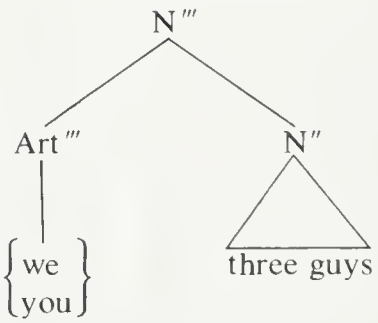
Postal effectively discounts the theory that the pronoun is head of this construction and the other elements are appositives (*we, who are three guys*). Thus the pronoun is apparently in article position. Since Postal has no syntactic features to work with, he resorts to describing the relationship between the two uses by deriving the pronominal use from an underlying *we ones*, so that *we* is always an underlying article. He proposes this mechanism as a general derivation for all definite pronouns.

In the present framework there is an alternative. *We* and you_{plur} can be assigned the feature matrix [+Subj, −Obj, +Det], leaving the feature *Comp* unspecified. Then they can be placed in either N or Art positions. No transformations or abstract underlying forms are necessary. Note that *we* and *you* can then be substituted into the determiner of an NP in two distinct ways:

(5.6) a.



b.



(5.6b) will remain in its underlying form, giving *we/you three guys*. (5.6a), however, will undergo *Poss*-Insertion to become *our/your three guys*.

To recapitulate, this section has shown that there are three semantic sorts of NP specifiers: demonstratives, quantifiers, and numerals. They are divided up among two NP specifier positions, in N‴ and in N″, and are of the categories Art and Q, respectively. Demonstratives always appear in N‴, but some quantifiers appear in N‴ and some in N″. In order to rule out structures with two quantifiers, one in each position, we appealed to the Specifier Constraint, a semantic condition. Finally, the use of *we* and *you* as articles has provided evidence for the feature system.

5.3. The Partitive

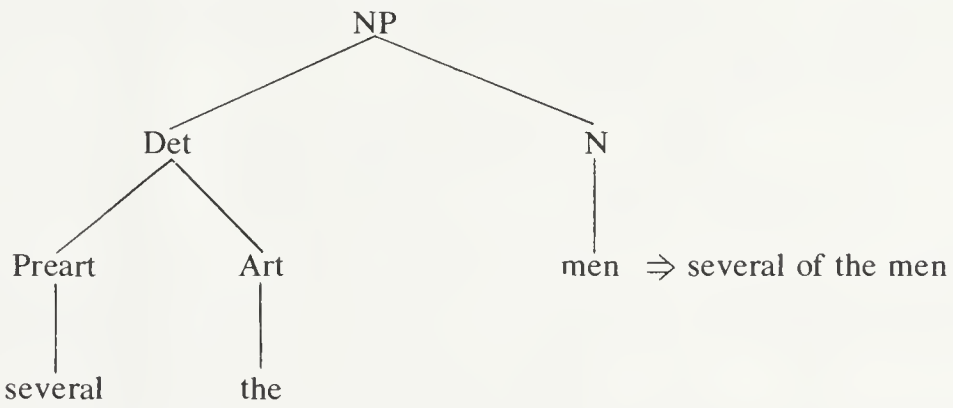
All of the quantifiers mentioned in the previous section occur not only with a simple noun (e.g. *many men, all horses*) but also with a definite *of*-phrase called a partitive

(e.g. *many of the men, all of the horses*). Some of them alter their form when used with a partitive (e.g. *every man, no men* but *every one of the men, none of the men*); this alternation will be discussed in section 5.3.4. All numerals allow a partitive phrase; so do the demonstratives *those* and *which,* and the adjective (?) *other* and superlatives (e.g. *the oldest of the men*). This section will deal with the structure of the partitive construction.[2]

5.3.1. Partitives as N′ Complements

Many sources (including Chomsky in "Remarks") account for the partitive construction by treating the noun following *of* as the head of the construction, and the quantifier preceding *of* as a "prearticle". *Of* is obligatorily inserted if the article is definite:

(5.7)



However, Jackendoff (1968) proposes a different structure, in which *of the N″* is part of the N′ complement. For justification, consider definite *of*-phrases after group nouns, e.g. *a group of the men, a gallon of the wine.* In these phrases, *of the N″* behaves like an ordinary object of the NP, for it always precedes N″ modifiers, as seen in (5.8), and the pronoun *ones* cannot substitute for the head alone, as seen in (5.9).

(5.8) a. a gallon of the wine $\left\{\begin{matrix}\text{in the kitchen}\\ \text{from Sicily}\end{matrix}\right\}$

b. *a gallon $\left\{\begin{matrix}\text{in the kitchen}\\ \text{from Sicily}\end{matrix}\right\}$ of the wine

(5.9) a. groups of the men from Siberia and ones from Japan (= groups of the men from Japan)

b. *groups of the men and ones of the women

Selkirk (1977) further points out that these *of the N″* can undergo Extraposition from NP, which regularly moves N′ complements.

[2] As will be seen, the arguments of Selkirk (1977) have contributed greatly to this section and the next. These are in turn based on Akmajian and Lehrer (1976). I understand that Jean-Claude Milner's recent doctoral dissertation reaches similar conclusions, but I have not had the opportunity to see it.

(5.10) (= Selkirk's (65))

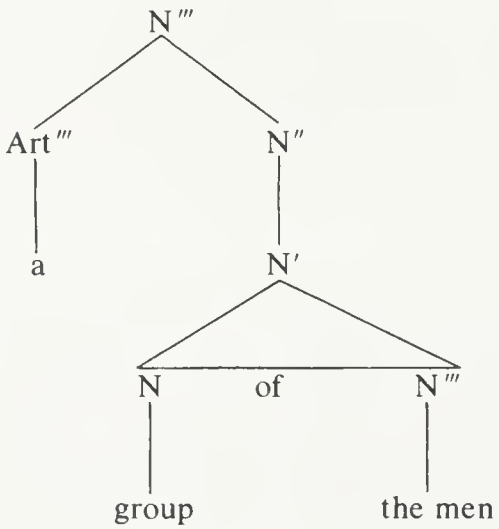
a. A lot of the leftover turkey has been eaten.
   ⇒ A lot has been eaten of the leftover turkey.

b. Only a handful of those questions concerning electromagnetism were asked.
   ⇒ Only a handful were asked of those questions concerning electromagnetism.

c. He gave a rather large number of his books by famous authors to Mary.
   ⇒ He gave a rather large number to Mary of his books by famous authors.

Thus a structure (5.11), which treats *of the N″* as a constituent in the N′ complement, is most plausible for group nouns.[3]

(5.11)



The partitive *of*-phrase serves the same function after quantifiers as after group nouns: it designates a set out of which certain individuals (or a certain subset) is selected. For count nouns, the *of*-phrase must be plural.[4]

Quantifiers and group nouns with partitives share three other properties. First, both prohibit quantification in the *of*-phrase (see section 5.3.3 for more discussion):

[3] Our account of NP objects says that *of N‴* is not a PP, since *of* is a transformationally inserted grammatical formative. The fact that Extraposition moves it like any other PP may be evidence for the position of Jackendoff (1974a), in which NP objects are really PPs with an empty P filled by *of*. I abandoned this theory in chapter 4 because the evidence for it was not as strong as I previously thought, in light of other revisions, and the complication of the object relation was so severe. Henceforth I present trees in a way neutral between the two theories. See also chapter 4, note 13, p. 70.

[4] This needs appropriate modification to deal with the semantics of mass nouns, but in any event the semantic parallelism holds: compare *a gallon of the wine* and *much of the wine*.

(5.12) $*\left\{\begin{array}{l}\text{a group}\\ \text{a number}\\ \text{many}\\ \text{few}\end{array}\right\}$ of $\left\{\begin{array}{l}\text{some men}\\ \text{all men}\end{array}\right\}$

Second, if an N″ quantifier or a group noun is preceded by *the* and has a definite *of*-phrase, a restrictive relative clause is necessary.

(5.13) a. *the group of the men
*the many of the men[5]
b. the group of the men that you met
the many of the men that you met

This relative clause is not attached to the *of*-phrase, as can be seen if the *of*-phrase is preposed:

(5.14) a. *Of the men that you met, $\left\{\begin{array}{l}\text{the group}\\ \text{the many}\end{array}\right\}$ aren't here anymore.
b. Of the men, $\left\{\begin{array}{l}\text{the group}\\ \text{the many}\end{array}\right\}$ that you met aren't here anymore.

Third, as Selkirk (1977) points out, the partitives after quantifiers, like those after group nouns, can undergo Extraposition from NP. The examples in (5.15) are comparable in acceptability to those in (5.10).

(5.15) a. (Not) much has been eaten of the leftover turkey.
b. Only a few were asked of those questions concerning electromagnetism.
c. He gave several to Mary of his books by famous authors.

Each of these arguments raises vexing problems for a theory like (5.7). First, if quantifiers can take relative clauses of their own, "prearticle" must really be some sort of NP. But (5.7) then incorrectly predicts that a relative clause on the prearticle should precede the head noun rather than follow it. Second, the preposing rule which derives (5.14b) would have to extract the head of the phrase, leaving the modifiers behind, an otherwise unprecedented operation. Third, Extraposition from NP would have to move a nonconstituent including the head, instead of a PP in the N′ complement.

By contrast, the theory that the partitive is an N′ complement explains the syntactic parallelisms exhibited in these examples. Parallel (though still unknown) constraints can be stated for the restrictions demonstrated in (5.12)–(5.13). Since relative clauses always follow N′ complements, the proper order of partitive and relative clause in (5.13) follows from the base rules. And since transformations can in general remove complements from their heads, the preposing in (5.14) is within the

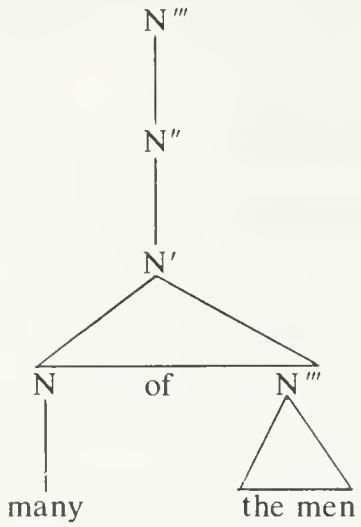
[5] I have no explanation for why *the group of them* is all right but **the many of them* is not.

province of traditional transformations, and Extraposition from NP can apply in normal fashion.

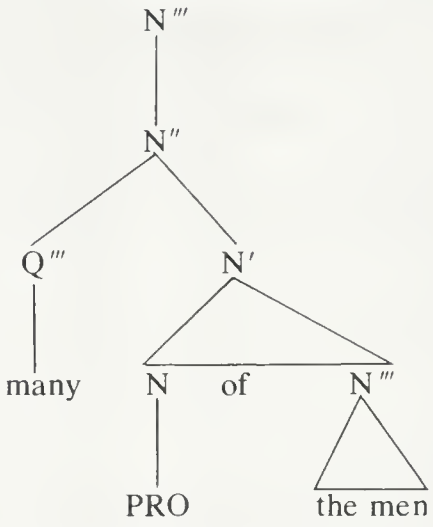
5.3.2. Position of the Quantifier in Partitive Constructions

This theory of the partitive leads to two possible deep structures for *many of the men*.

(5.16) a.



b.



(5.16a) is essentially the structure which Jackendoff (1968) chose, in part because empty nodes were not an available device at the time. Ideally it requires some explanation of why quantifiers cannot be preceded by adjectives; (5.16b) requires no such explanation. We will develop solution (5.16b) here; section 5.5 will show that numerals (but not quantifiers) do occur in construction (5.16a) and in fact can be preceded by adjectives in that use.

To develop the description of (5.16b), we need a projection rule to assign the empty head PRO an interpretation something like "unit(s)" when it is immediately to the right of any of the articles, quantifiers, or adjectives that govern partitives. All of these words must be lexically marked to govern the projection rule.[6]

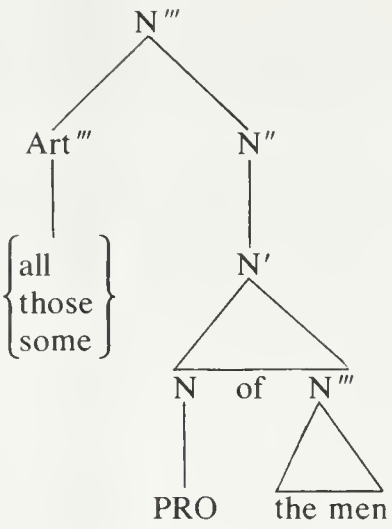
(5.17) *Partitive Projection Rule*

$$\text{PRO}_N \rightarrow \text{UNIT} \,/\, \begin{bmatrix} X \\ +\text{partitive} \end{bmatrix} \text{—}$$

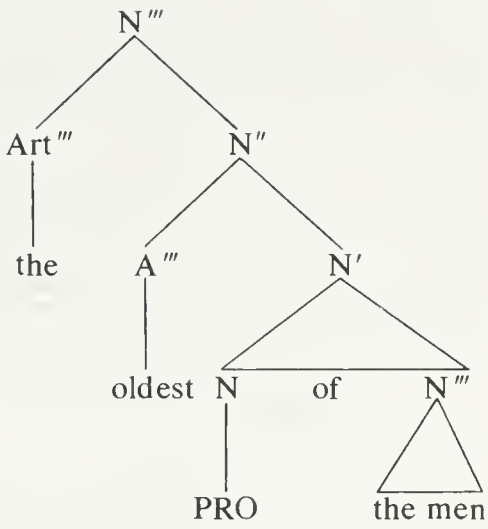
(5.17) interprets PRO in configurations such as (5.18a) and (5.18b), as well as in (5.16b).

[6] The interpretation here given roughly as UNIT must be general enough to apply to mass phrases too, e.g. *some of the wine*, where perhaps the interpretation AMOUNT is more appropriate than UNIT. Wald (1976) develops a theory of "slices", based on nonatomic mereologies, in which mass nouns and plurals can be dealt with homogeneously. The term common to UNIT for count nouns and AMOUNT for mass nouns is Wald's term SLICE.

(5.18) a.

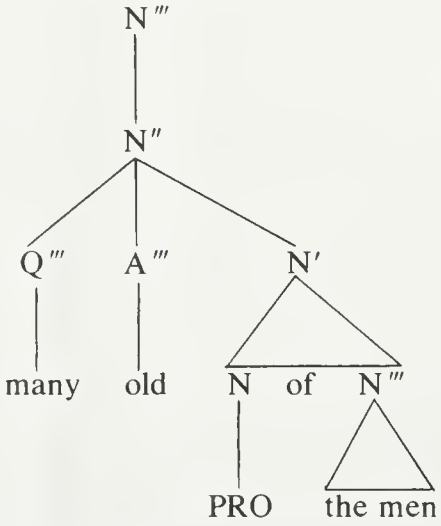


b.

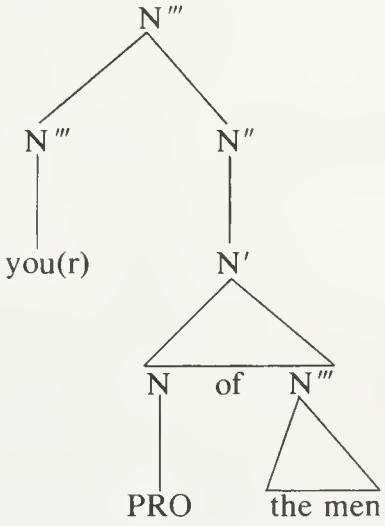


But in (5.19a) PRO cannot be interpreted because it is not adjacent to the partitive word *many*; and in (5.19b) PRO cannot be interpreted because there is no partitive word to govern the rule. Since an uninterpreted PRO produces anomaly, **many old of the men* and **your of the men* are eliminated.

(5.19) a.



b.



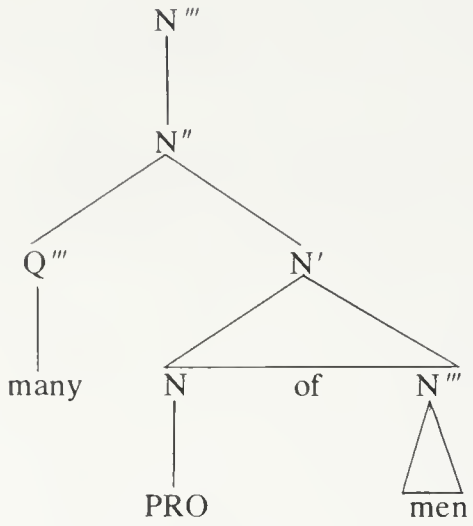
This subsection has shown, then, that (5.16b) is a practicable structure for *many of the men,* given the Partitive Projection Rule. Since we are saving (5.16a) for another construction, (5.16b) is our only choice.[7]

5.3.3. The Partitive Constraint

Given structure (5.16b) for *many of the men*, we must now ask what happens to this structure if the lower NP is indefinite, as in for example (5.20).

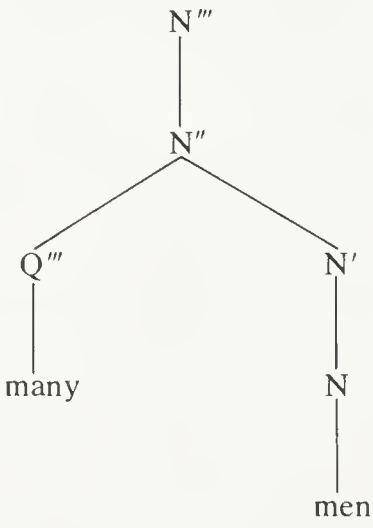
[7] It is worth mentioning that the ungrammaticality of **many old of the men* is a stumbling block to many earlier analyses of the partitive. After trying many alternatives, I have convinced myself that there is no purely syntactic solution. See further discussion in section 5.3.4, particularly footnote 8.

(5.20)



Jackendoff (1968) claims that the *of* in this construction obligatorily deletes, giving another possible source for *many men* besides the simple one of section 5.2, (5.21).

(5.21)



However, Selkirk (1977) argues that this claim makes a number of incorrect predictions. Hence only (5.21) is a correct structure, and some way must be found to eliminate the underlying form (5.20).

One of Selkirk's arguments involves cases where a definite N′ complement is acceptable, but the indefinite without *of* predicted by (5.20) is impossible:

(5.22) a. much of that story/*much story
b. as little of his speech/*as little speech
c. some of the book/*some book (as [−count])

(5.23) a. one of her brothers and sisters/*one brother and sister
b. one of the cattle/*one cattle
c. one of her successive failures/*one successive failure

Conversely, (5.20) predicts that *one single thing* has the impossible definite counterpart **one of the single things*. (5.21), on the other hand, deals with these facts easily, as part of the rule enforcing selectional restrictions and number agreement between quantifiers and their heads.

Second, Jackendoff (1968) claims that *Of*-Deletion corrects the number of the head noun when the quantifier is singular, so that underlying *each of men* becomes *each man* rather than **each men*. But Selkirk points out that it could not correct the number within a relative clause on the head noun. Thus we would incorrectly expect paradigms like (5.24).

(5.24) each of the politicians who were elected/*each politician who were elected

(5.21) would avoid this problem, since *politician* in *each politician* would be underlyingly singular, not plural.

Third, Extraposition from NP would be expected to apply in (5.20) just as it does in (5.16b), yielding sentences like (5.25), parallel to (5.15).

(5.25) a. *Not much has been eaten (of) leftover turkey.
b. *Only a few were asked (of) questions concerning electromagnetism.
c. *He gave several to Mary (of) books by famous authors.

On the other hand, (5.21) claims that *turkey, questions,* and *books* are heads of the subject NPs in (5.25), and that Extraposition from NP should move constituents that are one level further down into the NP. This prediction is correct:

(5.26) a. Only a few questions were asked concerning electromagnetism.
b. He gave several books to Mary by famous authors.

Thus the simpler (5.21), generated by the rules of section 5.2, is the only possible underlying structure for *many men*. Selkirk proposes therefore to filter out (5.20) by a generalization of the constraint illustrated in (5.12) that rules out **many of some men, *many of all men,* and various other specifiers in the partitive NP.

The nature of the constrainst that filters out (5.12) and (5.20) is not (to my knowledge) yet understood. In observational terms, it seems to be approximately (5.27).

(5.27) *Partitive Constraint*
In an *of-N‴* construction interpreted as a partitive, the N‴ must have a demonstrative or genitive specifier.

Thus, (5.27) rules out **many of all men, *many of some men,* and **few of many men,* as well as **many of men,* since they all lack demonstrative or genitive specifiers; but it permits *many of the men, few of the many men,* and *many of his friends*.

Notice that (5.27) refers to semantic functions in the specifier, not to syntactic positions. This argues that (5.27) is part of the semantic component, and further

motivation for it should be pursued in the semantics of specifiers and partitives. I am not prepared to do that here. However, the observation that the Partitive Constraint deals with the semantics of the construction removes any burden from the syntactic component of filtering out the offending structures.

This subsection has thus established that *many men* is generated by the syntax as (5.21); structure (5.20) does not undergo *Of*-Deletion, and though it is syntactically well-formed, it is ruled out by the Partitive Constraint, a part of the semantic component.

5.3.4. Substantivization

To complete the account of the partitive, we need to deal with the phonological change that takes place in certain articles that govern partitives:

(5.28) no men / none of the men
every man / every one of the men
each man / each (one) of the men

It is often thought that the change in form is a phonological coalescence of the article with *one*, so that *none of the men*, for instance, is underlyingly *no one of the men*. But since *none* can refer to mass nouns, as in *none of the rice*, this view is mistaken. We will treat the alternation rather as the addition of an inflectional marker *subst*, added by the following rule:

(5.29) *Substantivization I*
Art - $[_N$PRO] $\Rightarrow$ 1+*subst* - 2
OBLIGATORY

No+*subst* will be spelled as *none*; *every*+*subst* will be spelled as *every one*, etc. For those articles such as *all* which do not change form, the addition of *subst* will not change phonological shape.

This sort of rule changing phonological shape might be thought to violate the Lexicalist Hypothesis. However, since it occurs with only a limited number of morphemes and, unlike nominalization, has no semantic effect, we can regard it as a rule of inflectional morphology, like *Poss*-Insertion and Affix Hopping, which the Lexicalist Hypothesis permits.

In order to show that (5.29) is independently necessary, we will show that its application is not restricted to partitive constructions; other rules which interpret a PRO also create the appropriate configuration. One such rule can be called N′-Anaphora (N̄-Deletion in Jackendoff (1971)). This is the rule operative in the final NP of (5.30).

(5.30) John wanted to read the dossiers of famous linguists, $\left\{\begin{matrix}\text{and}\\\text{but}\end{matrix}\right\}$ he succeeded in reading $\left\{\begin{matrix}\text{Morris's}\\\text{yours}\\\text{these}\\\text{one}\\\text{every one}\\\text{none}\\\text{some}\\\text{many}\\\text{three}\\\text{one}\\\text{*ones}\end{matrix}\right\}$

The anaphoric element in this construction is a pro-N′ PRO which occupies the head of the NP, referring in (5.30) to the N′ *dossiers of famous linguists*. Any specifier is possible before PRO (as long as there is no intervening adjective); the impossibility of *ones* shows that *one* in this construction is the specifier *one,* not the noun.

In the N′-Anaphora construction, the same quantifiers change phonological shape as did in the partitive construction above. In addition, the possessive pronouns change shape: *my, your,* etc. are replaced by *mine, yours,* etc. Again, the change cannot be a result of coalescence with an underlying head *one*(*s*), since N′-Anaphora, unlike *one*(*s*)-Anaphora, can apply to antecedents with mass nouns as heads.[8]

(5.31) John wanted some rice, $\left\{\begin{matrix}\text{and}\\\text{but}\end{matrix}\right\}\left\{\begin{matrix}\text{yours}\\\text{none}\\\text{some}\end{matrix}\right\}$ was moldy.

The proper analysis appears to be an extension of (5.29) to include possessive pronouns.

(5.32) *Substantivization II*

$$\left\{\begin{matrix}\text{Art}\\\begin{bmatrix}\text{NP}\\\text{Pro}\end{bmatrix}+\text{Poss}\end{matrix}\right\} - [_{\text{N}}\text{PRO}] \Rightarrow 1+\textit{subst} - 2$$

OBLIGATORY

[8] The fact that this construction appears with mass nouns goes unnoticed by Postal (1966), who claims to derive all pronouns from underlying articles plus *one*. The solutions of Perlmutter (1970) and Jackendoff (1969c) suffer from the same objection, though they at least note it, with some puzzlement. Sommerstein (1972) has the opposite problem: he inserts *one* if the head has been deleted and there is no determiner or noun left for the number morpheme to attach to. Thus his rules incorrectly derive **My coarse sand is next to your fine one*. In the present theory these objections (raised also by Siegel (1974)) do not arise, because *one* and PRO are distinct anaphoric elements in NPs, not transformationally related.

Substantivization also applies when PRO is interpreted by the rule of N′- Gapping (see Jackendoff (1971)). This rule differs from N′-Anaphora in that it can apply only if the antecedent is directly conjoined to the gapped NP, and in that it permits N′ complements to be left behind.

(5.33) a. *Joe's destruction of the pictures was appalling, and Fred's of the records was even worse.
(N′-Anaphora: NPs not conjoined, but N′ complement left behind)
b. Joe's destruction of the pictures and Fred's of the records appalled everyone.
(N′-Gapping: NPs conjoined)

Again, the missing head can be regarded as a PRO; so, for example, the subject of (5.33b) is *Joe's destruction of the pictures* and *Fred's PRO of the records*.

N′-Gapping is of considerable interest to the theory of syntactic distinctive features, for it has properties very similar to the well-known rule of Gapping in sentences. These similarities are discussed in Jackendoff (1971), and I have nothing to add here. What is germane to the present discussion is that Substantivization applies in N′-gapped constructions as well, yielding *yours* and *none* in (5.34), for example.

(5.34) a. Joe's destruction of the pictures and yours of the records
b. few stories about Fred and none about Bill

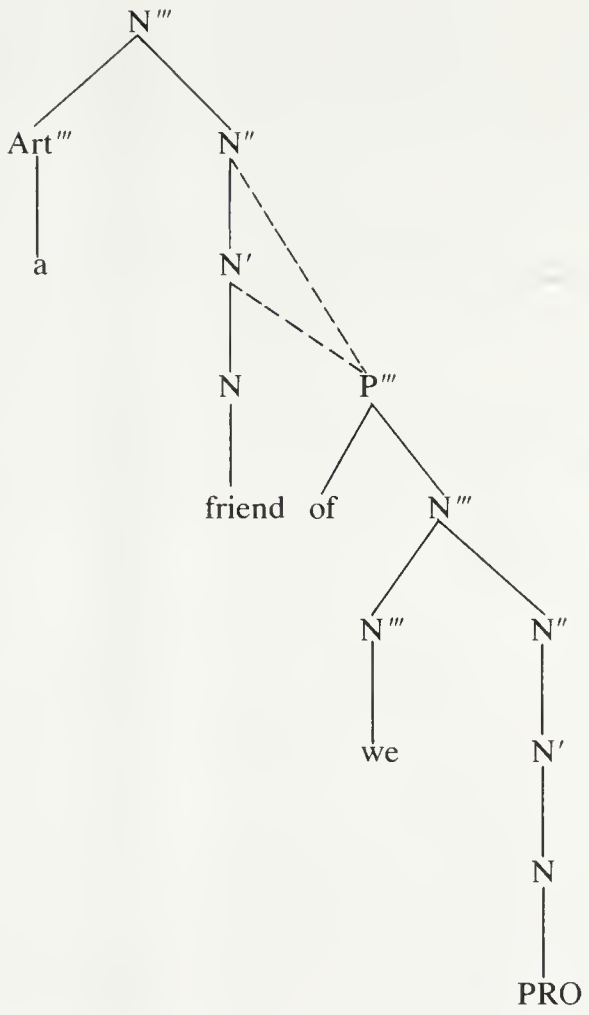
A fourth application of Substantivization is in "postposed genitive" constructions such as *a friend of ours*. This *of*-phrase occurs after certain N′-complements but appears to precede N″-complements:

(5.35) a. a picture of Bill of John's
?*a picture of John's of Bill
b. a picture of John's with a gold frame[9]
?*a picture with a gold frame of John's

The source for this construction proposed in Jackendoff (1969c), which I see no reason to abandon, comes out in the present framework as (5.36).

[9] (5.35b) shows that this *of*-phrase cannot be produced by extraposing the genitive phrase from the determiner to the end of the NP, as some have claimed, including Siegel (1974).

(5.36)



We in (5.36) undergoes *Poss*-Insertion and Substantivization to become *ours*.

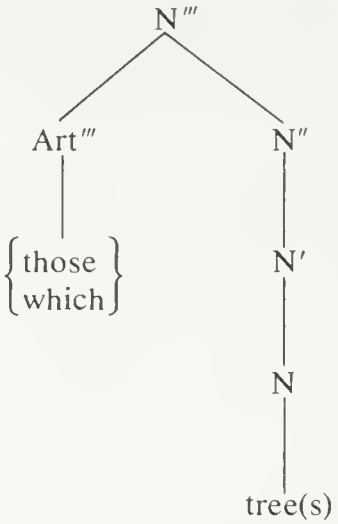
This source preserves the generalization that genitive NPs arise only in determiner position. However, the underlying position of the postposed genitive is not in the determiner of the main NP, from which it is then extraposed. Instead, it is in the determiner of a complement NP rather like a partitive. (5.36) is thus similar to the structure of *one of our friends*, except that *friend* is in the main NP rather than in the partitive. PRO must be interpreted by a special projection rule, devised for this construction. Again, an underlying *one*(*s*) instead of PRO is impossible, because of mass constructions such as *some information of theirs*.

We have thus shown that Substantivization is a well-motivated rule of English, applying to N''' specifiers followed by PRO, where PRO can be interpreted by the Partitive Projection Rule, N'-Anaphora, N'-Gapping, or the Postposed Genitive Projection Rule. Thus, the alteration of the form of *no, every,* etc., and the use of a PRO head are not peculiar to partitives, but rather are independently necessary. These results thus confirm the analysis of the partitive given above.

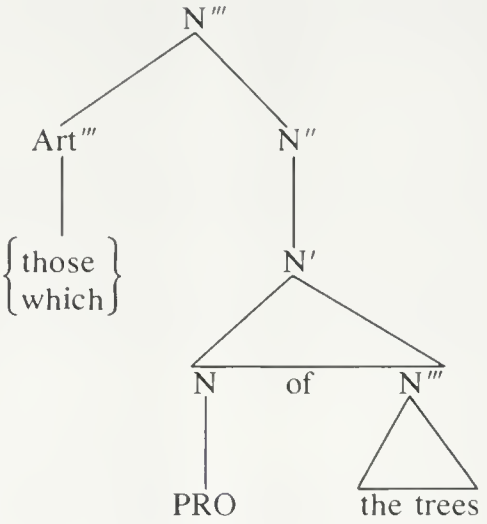
To sum up the results of this entire section on the partitive, we have shown that quantifiers and certain demonstratives occur in the following grammatical structures:

(5.37) *Demonstratives*

a.

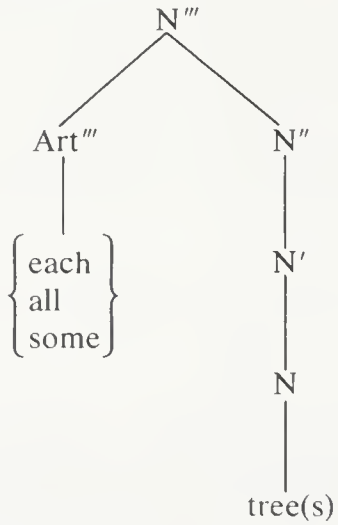


b.

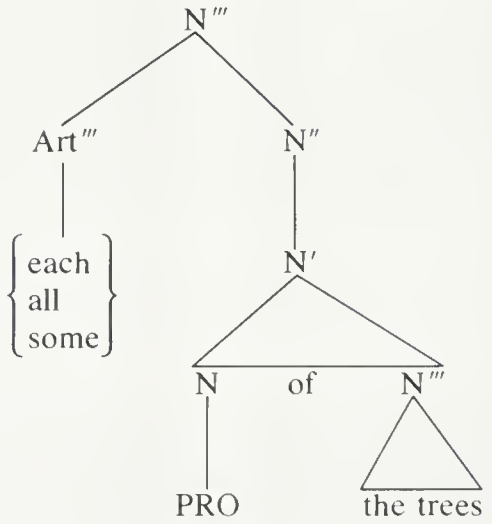


(5.38) *N‴ Quantifiers*

a.

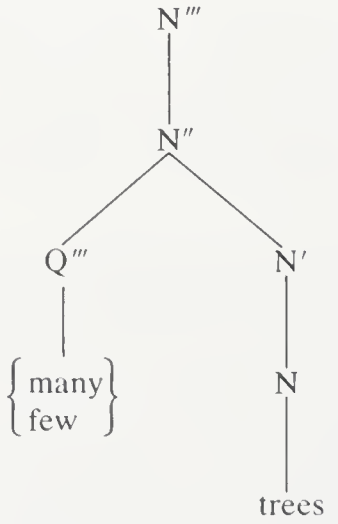


b.

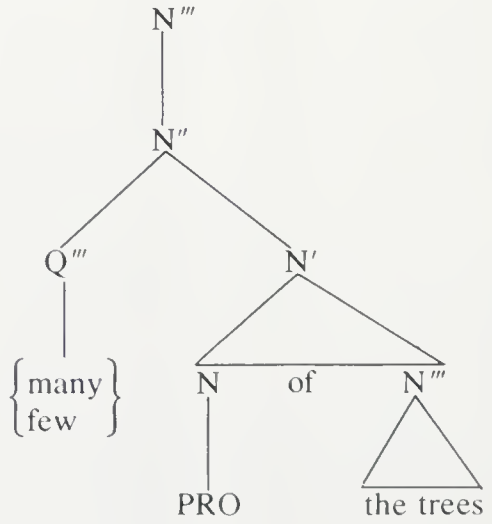


(5.39) *N″ Quantifiers*

a.



b.



Each of these specifiers is marked [+Partitive] and thus in the (b) cases PRO can be interpreted by the Partitive Projection Rule; the proper specifiers in the subordinate NPs are guaranteed by the Partitive Constraint. Certain of the N‴ specifiers in this construction change their form in case (b) through the application of the rule Substantivization.

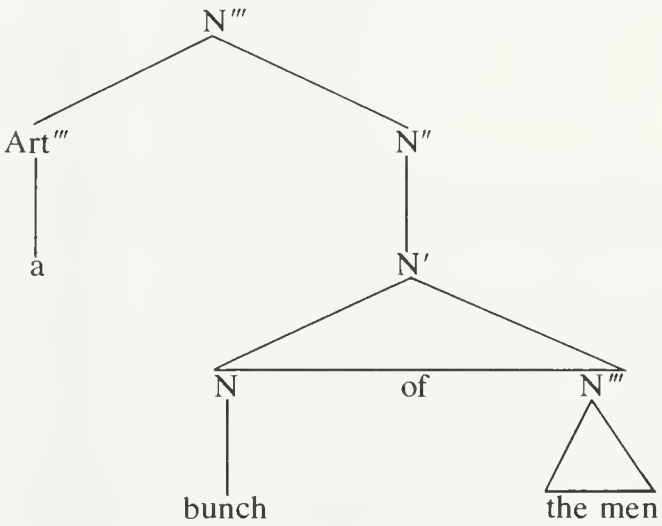
This completes our discussion of the partitive. We now turn to two other constructions in the NP specifier, pseudopartitives and numerals.

5.4. Pseudopartitives

5.4.1. Arguments for the Construction

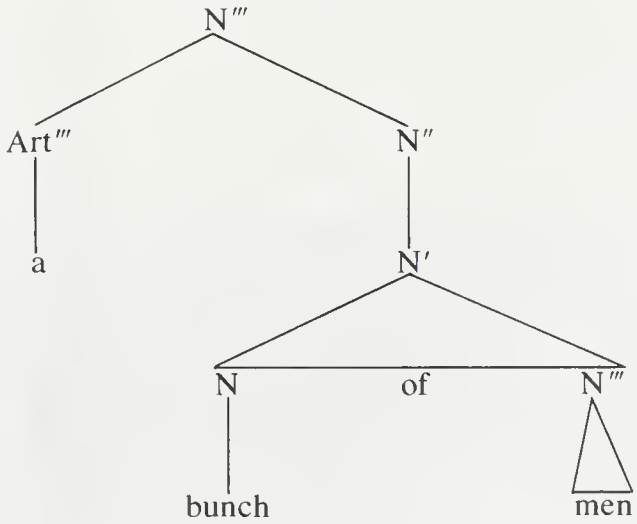
Section 5.3.1 argued that group nouns followed by partitives such as *a bunch of the men* have the following structure:

(5.40)



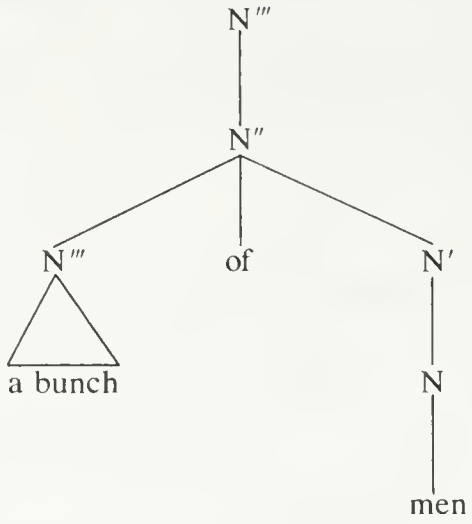
ackendoff (1968) assumes that when the lower NP is indefinite, as in *a bunch of men*, the structure is essentially the same:

(5.41)



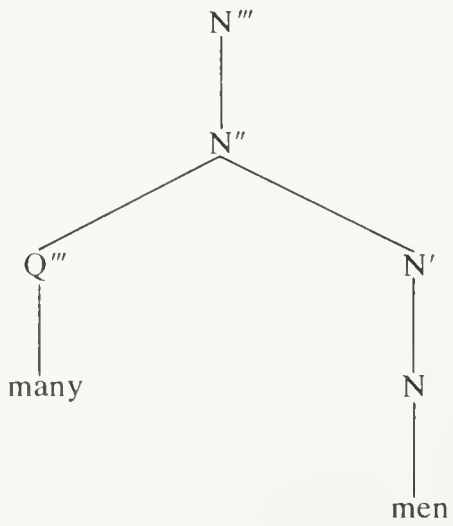
However, Selkirk (1977) argues that *a bunch of men* has a different structure, (5.42), which she calls the pseudopartitive. Her arguments are quite similar to those for quantifiers given in section 5.3.3.

(5.42)



In this structure, the grammatical formative *of* is a daughter of N″. Other than the *of*, (5.42) is parallel to the structure of *many men*, (5.21).

(5.21)



Before presenting three of Selkirk's arguments, let us see how a solution like (5.42) affects the grammar. Ignoring the *of* for a moment, we see that there is now a choice between Q‴ and N‴ in the N″ specifier. In a theory without syntactic distinctive features, this would involve a complication of the grammar. But since, in the present theory, Q‴ and N‴ differ only in the feature *Comp*, the choice between Q‴ and N‴ makes the grammar actually one feature simpler than if Q‴ alone were possible. The rule for N″ thus reads like this:

$$(5.43)\quad N'' \rightarrow \left(\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}'''\right) - (A''')^* - N' - \ldots$$

Chapter 6 will elaborate on this rule considerably.

Returning to the *of*, we have the choice of introducing it in underlying structure (with suitable use of angle brackets in (5.43) to make sure it appears only in the N''' case), or by a simple local transformation. Section 6.2 will give one argument that it should appear as a specified grammatical formative in the phrase structure rule; we will assume this position in advance, though little depends on it.

Selkirk's first argument for (5.42) concerns the Partitive Constraint of section 5.3.3. She points out that group nouns with partitives are apparently subject to the Partitive Constraint, since the specifier of the NP after *of* must contain a demonstrative or genitive:

(5.44) *a bunch of some men
*a gallon of much wine
*a number of all objections

(5.45) a bunch of the men
a gallon of our wine
a number of those objections

Since one of the things the Partitive Constraint rules out is a null determiner in the partitive phrase, it rules out (5.41), unless (5.41) is an otherwise unmotivated exception. On the other hand, (5.42) contains no NP interpreted as a partitive, so the Partitive Constraint does not apply to it.[10]

Selkirk's second argument concerns the application of Extraposition from NP. Recall that in a definite partitive with structure (5.40), Extraposition from NP moves the *of*-phrase:

(5.10) a. A lot of the leftover turkey has been eaten.
⇒ A lot has been eaten of the leftover turkey.
b. Only a handful of those questions concerning electromagnetism were asked.
⇒ Only a handful were asked of those questions concerning electromagnetism.

[10] Note that there is another construction, exemplified by *a group of three men*, in which the *of*-phrase does not contain a demonstrative specifier. But this *of*-phrase does not have a partitive interpretation, since *a group of three men* means not a group taken out of three men, but a group consisting of three men. Hence the Partitive Constraint does not apply, and a nondemonstrative specifier is permitted. The possibility of such a "consistive" interpretation is what makes *a pot of steel* ambiguous: one reading is like *a group of men*, has structure (5.42), and denotes a certain amount of steel; the other reading is like *a group of three men*, has structure (5.41), and denotes a pot made out of steel. In fact, the consistive reading seems to *require* a nondemonstrative specifier and thus is mutually exclusive with the partitive reading. The Partitive Constraint may thus in fact be part of the structural description of the semantic rule that differentiates between these two readings of the N' complement.

c. He gave a rather large number of his books by famous authors to Mary.
⇒ He gave, a rather large number to Mary of his books by famous authors.

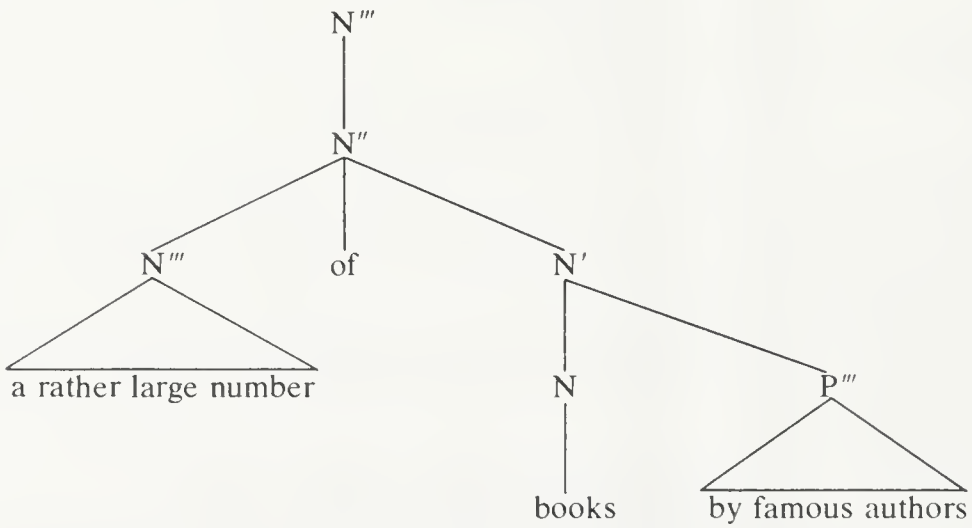
If (5.41) were the structure for *a bunch of men,* we would expect Extraposition from NP to apply in precisely parallel fashion. However, it does not:

(5.46) a. ?*A lot has been eaten of leftover turkey.
b. ?*Only a handful were asked of questions concerning electromagnetism.
c. ?*He gave a rather large number to Mary of books by famous authors.

On the other hand, structure (5.42) correctly predicts the unacceptability of (5.46), since it claims that, for example, *leftover turkey* is not an N‴ in the N′ complement.

Furthermore, structure (5.42) predicts a different result. It requires *a rather large number of books by famous authors* to have this structure:

(5.47)



Thus, if Extraposition from NP applies to anything in this structure, it should apply to *by famous authors*. This prediction is correct:

(5.48) a. Only a handful of questions were asked concerning electromagnetism.
b. He gave a rather large number of books to Mary by famous authors.

Hence (5.42), not (5.41), predicts the correct application of Extraposition from NP.

A third argument concerns the interpretation of relative clauses. Selkirk points out that in (5.49a) the relative clause can apply to either *a number* or *those daffodils,* and the interpretations are quite distinct, whereas in (5.49b) there is only one interpretation.

(5.49) a. She bought him a number of those daffodils, only two of which were faded.
b. She bought him a number of daffodils, only two of which were faded.

This difference is easily accounted for if *a number of those daffodils* includes two N‴ to which relative clauses can be attached, but *a number of daffodils* contains only one such N‴. This is predicted by the theory that *a number of those daffodils* has structure (5.40) and *a number of daffodils* has structure (5.42). It is not predicted by the theory that *a number of daffodils* has structure (5.41).

We conclude therefore that *a group of men* is not a true partitive. The syntax does in fact generate the string ambiguously, with partitive structure (5.41) and pseudopartitive structure (5.42), but (5.41) is ruled out semantically by the Partitive Constraint.

5.4.2. *Pseudopartitive Specifier Constraints*

Since pseudopartitives contain an N‴ in the N″ specifier, there is nothing to stop this subordinate N‴ from having its own specifiers. Thus the phrase structure rules cannot be prevented from generating such ungrammatical combinations of specifiers as these:

(5.50) $$*\Big[_{N'''}\ \Big\{\begin{matrix}\text{John's}\\ \text{all}\\ \text{some}\\ \text{those}\end{matrix}\Big\}\ \Big[_{N''}\ \Big[_{N'''}\ \Big\{\begin{matrix}\text{John's}\\ \text{a}\\ \text{every}\\ \text{that}\end{matrix}\Big\}\ \text{group}\Big]\ \text{of}\ \Big[_{N'}\ \text{trees}\Big]\Big]\Big]$$

It is clear that only one N‴ specifier may be allowed to precede the group noun. The problem is whether the permitted specifier is always that of the matrix N‴, always that of the subordinate N‴, or sometimes one and sometimes the other.

First consider articles. It appears that when a group noun in a pseudopartitive is preceded by an Art‴, it is always the Art‴ of the group noun, not that of the matrix N‴. To see this, notice that when the group noun is singular and the head noun is plural, an Art before the group noun must be one that permits a singular head.

(5.51) a. $\Big\{\begin{matrix}\text{a}\\ \text{every}\\ \text{this}\\ \text{that}\end{matrix}\Big\}$ group of people

b. $*\Big\{\begin{matrix}\text{all}\\ \text{we}\\ \text{you}\\ \text{these}\\ \text{those}\end{matrix}\Big\}$ group of people

On the other hand, if the group noun is plural, the Art is one that permits a plural head.

(5.52) $\Big\{\begin{matrix}\text{all}\\ \text{we}\\ \text{you}\\ \text{these}\\ \text{those}\end{matrix}\Big\}$ groups of people

Since the number of the group noun controls the choice of articles, the assumption that the article is a specifier of the group noun keeps the statement of these particular agreement constraints the simplest.

Similarly, where quantifiers precede a group noun in a pseudopartitive, they appear to quantify over instances of the group noun rather than over the head noun. For example, *all groups of people* quantifies over instances of *groups,* not of *people.* Compare this to *all three people*, which quantifies over instances of people (of whom there are three), and which we have assumed has *all* in the N‴ specifier of *people.* If this semantic difference is to be represented syntactically, *all* in *all groups of people* must be a specifier of *groups.*

Thus the evidence seems to indicate that an NP containing a group noun in its N″ specifier cannot have an Art‴ in its N‴ specifier. Can it have an N‴ instead? The answer seems to be yes. To see this, consider group nouns such as *number* and *quantity* which place restrictions on their specifiers. When used in genuine partitives, they cannot be used with a possessive:

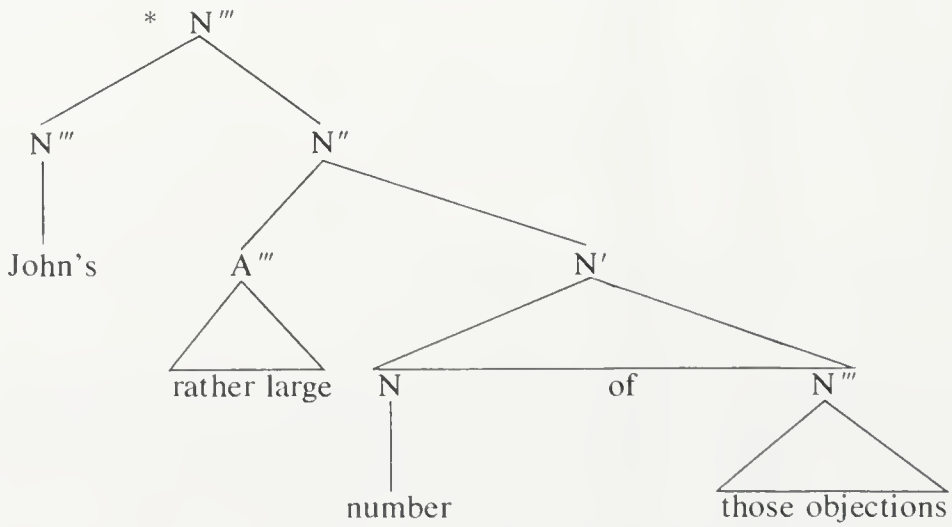
(5.53) *John's rather large number of those objections
*Fred's substantial quantity of that wine

But possessives can precede them in pseudopartitives:

(5.54) John's rather large number of objections
Fred's substantial quantity of wine[11]

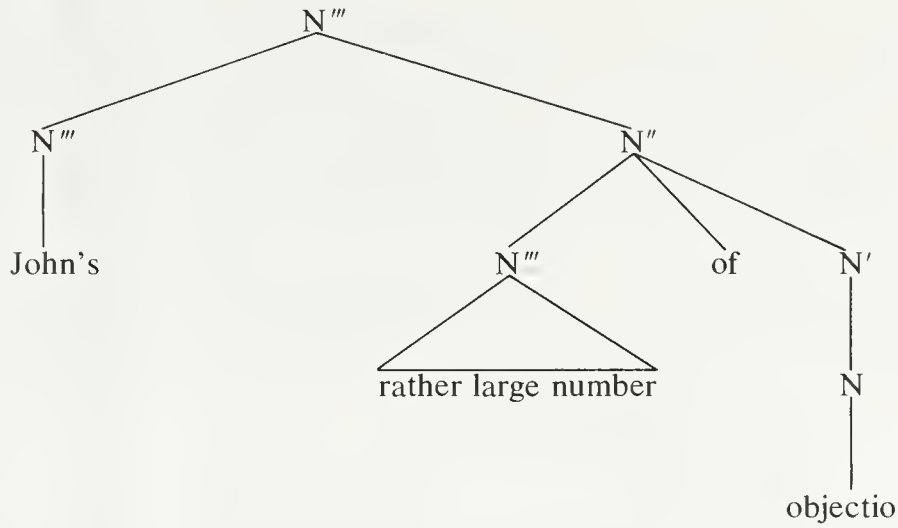
This difference can be explained if *number* and *quantity* exclude possessives from their N‴ specifiers: in the true partitives (5.53), the possessives will have to appear improperly in the specifiers of *number* and *quantity,* as in (5.55a), but in (5.54) they may be in the specifier of *objections* and *wine*, as in (5.55b).

(5.55) a.



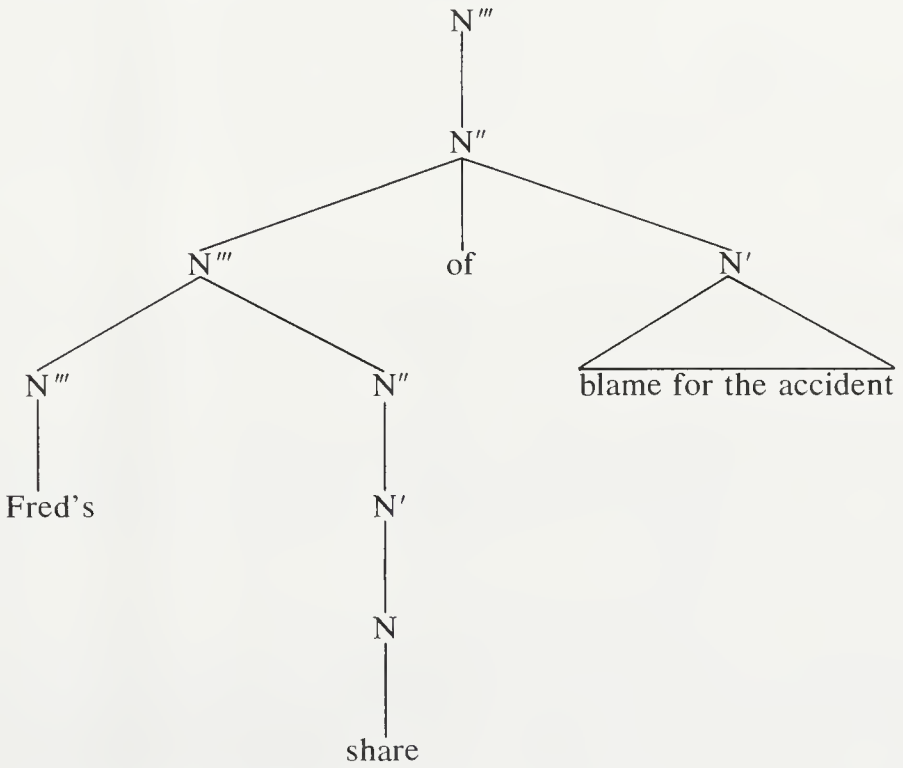
[11] I have no explanation for the fact that (5.54) becomes much worse if the adjectives are omitted. Note that this paradigm is parallel to the one with N″ quantifiers, since we have **John's many of those objections,* like (5.53), but *John's many objections,* like (5.54).

b.



Thus it is possible to have a genitive in the matrix N‴ specifier. There also seem to be cases where a genitive in the group-noun specifier is necessary. The noun *blame* does not permit a genitive specifier, since, for example, only *the blame for the accident* is possible, not **Fred's blame for the accident*. Nonetheless, the pseudopartitive construction with *blame* does permit a genitive, since *Fred's share of blame for the accident* is possible. The simplest account of this seems to be to attribute the genitive to the specifier of the group noun:

(5.56)



Thus, in pseudopartitives, only the group noun may contain an Art‴, but apparently either the group noun or the head may contain a genitive. Other than one

combination which will be useful in the next section, I leave precise statements of these curious restrictions for future research.

The one combination of interest is this: just in case the matrix NP has an N''' specifier, as in (5.54), the group noun is left with no surface specifier at all. This is a little strange, since singular count nouns like *number* and *quantity* normally require some specifier. One way to make the situation a little better is to claim that the underlying form of (5.54) is *John's* [$_{N'''}$*a rather large number*] *of objections, Fred's* [$_{N'''}$*a substantial quantity*] *of wine*, with the indefinite article as the specifier of the group noun. Then the following local transformation would delete the indefinite article:

(5.57) *Pseudopartitive* a-*Deletion*
$[_{N'''}N''' - a - X] \Rightarrow 1 - \phi - 3$
OBLIGATORY

Thus we can claim that the group nouns in (5.54) do have the specifier they require in underlying form, but it is deleted.

This solution does not deal with the possibility of other specifiers of the group noun cooccurring with a matrix genitive. I assume they are eliminated by a surface constraint which is necessary anyway to eliminate such combinations as two genitives in a row. Thus *a* is the only group noun specifier cooccurring in underlying structure with matrix genitives, because it is deleted and hence the surface constraint is satisfied.

To sum up our discussion of pseudopartitives, we have argued that the N'' specifier contains an N''' position which alternates with the Q''' position there. It may be occupied by group nouns which are followed by the specified grammatical formative *of*. Furthermore, the group noun may contain its own specifiers, and there is a complicated set of constraints coordinating the group-noun specifier and the matrix N''' specifier.

5.5. Numerals

We have so far neglected the syntax of numerals. They come in two varieties, which I will call *cardinals* and *seminumerals*. Cardinals are words like *three* and *seven*, which need not be preceded by an article; seminumerals are words like *dozen* and *hundred*, which require an article before them.

The grammar of numerals is somewhere in between that of quantifiers and that of group nouns. This section will argue that numerals are nouns which trigger certain local transformations whose effect is to make the superficial behavior of numerals closer to that of quantifiers.

5.5.1. Comparison to Quantifiers and Group Nouns

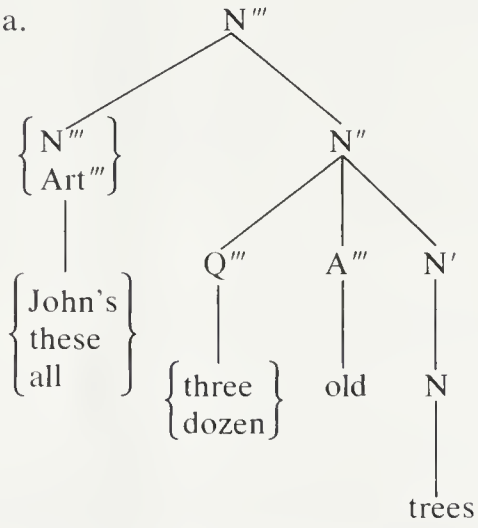
This subsection will demonstrate the similarities of numerals to group nouns and quantifiers, by showing how numerals behave in paradigms discussed in previous sections.

First, like quantifiers and pseudopartitive group nouns, numerals appear between N''' specifiers and A'''s.

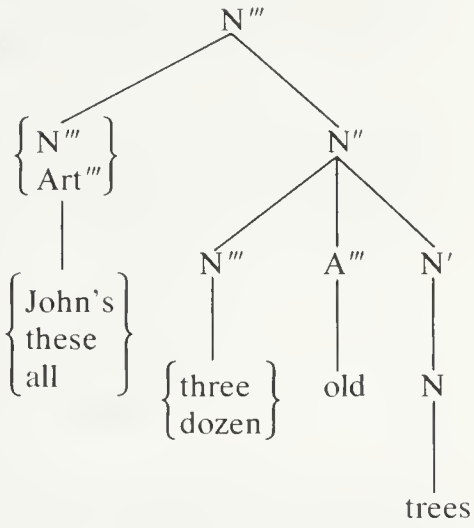
(5.58) $\left\{\begin{matrix}\text{John's}\\\text{these}\\\text{all}\end{matrix}\right\}\left\{\begin{matrix}\text{three}\\\text{dozen}\end{matrix}\right\}$ old trees

This suggests that they are in the N″ specifier, either as quantifiers (5.59a) or as group nouns (5.59b).

(5.59) a.

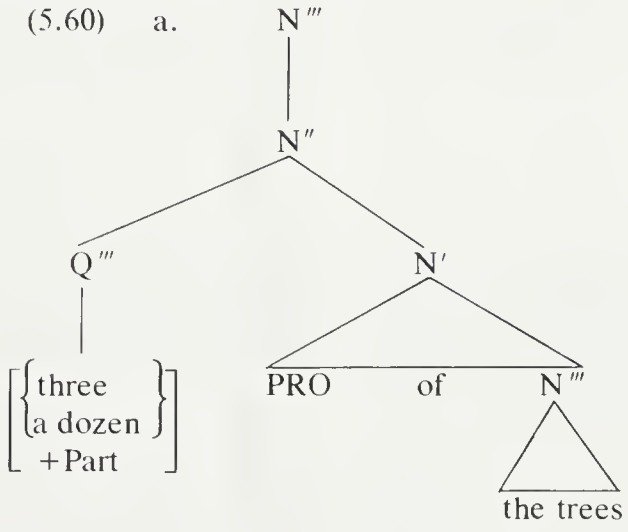


b.

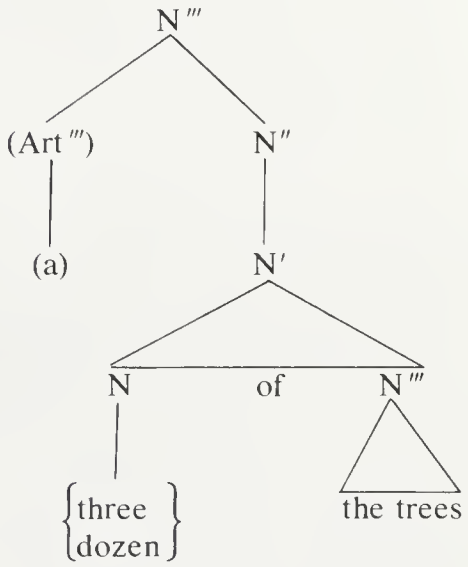


Numerals also occur in partitive constructions, with structure (5.60a) if they are quantifiers and (5.60b) if they are nouns.

(5.60) a.



b.



The Partitive Constraint applies to these constructions,

(5.61) $*\left\{\begin{matrix}\text{three}\\\text{a dozen}\end{matrix}\right\}$ of $\left\{\begin{matrix}\text{some men}\\\text{all men}\end{matrix}\right\}$

as does the constraint on quantifiers and group-noun partitives that if they are definite, a relative clause is required.

(5.62) a. *the {three / dozen} of the trees

b. the {three / dozen} of the trees that weren't cut down

Extraposition from NP can be applied to the partitive:

(5.63) {Three / A dozen} weren't cut down of the trees you planted.

By contrast, in the simple construction (5.58) (as in the parallel cases with quantifiers and group nouns), Extraposition from NP applies not to the noun being counted, as in the partitive, but to a PP one level further down:

(5.64) a. *{Three / A dozen} weren't cut down (of) trees you planted.

b. {Three / A dozen} pieces were eaten of the leftover turkey.

Thus numerals parallel the behavior of quantifiers and group nouns. If they are quantifiers, the simple use (5.59a) is straightforward, and in the partitive use (5.60a) they govern the Partitive Projection Rule. If they are nouns, the partitive use (5.60b) is straightforward, and the simple use (5.59b) is like a pseudopartitive (though lacking the *of*). The next question we will address is which of these analyses is correct. So far the analysis of numerals as quantifiers looks more promising, since the paradigm (5.59a)–(5.60a) is just like that of *many* and *few*. However, the next subsection will argue that they are nouns.

5.5.2. *Numerals as Nouns*

The argument that numerals are nouns is based on their specifier system. The quantifiers *many*, *few*, and *little* can be preceded by degree words such as *so*, *as*, *too*, and *how*; chapter 6 will explore this specifier system in considerable detail. Numerals, of course, can never be preceded by degree words:

(5.65) *{so / as / too / how} {three / dozen}

On the other hand, nouns such as *group*, *bunch*, and *number* are preceded by typical nominal modifiers such as articles and adjectives; if numerals are nouns, they should have such specifiers too.

The fact that the seminumerals *dozen* and *hundred* always require a nominal specifier (most often *a*) is a hint that we should look for other evidence that numerals are nouns. The clearest evidence comes from a construction little noted in the literature

(although it appears in Jespersen), in which even cardinals are preceded by adjectives and by the indefinite *singular* article:

(5.66) a beautiful two weeks
a dusty four miles of road
another whole seventeen pages

The phrases are close syntactic parallels to group noun pseudopartitives like (5.67).

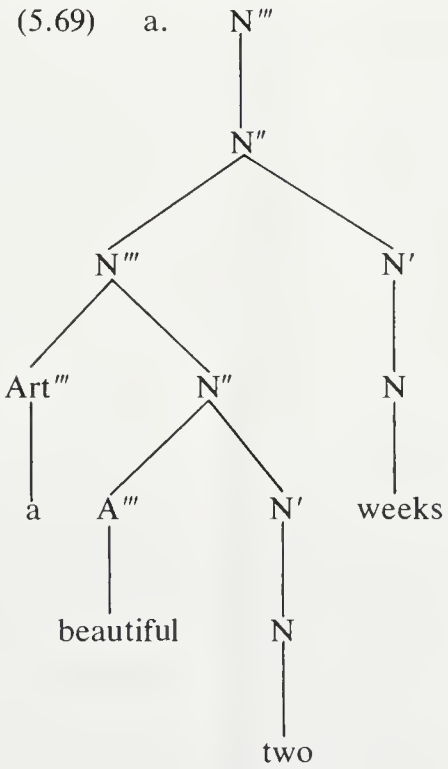
(5.67) a tremendous group of people
a useless couple of days
another whole bunch of bananas

A similar parallel obtains with true partitives:[12]

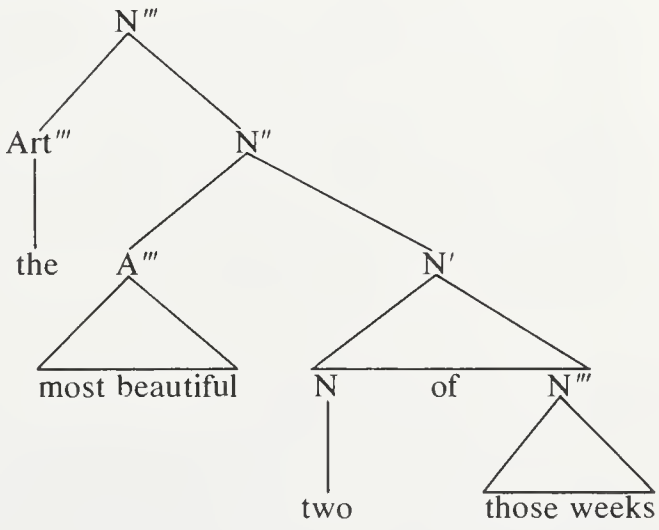
(5.68) a. the most beautiful two of those weeks
a dusty four of those many miles of road
another whole seventeen of SPE's pages
b. a tremendous number of the senators
a useless couple of the remaining days
another whole bunch of those bananas

Since quantifiers cannot be preceded by adjectives, but nouns certainly can, it appears most expedient to treat the numerals in these constructions as nouns, and therefore to assign (5.66) and (5.68a) the following structures:

(5.69) a.



b.



[12] I have no explanation for why good examples of partitives like (5.68a) are so hard to find.

If numerals are indeed nouns, the next step in the analysis must be to account for the differences between numerals and group nouns; the next subsection will deal with this problem. But first it is worth mentioning two special cases, *one* and *few*.

We have repeatedly alluded to two different morphemes spelled *one*, a numeral and a pronoun. The numeral can govern a partitive and cannot be pluralized; the pronoun allows no N′ *of*-complement and has the plural *ones*. If the numeral *one* were like other numerals in being a noun, we would expect, parallel to (5.66), the ungrammatical **a beautiful one day*. In order to prevent this, we will claim that the numeral *one*, unlike *two*, *three*, etc., is a quantifier and hence takes no nominal specifiers. The pronoun *one*, on the other hand, takes nominal specifiers but is not a group noun, so it cannot occur in structures (5.69a,b) either.

Section 5.2.1 hinted at the distinction between *few* and *a few*. We can now spell it out more clearly. *A few*, because of its indefinite article, can be identified as a seminumeral, hence a noun. As confirmation, we note that it can be preceded by adjectives, e.g. *a miserable few people showed up*. On the other hand, *few* is a quantifier, as can be seen by its ability to take degree words (*so few*, *as few*, etc.) and adverbs (e.g. *We met miserably few people*). Similarly, *a little* is a seminumeral and *little* is a quantifier.[13]

We now turn to the adjustment rules which distinguish cardinals, seminumerals, and group nouns from each other.

5.5.3. Three Adjustment Rules

First let us distinguish cardinals from seminumerals. Seminumerals are like group nouns in that they are always preceded by a specifier, in the unmarked case the article *a*. (5.66) and (5.68) showed that cardinals also can occur with articles if an adjective intervenes. To assimilate cardinals to the normal case of group nouns and seminumerals, we will assume that cardinals too always have an article in underlying structure, but that it is deleted by the following lexically governed local transformation:

(5.70) *Cardinal* a-*Deletion*

$$a - \begin{bmatrix} \text{N} \\ +\text{Card} \end{bmatrix} \Rightarrow \phi - 2$$

OBLIGATORY

According to this analysis, the underlying structure of *two weeks* is *a two weeks*, parallel to *a dozen weeks* and *a bunch of weeks* (we deal with *of* in a moment); (5.70) deletes *a* before *two*. *A beautiful two weeks*, parallel to *a beautiful dozen weeks* and *a beautiful bunch of weeks*, will not be affected by (5.70) because of the adjective intervening between the article and the numeral. Similarly, in *That two weeks we spent in Bermuda was awful*, the singularity of the article shows it is a specifier of the

[13] One exception to this clear distinction is *a very few*, which has both a nominal and a quantifier specifier. I have no explanation.

numeral; since the article is not *a*, it does not delete. (5.70) thus accounts for the difference between cardinals and seminumerals.

Next we must distinguish numerals from group nouns. As pointed out in section 5.5.1, they differ in that group nouns but not numerals are followed by *of* when they occupy the N″ specifier. Under the assumption that the pseudopartitive *of* is an underlying specified grammatical formative (a point that will be somewhat justified in section 6.2), the simplest way to account for this difference is to posit another local transformation, governed by all numerals:[14]

(5.71) *Numeral* of-*Deletion*

$$[_{N''}X - \begin{bmatrix} N \\ +Num \end{bmatrix} - of - Y - N' - Z] \Rightarrow 1 - 2 - \phi - 4 - 5 - 6$$

OBLIGATORY

These two rules yield the following derivations:

(5.72)

	a bunch of weeks	a [hundred, +Num] of weeks	a [six, +Num, +Card] of weeks
apply (5.71)	↓ n.a.	↓ a hundred weeks	↓ a six weeks
apply (5.70)	↓ n.a.	↓ n.a.	↓ six weeks

There are a few lexical doublets that show how superficial the presence of the feature [+Num] is. In my dialect, the word *couple* alternates freely between group noun and seminumeral status, since I can say either *a couple of trees* or *a couple trees*.[15] In the majority dialect, *dozen* is a seminumeral; but there are dialects which allow *a dozen of eggs*, where *dozen* is a group noun. Though *hundred*, *thousand*, and *million* are seminumerals, *hundreds*, *thousands*, and *millions* are group nouns; thus we have the contrast of *six million men* (numeral) vs. the somewhat fusty *six millions of men* (group noun). In the present analysis, these differences follow solely from the presence or absence of the exception feature [+Num].

These two local transformations, Cardinal *a*-Deletion and Numeral *of*-Deletion, are sufficient to account for the differences between numerals and group nouns observed in

[14] Under any other assumptions about how the pseudopartitive *of* is introduced, the exception feature [+Num] will play a similar role in preventing *of* from occurring in the surface. I assume there is a separate idiosyncratic *Of*-Deletion rule applying in the partitives *all of the men*, *both of the men* to form *all the men*, *both the men*.

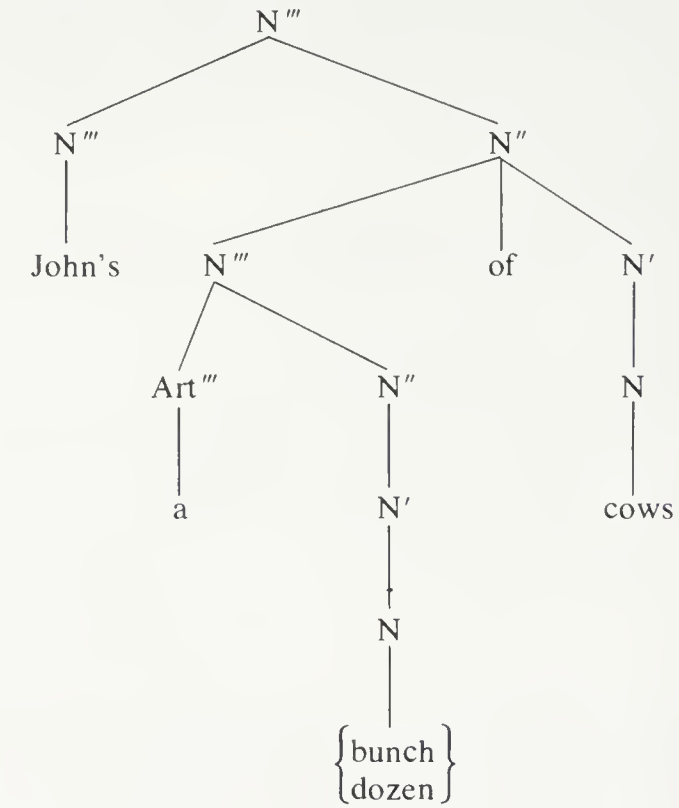
[15] There is also the intermediate reduction *a coupla trees*, occurring with other group nouns as well, e.g. *a buncha trees*. This reduction depends on *of* being a daughter of N″. Notice that in the following example, the reduction is impossible:

(i) We bought a couple (of) pictures of Fred, and a $\left\{ \begin{matrix} \text{couple of} \\ ?*\text{coupla} \end{matrix} \right\}$ Bill.

The reason is that the final NP in this S is the result of N′-Gapping and hence has the structure $[_{N''}[_{N''}$a couple $[_{N'}$PRO of Bill]]]. The presence of PRO inhibits the reduction of *of*.

section 5.5.1. To complete our account of the local rules applying to numerals, we observe that Pseudopartitive *a*-Deletion, developed in section 5.4.2, applies also to seminumerals, since the derivations of *John's bunch of cows* and *John's dozen cows* are precisely parallel in this respect:

(5.73)



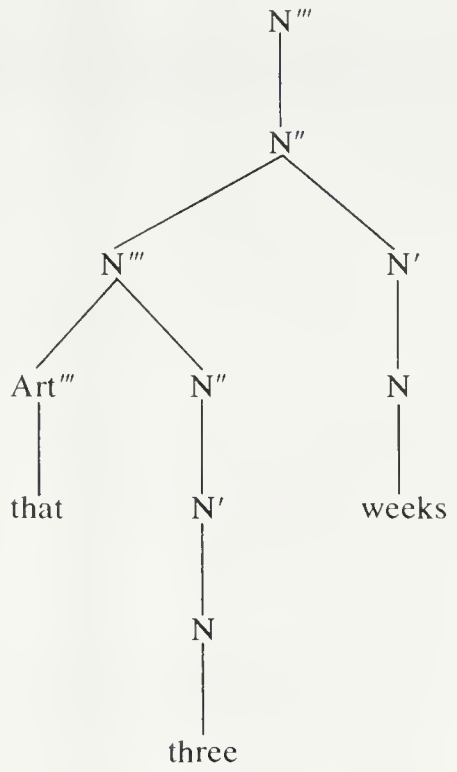
(*a*–Deletion) ⟹ John's { bunch of / dozen of } cows

(*of*–Deletion) ⟹ John's { bunch of / dozen } cows

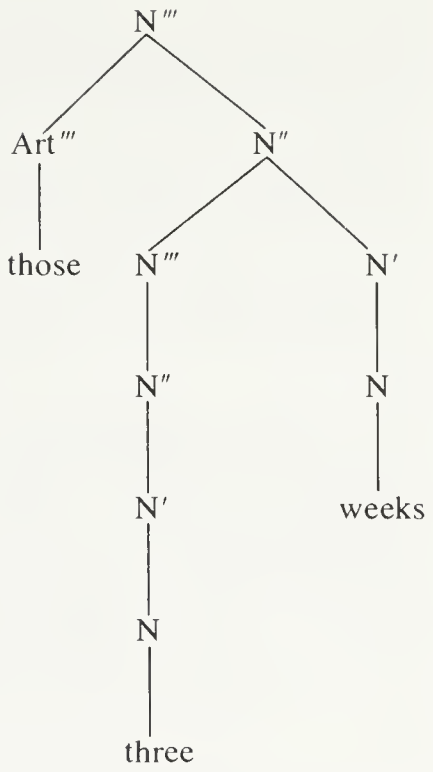
However, there is a difference between group nouns and numerals. Section 5.4.2 showed that when a group noun in a pseudopartitive is preceded by an article, the article must be in the specifier of the group noun, not in the matrix NP. For example, *that group of men* has singular *that* in the specifier of *group*; but **those group of men*, with plural *those* in the specifier of *men*, is ungrammatical. Numerals apparently lack this constraint, since we have both *that three weeks*, with *that* in the specifier of *three*, and *those three weeks*, with *those* in the specifier of *weeks*.[16]

[16] As further justification for these structures, notice the distinction between *every three weeks* and *all three weeks*. *Every* requires a singular head, and thus cannot be in the specifier of *weeks*. Moreover, the meaning of *every three weeks* is not 'every one of three weeks' but something like 'every group of three weeks': *every* quantifies instances of *three*. Thus the evidence points to *every* being in the specifier of *three*, like *that* in (5.74a). By contrast, *all* requires a plural head, and in *all three weeks* it quantifies instances of *weeks*. This means that it must be in the specifier of *weeks*, and hence that *all three weeks* has structure (5.74b).

(5.74) a.



b.



Now since seminumerals normally appear at the surface with their own determiners, we would expect, parallel to (5.74b), **those a dozen weeks*. Instead, we find *those dozen weeks*. Apparently Pseudopartitive *a*-Deletion must be extended so as to delete *a* in this configuration as well.

In the revision of Pseudopartitive *a*-Deletion, the feature system collapses N‴ and Art‴ in the first term of the rule:

(5.75) *Pseudopartitive* a-*Deletion*, revised

$$\left[{}_{N'''}\left[\begin{matrix} +\text{Subj} \\ -\text{Obj} \\ +\text{Det} \end{matrix} \right]''' - a - X \right] \Rightarrow 1 - \phi - 3$$

OBLIGATORY

Thus the underlying structure of *those dozen weeks* is *those a dozen of weeks*, which undergoes Pseudopartitive *a*-Deletion and Numeral *of*-Deletion to reach the surface form.

This completes our discussion of numerals. We have shown that, despite their similarities to quantifiers, their specifier structure proves them to be nouns very much like group nouns. Most differences between them and group nouns, and between cardinals and seminumerals, are accounted for by simple local transformations.

5.6. Summary

Here is a list of NP specifiers classified in terms of the syntactic and semantic properties discussed in this chapter.

Articles

Semantically demonstratives: *a, the, this, that, these,* $\begin{bmatrix}\textit{those}\\+\text{Part}\end{bmatrix}$, $\begin{bmatrix}\textit{which}\\+\text{Part}\end{bmatrix}$, *what, we,* you_{plur}

Semantically quantifiers (all [+Partitive]): *any, all, no, some each, every, either, neither, both*

Quantifiers (all [+Partitive]): *many, much, several, few* (negative), *little* (negative), *one* (nonpluralizable)

Nouns

Group nouns: *group, gallon, bunch, number, lot, score, hundreds, thousands, millions, couple* (some dialects), *dozen* (some dialects)

Seminumerals ([+Num, −Card]): *hundred, thousand, million, few* (nonnegative), *little* (nonnegative), *couple* (some dialects), *dozen* (most dialects)

Cardinals ([+Num, +Card]): *two, three,* etc.

Pronouns: *we, you, one* (pluralizable), PRO

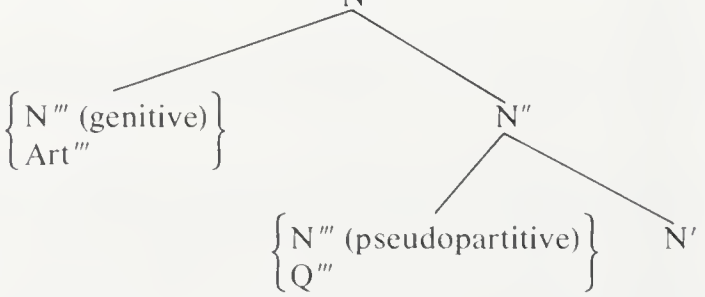
The following two phrase structure rules are responsible for the NP specifier system:

(5.76) $\quad N''' \rightarrow (\begin{bmatrix}+\text{Subj}\\-\text{Obj}\\+\text{Det}\end{bmatrix}''') - N''$

(5.77) $\quad N'' \rightarrow (\begin{bmatrix}+\text{Subj}\\-\text{Obj}\\-\text{Det}\end{bmatrix}''') - (A''')^* - N' - \ldots$

In (5.76), the feature matrix can be expanded as N‴, which becomes a genitive, or as Art‴, which is a demonstrative or N‴ quantifier. In (5.77), the feature matrix can be expanded as N‴, which is a group noun or numeral in a pseudopartitive construction, or as Q‴, an N″ quantifier.

(5.78)

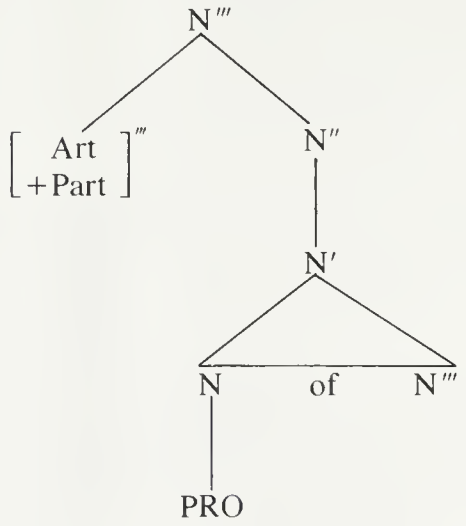


Certain configurations generated by applying these phrase structure rules freely are filtered out by constraints we have mentioned. The combination of an N‴ quantifier and an N″ quantifier is ruled out by the Specifier Constraint. The combination of an Art‴ in N‴ and a group noun in N″ is ruled out by a constraint justified in section 5.4.2;

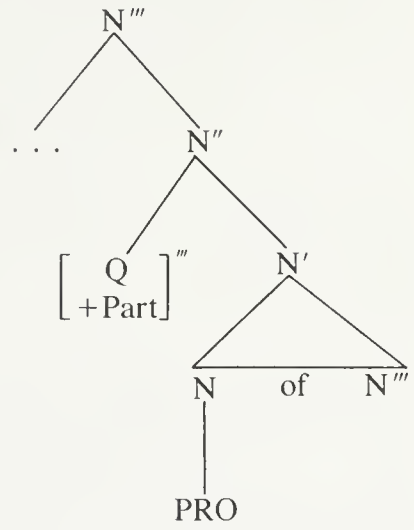
but an Art‴ in N‴ followed by a numeral in N″ is not ruled out, as we showed in section 5.5.3. Further constraints are necessary and have been left for future research.

Partitive constructions take one of these three forms:

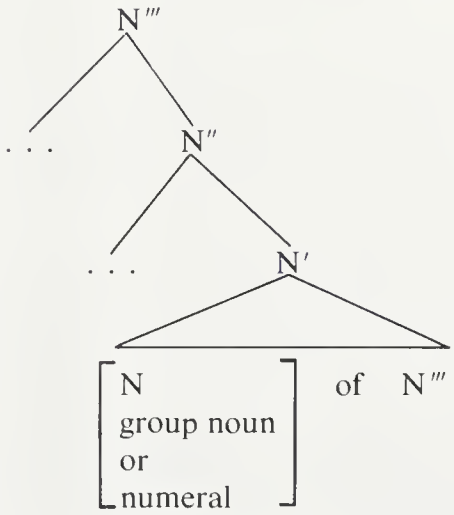
(5.79) a.



b.



c.



The Partitive Constraint restricts the possible specifiers of the complement N‴ to those which contain a genitive or demonstrative. In (5.79a,b), PRO is interpreted by the Partitive Projection Rule, which gives PRO the reading UNIT or AMOUNT.

We have developed four local transformations. The first, Substantivization, changes the form of various articles and possessive pronouns when they are directly followed by PRO. For example, *no* becomes *none*, *every* becomes *every one*, and *my* becomes *mine*. This rule applies regardless of the interpretation of PRO; thus the changes can be observed not only in partitive constructions such as (5.79a), but also in instances of N′-Anaphora, N′-Gapping, and postposed genitives.

The second local transformation, Pseudopartitive *a*-Deletion, deals with situations in which a group noun or seminumeral occupies an N″ specifier, and the N‴ specifier is

nonempty. This rule deletes the indefinite article of the group noun or seminumeral, giving derivations such as these:

(5.80) John's [a rather large number] of objections
⇒ John's rather large number of objections
John's [a dozen] trees ⇒ John's dozen trees
$\left\{\begin{matrix}\text{those}\\\text{all}\end{matrix}\right\}$ [a hundred] days ⇒ $\left\{\begin{matrix}\text{those}\\\text{all}\end{matrix}\right\}$ hundred days

The other two local transformations deal with the differences between group nouns, seminumerals, and cardinals. We assume all three of these have underlying articles, and that all are followed by *of* when in the N″ specifier (pseudopartitive position). Numeral *of*-Deletion eliminates the *of* after all numerals, giving contrasts like *a pair of men* (group noun) vs. *a few men* (seminumeral). Cardinal *a*-Deletion deletes *a* immediately before cardinals, giving contrasts like *a couple weeks* (seminumeral) vs. *two weeks* (cardinal); the rule does not apply if there is an intervening adjective, so that the article appears both with seminumerals (*a beautiful few weeks*) and cardinals (*an ugly two weeks*).

The relevance of all this to the general theory of phrase structure presented here is this: first, the NP specifier system has long been an area subject to wild unsystematic proposals about phrase structure. We have shown that it can be analyzed quite consistently within the highly restrictive theory of phrase structure presented in chapter 3, and that its irregularities are accounted for either by local transformations or by nonsyntactic means.

Second, we have seen that the specifier system is much less consistent in its semantics than the complement system. Furthermore, a reader of this chapter who is concerned with semantics will have seen how little is really known about the semantics of specifiers, despite nearly a century of study of the formal logic of quantifiers and demonstratives. It is hoped that our systematic survey of the syntactic possibilities of specifiers will serve as a source of irritation to those in linguistic semantics who think they know something about quantification.

6: Specifiers of X‴

6.1. Introduction

This chapter presents an analysis of the specifiers of degree, developing two major cross-categorial phrase structure rules for specifier systems which introduce measure phrases and quantifier phrases in X″ and degree phrases in X‴. We will show that the minor lexical categories Q and Deg enter into cross-category generalizations with major categories, justifying the use of Q‴ and Deg‴ predicted by the Uniform Three-Level Hypothesis of phrase structure.

We draw heavily on Bresnan's (1973) lengthy and insightful discussion of specifier systems, which makes copious use of cross-category generalization and the $\bar{X}$ Convention. We will propose a number of improvements on Bresnan's analysis which in turn confirm the general theory of phrase structure adopted here.

The general procedure behind the analysis is the same one we followed in chapter 4. A particular construction found in one category is checked for its occurrence in other categories. An analysis is then developed consistent with the facts in all categories and with the phrase structure schema. In particular, several crucial hypotheses are verified here by reference to PP and its specifiers, to which Bresnan does not appeal, and which turn out to be a surprisingly rich source of evidence.

We will build up the specifier system piecemeal, starting with measure phrases, then going on to quantifier phrases. There follows a detailed discussion of the occurrence of degree phrases and how recursion takes place in their specifiers.

6.2. Measure Phrases

Both adjective phrases and prepositional phrases can be modified by a prehead noun phrase often called a *measure phrase*.

(6.1) a. two feet long
seven miles wide
b. five miles down the road
three inches along the seam

By means of conjunction and various preposing transformations, it can be shown that the phrases in (6.1) are constituents.

(6.2) a. The pool is sixty feet long and thirty feet wide.
b. Six feet tall you'll never be!
c. His house is to the left and six miles down the road.
d. Four hundred yards up the street they encountered a panther.

Apparent counterparts of these construction in NPs are such phrases as *an inch of rope* and *five miles of road*, especially since the parallel between, for example, *two feet long* and *two feet of length* is semantically appealing. As mentioned in section 5.4, Selkirk (1977) delineates the cases in which these measure phrases are in fact specifiers, followed by *of*: those in which no article follows the *of*. Thus *two feet of rope* has the desired structure, with *two feet* as a measure phrase and *rope* as the head; but *two feet of the rope* has the partitive structure in which *of the rope* is in the N′ complement.

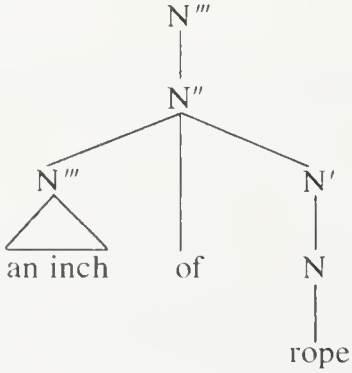
The *of* in the cases we are interested in here is therefore not a preposition but a specified grammatical formative like *Poss*, introduced either by the base or a local transformation. To decide which source is correct, observe that there is another NP construction, not often discussed, which fills the syntactic paradigm (6.1) even better than *an inch of rope* does.

(6.3) a. This mixture is two parts alcohol and three parts water.
b. The book was (one) half garbage.

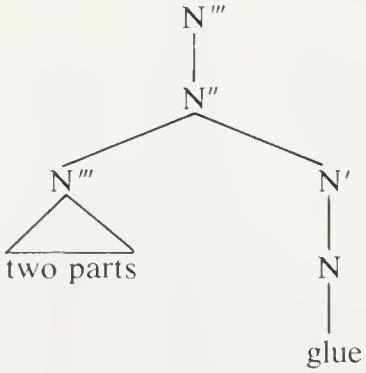
Expressions like *two parts alcohol* could be little but NPs, since they are made up of a numeral and two nouns; but they may be used only in predicate positions, e.g. after *be*, *seem*, *stay*, and so forth. Thus there are two uses of measure phrases in the specifier of NP, with different interpretations. If the *of* in the use first mentioned is base-generated, we can provide a deep structure cue for the difference in interpretation.

Where should the measure phrase be attached within NP? It precedes adjectives, which are in N″: *two gallons of ordinary water, three parts clear alcohol*, not **ordinary two gallons of water*, **clear three parts alcohol*. Thus it is either in N″ or N‴. If we attach it to N″, we automatically account for the fact that measure phrases do not receive the *Poss* invariably attached to N‴ in the N‴ specifier. Thus we arrive at the structures (6.4a,b).

(6.4) a.

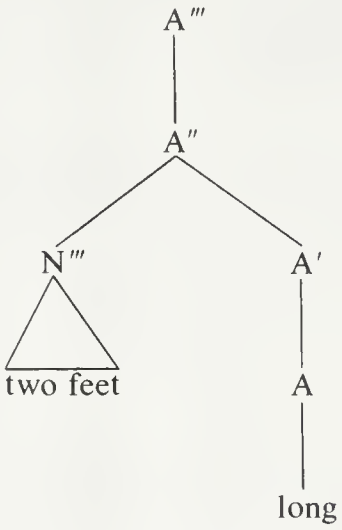


b.

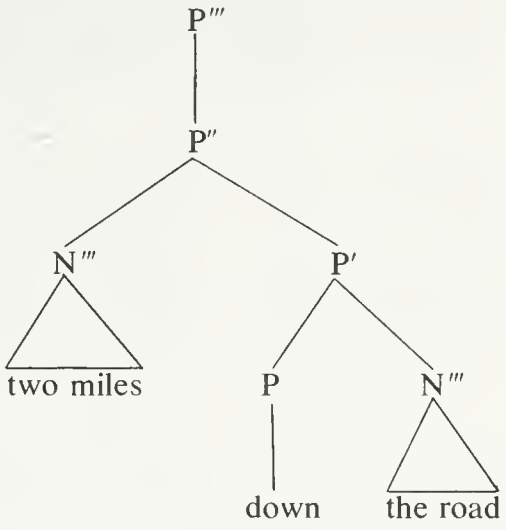


Given (6.4) as the structure for measure phrases in NP, we are constrained by the $\bar{X}$ Convention to propose structures (6.5a,b) for (6.1a,b). We then have the partial phrase structure rules (6.6) and a grammatical relation (6.7).

(6.5) a.



b.



(6.6) a. $N'' \rightarrow (N''' - (of)) - N'$
b. $A'' \rightarrow (N''') - A'$
c. $P'' \rightarrow (N''') - P'$

(6.7) A *measure phrase* is the italicized constituent in the configuration $[_{X''} N''' \ldots X' \ldots]$.

The projection rules for measure phrase will require that it consist of a quantified count noun, with further selectional restrictions on the noun depending on the nature of the X′ to which it is a sister. The difference between the two interpretations of measure phrase in NP will depend on the presence or absence of the specified grammatical formative *of*.

The major category which is still missing in the paradigm above is V. There is nothing preceding the verb which looks like the constructions in (6.1)–(6.4), but measure phrases do appear postverbally.

(6.8) a. Jill ran around the track three times.
b. Fran stayed in Africa three years.[1]

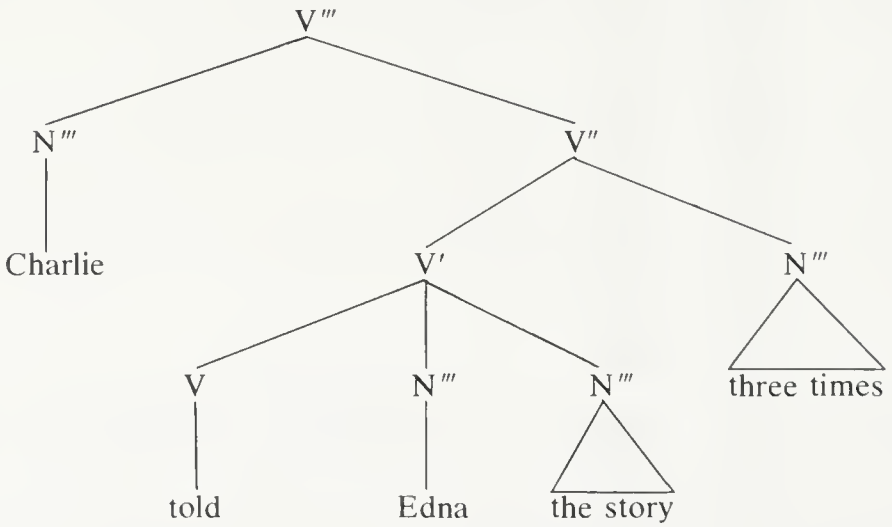
Note that the measure phrase appears after a PP in both of these examples, a new position for NP. In fact the measure phrase must appear after all strictly subcategorized phrases (barring Heavy NP Shift), as shown by (6.9).

[1] The measure phrase in (6.8b) is intuitively quite close to the PP *for three years*, and it is quite natural to assume that it is the result of a *For*-Deletion. On the other hand, since *three times* does not appear to be the result of a reduction, there is a base position available for *three years*, and we might as well make use of it, sparing ourselves a rather idiosyncratic transformation.

(6.9) a. Charlie put the book on the table three times.
b. ?Charlie put the book three times on the table.
c. *Charlie put three times the book on the table.
d. Charlie told Edna the story three times.
e. *Charlie told Edna three times the story.
f. *Charlie told three times Edna the story.

The position of the measure phrase is thus the same as that of postverbal manner adverb, so it must be a daughter of V″, generated by the partial phrase structure rule (6.11).

(6.10)



(6.11) $V'' \rightarrow V' - (N''')$

The posthead position of the measure phrase appears also as a second option in PP, as in *down the road three miles, up the river a way*. By the $\bar{X}$ Convention, this must be attached to P″, parallel to V″, and there must be a partial phrase structure rule (6.12).

(6.12) $P'' \rightarrow P' - (N''')$

There is a curious partial generalization here. All major categories permit a measure phrase as a daughter of X″, but it is a left sister of N′ and A′, a right sister of V′, and either a left or right sister of P′. Section 4.5 discussed the problems involved in expressing this partial generalization in terms of conventionally stated phrase structure rules, and I have nothing to add here.

In order to generalize the grammatical relation *measure phrase* to all categories we must apparently resort to a definition of the *Aspects* sort, which involves not order but only domination.

(6.13) A *measure phrase* is an N‴ immediately dominated by X″. (In the *Aspects* notation, [N‴, X″].)

This section has shown, then, that NPs with the function of measure phrases occur in the X″ of all major lexical categories, providing another significant cross-category generalization. We next show that these NPs alternate with QPs in the same position.

6.3. Quantifier Phrases in X″

Section 5.2 claimed that the quantifiers *many, few, much, little,* and *several* are daughters of N″. They differ from the articles *all, each, every, some,* etc., which are daughters of N‴ and cannot be preceded by genitives or other articles. The operative phrase structure rule was (6.14).

(6.14) $N'' \rightarrow (Q''') - (A''')^* - N' - \ldots$

Since the N″ quantifiers do not cooccur with measure phrases, we can immediately generalize (6.14) with (6.6a) to form (6.15).

(6.15) $N'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \\ \langle +\text{Comp} \rangle \end{bmatrix}''' - \langle (\text{of}) \rangle) - (A''')^* - N' - \ldots$

In (6.15), the realization of the constituent broken down into features has two possibilities. If it is [+Comp], the features spell out N‴ and the *of* is possible. If it is [−Comp], the features spell out Q‴ and the *of* cannot occur.

This generalization of measure phrase and Q‴ carries over to the other major categories. In sentences, we find QPs postverbally, as in *You talk too much.* Since these QPs are not strictly subcategorized, they belong in V″, so we can generalize (6.11) to (6.16).

(6.16) $V'' \rightarrow V' - (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''')$

In PPs there are quantifiers *far* and *long* (as in *far down the road* and *long after the accident*) which alternate with measure phrases, and which, as we will see, take the same modifiers as *many* and *much*. This justifies generalizing (6.6c) to (6.17).

(6.17) $P'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - P'$

The QP, like the measure phrase, may follow the P′ as well, as in *down the road too far.* This use is subject to the restriction that the QP must contain a specifier of its own,

since **down the road far* is out. I have no explanation for this restriction; for present purposes I will conveniently ignore it, generalizing (6.12) to (6.18).

$$(6.18)\quad P'' \rightarrow P' - (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}^{m})$$

Having dealt with the generalization of measure phrase and QP in N″, V″, and P″, we turn to A″. The situation in A″ is more complex than the other three cases, and presents a difficult choice. Bresnan (1973) chooses one possibility, which, as we will see, has repercussions for her entire analysis of the degree system. We will show that a different choice provides a more satisfactory analysis throughout. The rest of this section will outline the problem, and its thread will be repeatedly picked up in the rest of the chapter.

The problem is this: on the whole adjectives do not take simple quantifiers; for example, **much long* and **little long* are ungrammatical. However, in the comparatives *more beautiful* and *less beautiful, more* and *less* appear to be the comparative forms of the quantifiers *much* and *little*. The question is how to account for this curious alternation.

Bresnan describes it by claiming that *much* deletes obligatorily before adjectives, leaving only its modifiers. For example, underlying *as much beautiful* becomes *as beautiful*; underlying *much beautiful* becomes *beautiful*. In *more beautiful, much* becomes *more* by affixation of the comparative *-er*; since it is no longer the phonological *much*, it does not delete.

One immediate difficulty with Bresnan's account is the nonexistence of **little beautiful*, since a deletion of the quantifier *little* here could hardly result in anything with the expected meaning. Yet since *less beautiful* is to be derived from underlying *-er little beautiful*, some way must be found to rule out the noncomparative form.

There seem to be two ways out of this difficulty. First, Bresnan could be right, and all adjectives could take quantifiers, but various deletions and semantic constraints could prevent all forms but the proper comparative forms. Under this theory, (6.6b) would immediately generalize to (6.19), parallel to (6.15) and (6.17).

$$(6.19)\quad A'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}^{m}) - A'$$

As an alternative to Bresnan's theory, I propose that usually adjectives do not take quantifiers, and that the exceptions *more* and *less* are not quantifiers, but arise from a different source, to be discussed in section 6.4.3. Under this theory, (6.6b) does not at first glance appear to generalize to (6.19).

However, Bresnan points out some evidence that there is a generalization anyway (note 4, p. 278): there are at least two adjectives which appear with surface quantifiers.

(6.20) a. You and I are {much / little / far} different from Bill.

b. You and I are {much / little} alike.

In Bresnan's theory, these adjectives must be lexically marked to take optional rather than obligatory *Much*-Deletion. However, this fails to explain why *little* and *far* can occur with these adjectives and not with others.

Under the alternative view I have proposed, there is another account: these adjectives are exceptions to the generalization that adjectives do not take quantifiers. Such exceptionality is not too surprising, in view of the fact that adjectives that permit measure phrases must also be lexically marked and are almost equally rare (*high, wide, long, tall, old, thick, broad,* and perhaps a few others). In order for these exceptional adjectives with quantifiers to exist, the phrase structure rules must allow a quantifier, and thus the generalization of measure phrase and QP in (6.19) is motivated.

This view of QP in AP eliminates all need for Bresnan's *Much*-Deletion transformation, along with the difficulties it poses for the occurrence of *little* and *far*. Rather, the few cases where there is surface evidence for a QP are taken as exceptions in underlying structure, provided for by general phrase structure rules. One would hope that the exceptionality ultimately proves to be semantically motivated, leaving the syntax and lexicon maximally simple.

Whether we take Bresnan's view or the present one on the presence of quantifiers in A'', this section has shown the need for phrase structure rules (6.15), (6.16), (6.17), (6.18), and (6.19). We thus see that the base rules for measure phrases and quantifiers are fully general among the major lexical categories, aside from the problem of order in V'' and P''. The parallelism of the X'' specifiers justifies us in considering all these rules to be special cases of a single phrase structure rule. Furthermore, the recurring parallelism of NP and QP in the X'' specifier is evidence for the feature analysis that relates the lexical categories N and Q.

6.4. Degree Phrases

There are a number of words we will classify as Deg which, among other things, precede the quantifiers of the preceding section. (6.21a) illustrates them in QP within N''', (6.21b) in QP within P''', and (6.21c) in QP within V'''.

(6.21) a. {how / this / that / so / too / as} {{many / few} people / {much / little} dust}

b. {how / this / that / so / too / as} {far along the road / long after the accident}

c. He talks {{this / that / so / too} {much / little} / as {much / little} as Bill}

How {much / little} does he talk?

This section will develop the phrase structure rule generating them, showing that it too is subject to an interesting cross-category generalization.

6.4.1. Attachment of Deg‴ to Q‴

There is little strong evidence for where the degree words in (6.21) are attached to Q‴. However, the $\bar{X}$ Convention suggests the hypothesis that they are Q‴ specifiers. This accounts for the parallelism in the use of *this* and *that*, which as specifiers in N‴ are of the category Art but with similar demonstrative meaning. Under this theory, the N‴ specifier will then contain a [+Subj, −Obj, −Comp, +Det]‴ and the Q‴ specifier will contain a [−Subj, −Obj, −Comp, +Det]‴. The lexical items *this* and *that* will either occur in both categories, differing by the feature *Subj*, or, preferably, they will simply be unmarked for this feature. This parallelism between Deg and Art is the reason for assigning Deg the feature analysis proposed in chapter 3. A further bit of evidence for the parallelism arises in section 6.4.4.

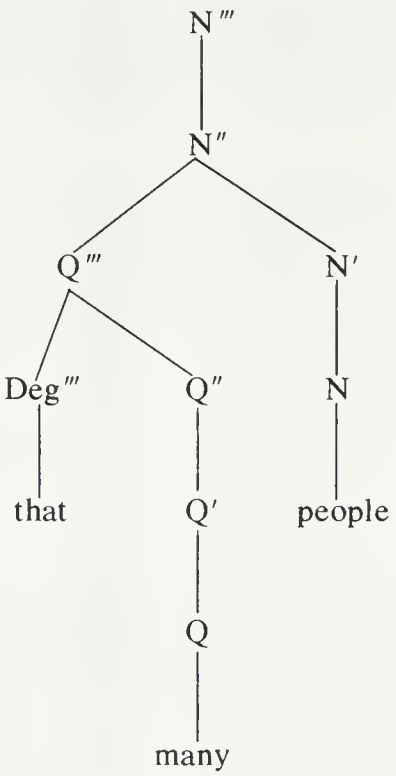
Bresnan suggests a different attachment of degree words, to what in our terminology would be Q″. Section 6.5 will call into question the validity of her argument for this structure. Anticipating the refutation of her argument, I will adopt the solution suggested by the $\bar{X}$ Convention and generate degree phrases in QP with the partial phrase structure rule (6.22). Its relationship to the partial phrase structure rule (6.23) for N‴ is obvious.

(6.22) $Q''' \rightarrow (Deg''') - Q''$

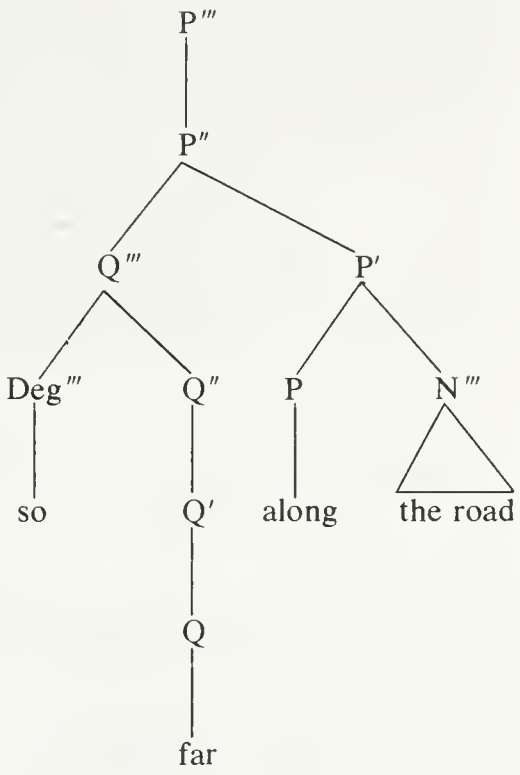
(6.23) $N''' \rightarrow (Art''') - N''$

Example structures are given in (6.24).

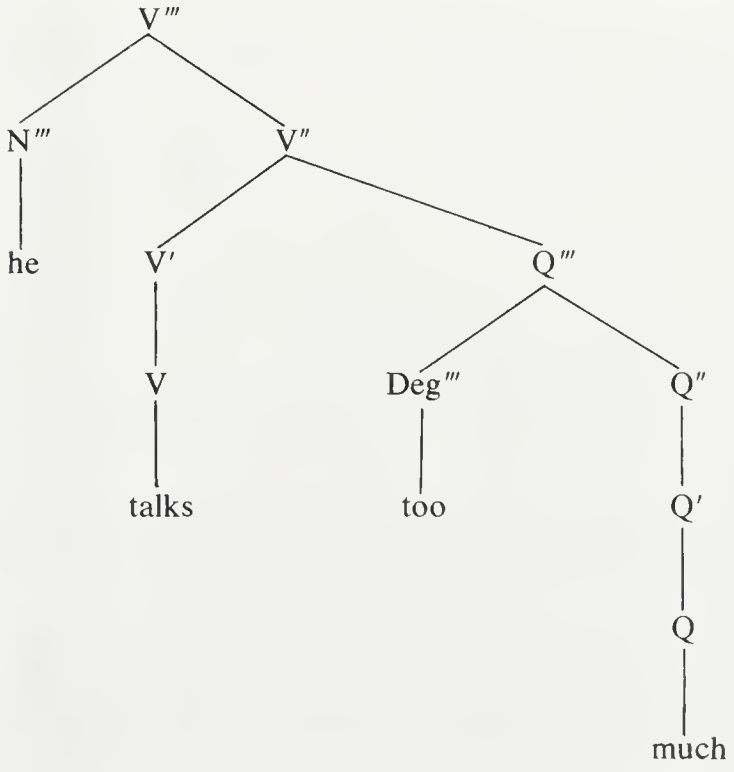
(6.24) a.



b.



c.



There are three questions that must be dealt with in implementing this structure. The most crucial is how the degree phrase is applied to A‴ and Adv‴. The next subsection deals with this. There are also consequences for the treatment of the words *more*, *less*, and *enough*, which we turn to in sections 6.4.3 and 6.4.4.

6.4.2. The Degree System of Adjectives and Adverbs

Section 6.3 claimed that adjectives (with a few rare exceptions) do not occur with quantifiers. Therefore the degree words *as*, *so*, *too*, *this*, *that*, and *how* in A‴ and Adv‴, as in (6.25), must be generated by a phrase structure rule that introduces degree phrases directly in the specifier of A‴ and Adv‴.

(6.25) { as / so / too / this / that / how } { tall / quickly }

The $\bar{X}$ Convention suggests that the degree phrase be generated in the A‴ and Adv‴ specifier, as in (6.26), so that the rules generalize with (6.22) and to a lesser extent with (6.23).

(6.26) a. A‴ → (Deg‴) – A″
b. Adv‴ → (Deg‴) – Adv″

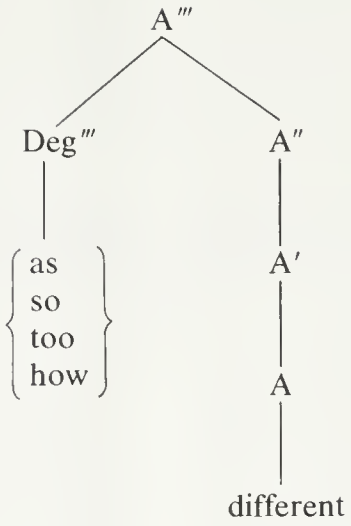
This analysis thus claims that adjectives appear with the same degree words as *much* because of the similarity of the phrase structure rules (6.22) and (6.26).

Bresnan expresses this generalization a different way: by quantifying adjectives with a *much* which is modified by a Deg‴. This *much* is deleted transformationally, leaving only the degree phrase at the surface. Section 6.3 pointed out the difficulty in this view: it requires all adjectives to allow a QP, thus leaving unexplained the distribution of the quantifier *little*. Thus Bresnan's means of expressing the generalization is questionable, whereas in the present analysis, such a distribution is an automatic consequence of the proposed base structure

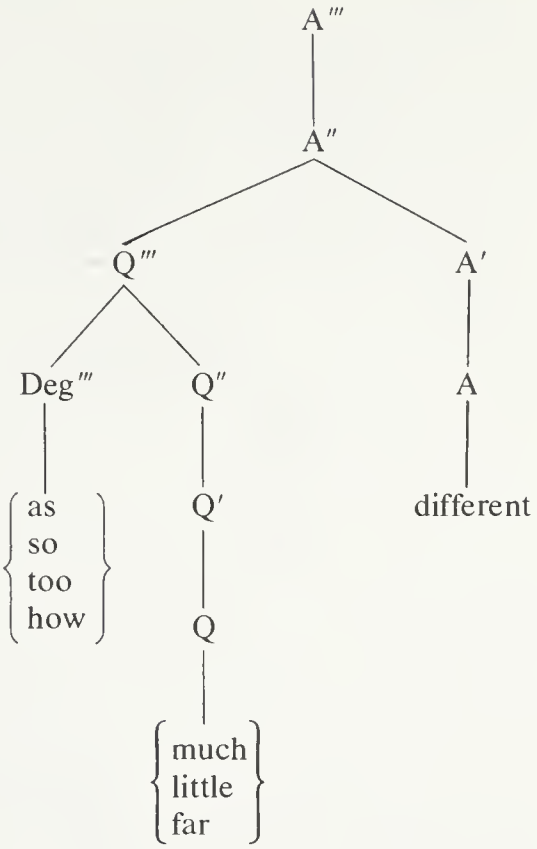
A complication Bresnan points out turns out to follow immediately from the present theory. She observes (note 4, p. 278) that those adjectives such as *alike* and *different* that occur at the surface with *much* have two possible forms with degree words, for instance *as different* and *as much different*. She uses this as evidence that *Much*-Deletion is optional with these adjectives. The present theory permits a more direct account, using phrase structure rules alone: the two base forms in (6.27).[2]

[2] A third possibility would attach Deg‴ to A‴ and Q‴ to A″. I have no arguments against such a structure, except for the possibility that combinations of quantifiers and degree words as independent specifiers may be excluded by some generalization of the Specifier Constraint of section 5.2. I know of no cases where both are needed; but I understand the semantics of degree phrases and quantifiers sufficiently poorly that I am loath to make a strong commitment as to how the unnecessary combinations are to be eliminated.

(6.27) a.



b.



Hence the alternation follows not from a transformational exception, but from the unusual subcategorization possibilities of these adjectives. This provides further evidence that Deg‴ is generated in A‴ and Adv‴ as well as in Q‴, yielding a further cross-category generalization in the specifier system.

6.4.3. Comparative Specifiers

Section 6.3 pointed out that the hypothesis that most adjectives do not take quantifiers has as a consequence that *more* in *more beautiful* cannot be the comparative of the quantifier *much*. Let us suppose instead that it is a Deg in the A‴ specifier, parallel to *as, too, so,* and *how* in (6.27a).

If *more* is a Deg, though, we expect it also to occur in the Deg‴ within a Q‴, giving the incorrect forms **more many*, **more much*, etc. We correct this problem by providing the idiosyncratic spelling rules (6.28) to yield the existing forms.[3]

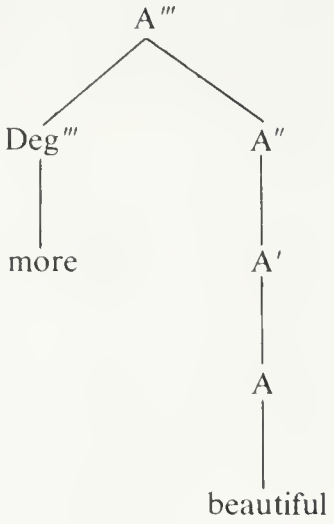
[3] Note that these rules are not a disadvantage with respect to Bresnan's theory, since she needs similar rules to attach her comparative morpheme *-er* to quantifiers:

-er – much ⇒ more
-er – many ⇒ more
etc.

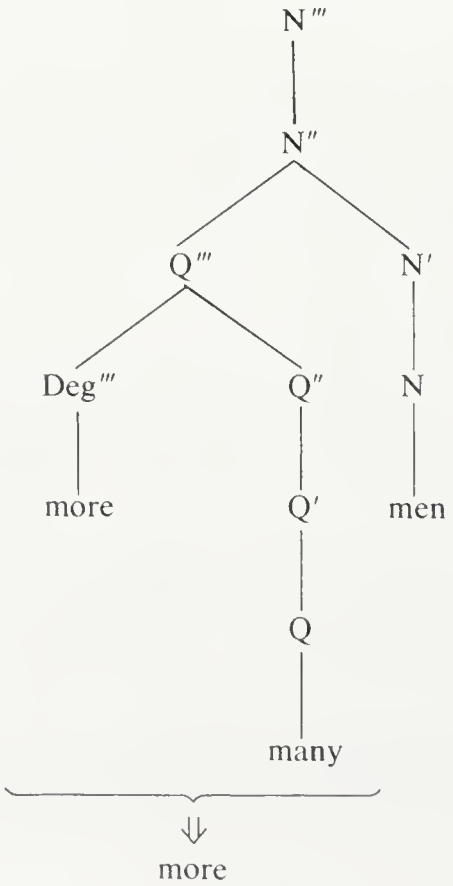
(6.28) more - much ⇒ more
more - many ⇒ more
more - little ⇒ less
more - few ⇒ fewer
more - far ⇒ farther

The theory that *more* is a Deg thus produces structures such as these:

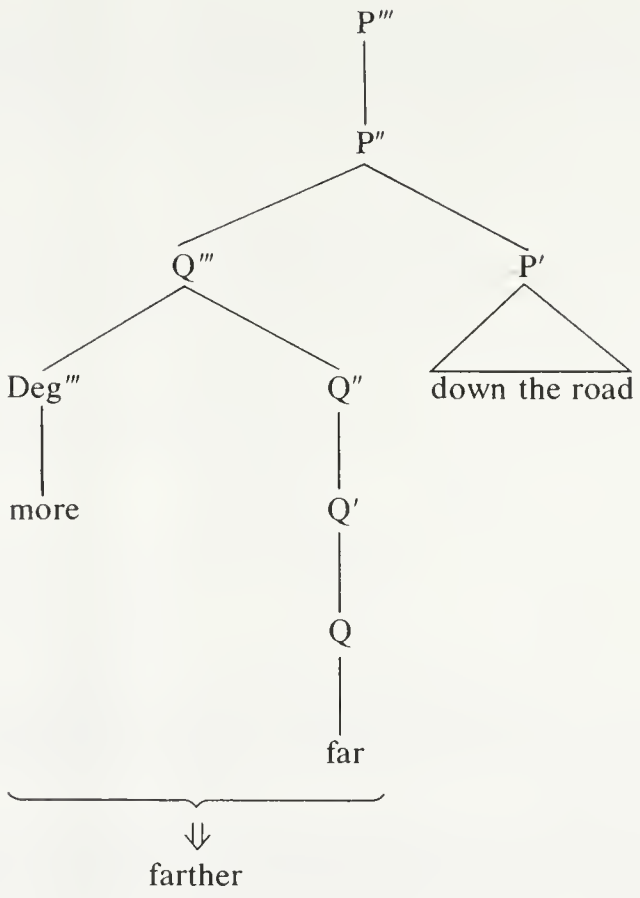
(6.29) a.



b



c.

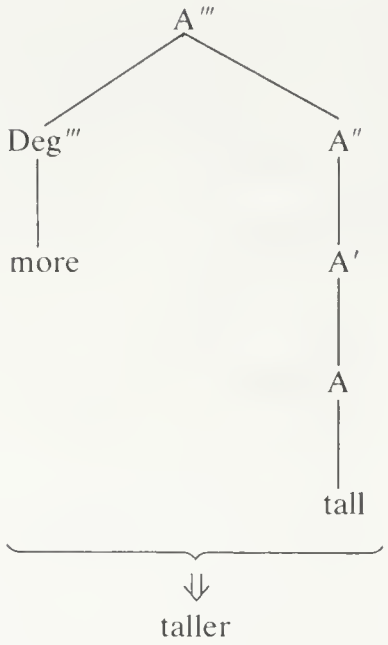


As a consequence of the theory that *more* in *more beautiful* is not the comparative of *much*, *less* in *less beautiful* cannot be the comparative of the quantifier *little* either. Rather, *less* must also be able to be a Deg, and *less beautiful* is then identical in structure to (6.29a).

We further predict that *less* should be applicable to quantifiers as well as to adjectives, substituting for *more* in constructions like (6.29b,c). One such possibility exists—*less far*, as in *less far down the road*; but there are no **less much*, **less many*, **less little*, or **less few*. Thus, while Bresnan's theory must find a way to generate *less far* and eliminate **little beautiful*, the present theory must find a way to rule out the other four *less*-Q combinations. The four bad forms, if they existed, would presumably be synonymous with the existing forms *more little* > *less*, *more few* > *fewer*, *more much* > *more*, and *more many* > *more*, respectively, which somehow seem "less negative". This fact may be useful someday in giving an interesting account for what must remain for now an unprincipled filter.

Now consider the rule that forms comparative adjectives such as *taller*, presumably from the source *more tall*. As is well known, this rule applies only to a particular class of lexically marked adjectives, mostly monosyllabic. In the present theory it applies to this structure:

(6.30)



Observe that this configuration is precisely parallel to that in the QPs of (6.29b,c). Thus the rule forming comparative adjectives generalizes via the $\bar{X}$ Convention with the rule forming comparative quantifiers.

In Bresnan's system, there is of course also a rule that turns *more* – *A* into *A-er*. However, in her theory the *more* – *A* rule attaches the quantifier *more* to the head and thus has little to do with the rule forming comparative quantifiers, which attaches the Deg *-er* to the head (see footnote 3, p. 277). Thus the present theory expresses a further generalization missed by Bresnan's analysis, justifying the assignment of *more* to the category Deg, and further justifying the absence of QP in most APs.

As a final note, observe that all the arguments given here for the comparative Deg *more* apply equally to the superlative Deg *most*, which also occurs in the specifiers of A‴, Adv‴, and Q‴. Like *more*, *most* undergoes spelling rules that attach it to quantifiers and lexically designated adjectives; the environment is exactly the same as with *more*, so the rules generalize. The word *least*, like *less*, occurs in the specifier of A‴ and Adv‴ and is therefore also a Deg. Since *least* does not appear comfortably with any quantifiers at all (even **least far* is unacceptable), we can assume either a generalization of the filter for *less* or some sort of semantic constraint. Thus the superlatives behave almost precisely like the comparatives.

6.4.4. Enough

We now turn to the most syntactically exceptional word in the specifier system, *enough*. It provides a great deal of evidence for the analysis of preceding sections.

As Bresnan points out, *enough* is peculiar in that it obligatorily follows adjectives, adverbs, and quantifiers that it modifies, and it optionally follows nouns. This behavior is unlike that of all other modifiers it alternates with, which invariably precede their heads.

(6.31) a. beautiful enough / $\left\{\begin{matrix}\text{so}\\\text{as}\\\text{too}\end{matrix}\right\}$ beautiful

b. quickly enough / $\left\{\begin{matrix}\text{so}\\\text{as}\\\text{too}\end{matrix}\right\}$ quickly

c. $\left\{\begin{matrix}\text{far}\\\text{few}\end{matrix}\right\}$ enough / $\left\{\begin{matrix}\text{so}\\\text{as}\\\text{too}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{far}\\\text{few}\end{matrix}\right\}$

d. $\left\{\begin{matrix}\text{enough pudding}\\\text{pudding enough}\end{matrix}\right\}$ / $\left\{\begin{matrix}\text{much}\\\text{little}\end{matrix}\right\}$ pudding

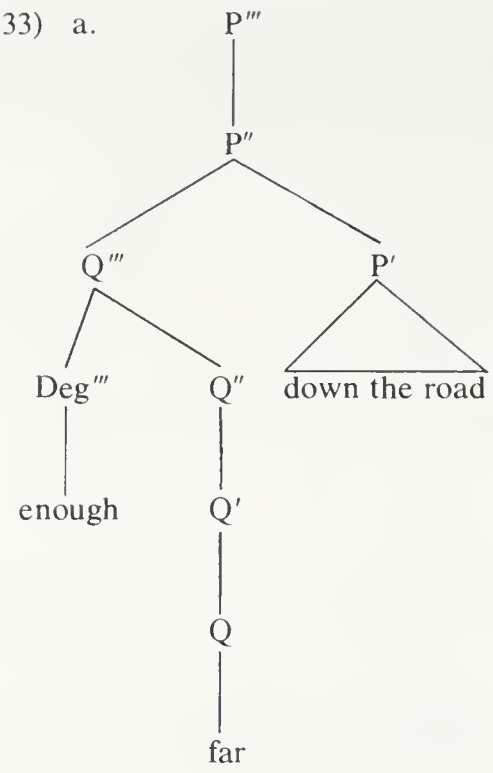
Bresnan proposes a local transformation to account for this, moving *enough* around its head. In present terms, this rule can be stated as (6.32), where X^0 denotes a lexical category.

(6.32) Enough-*Shift*
enough - $X^0 \Rightarrow \phi$ - 2 + 1
OBLIGATORY if X^0 = A, Adv, Q
OPTIONAL if X^0 = N

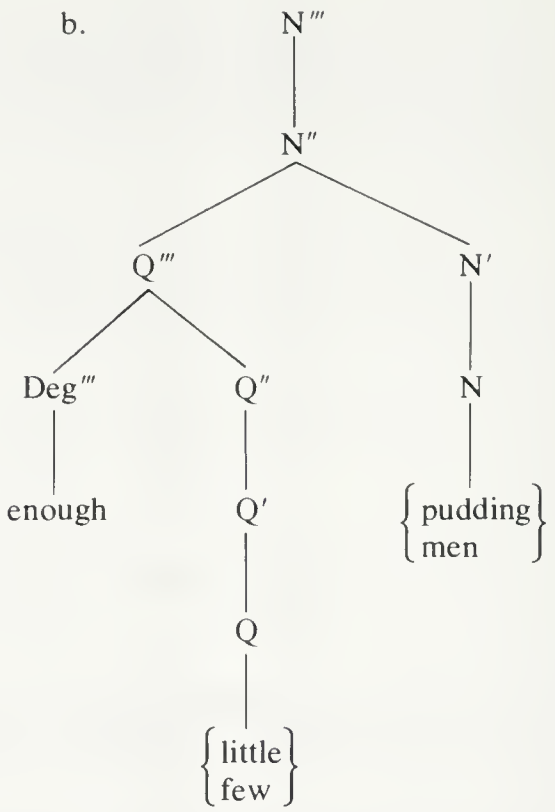
Bresnan treats *enough* as a Q, on the strength of similarities like $\left\{\begin{matrix}\textit{much}\\\textit{enough}\end{matrix}\right\}$ *pudding* and $\left\{\begin{matrix}\textit{many}\\\textit{enough}\end{matrix}\right\}$ *men*. *Small enough* is taken to have the underlying structure *enough small*, undergoing *Enough*-Shift; the parallel **much small* becomes *small* by obligatory *Much*-Deletion.

Consider, however, the alternative view that *enough* is a Deg. This would explain its cooccurrence with quantifiers in expressions like *far enough down the road, little enough pudding,* and *few enough men,* which would then have the following structures.

(6.33) a.



b.



Since Bresnan points out (note 3, p. 277) that her analysis is unable to generate these forms, we have already found a significant advantage to our analysis.

An apparent difficulty for our view that *enough* is a Deg is that it predicts the existence of the forms **many enough men* and **much enough pudding,* parallel to (6.33b). To correct this discrepancy, we introduce a rule to delete *many* and *much* after *enough*,[4] changing these incorrect forms to *enough men* and *enough pudding*. This rule is much more restricted than Bresnan's very general rule of *Much*-Deletion, since it localizes the irregularity in a particular combination of lexical items.

(6.34) Many–Much *Deletion*

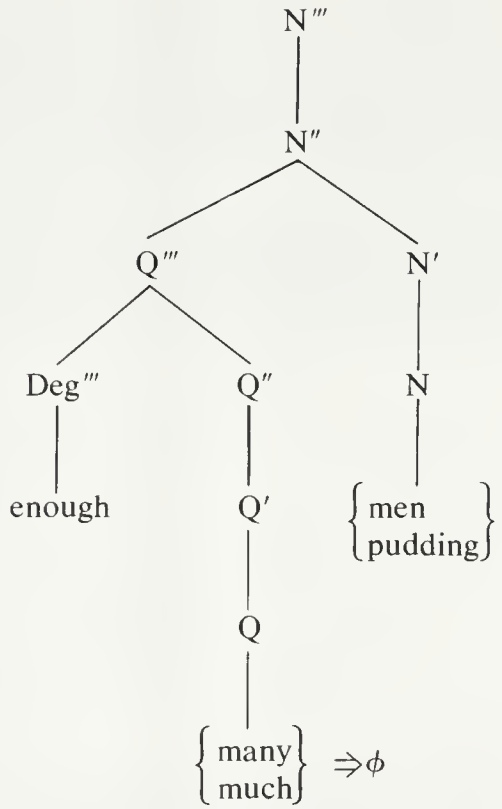
$$enough - \begin{Bmatrix} many \\ much \end{Bmatrix} \Rightarrow 1 - \phi$$

OBLIGATORY

Thus *enough men* and *enough pudding* have the structure (6.35).

[4] This assumes the rule precedes *Enough*-Shift. Alternatively, it could delete *many* and *much* before *enough*, if it were ordered after *Enough*-Shift. The choice does not seem crucial.

(6.35)



If the Q in (6.35) is changed to *few* or *little*, we get the forms *few enough men* and *little enough pudding* by *Enough*-Shift. If the Deg is changed to *as*, we get *as many men* and *as much pudding*, since neither *Many–Much* Deletion nor *Enough*-Shift applies.

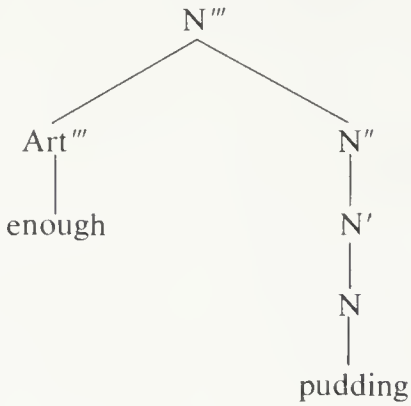
Bresnan's arguments that *enough* is a Q are based on a large number of parallels between it and *more*, which, as we mentioned, she takes also to be a Q, the comparative of *much* and *many*. However, our analysis also accounts for these parallels. There are two relevant cases. In the specifiers of adjectives and adverbs, *enough* and *more* are both Deg‴ attached to X‴, accounting for one set of parallelisms. In the specifiers of nouns, prepositions, and verbs, however, surface *enough* and *more* are dominated by Q‴ and are derived from underlying *enough* $\left\{\begin{matrix}much\\many\end{matrix}\right\}$ and *more* $\left\{\begin{matrix}much\\many\end{matrix}\right\}$, respectively, by the local transformations we have proposed. Hence the parallelism of *enough* and *more* is a consequence of our analysis as well as Bresnan's.

In addition to accounting for the combinations *little enough, few enough,* and *far enough*, the theory that *enough* is a Deg has other advantages. A minor dividend concerns the complement it binds (as in *enough pudding to sink a ship*). In note 11, p. 289, Bresnan claims that although *enough,* like the degree words *so, as,* and *too,* governs a complement clause, this is not an argument that it is a Deg. As justification, she points out that *sufficiently,* an adverb, also governs a complement (e.g. *It's*

sufficiently long for us to use it); so governing a degree clause is not a sufficient condition for something to be a Deg. But in fact, there are no other *quantifiers* that govern complements, so *enough* would be in any event an unusual quantifier, though not an atypical Deg. In our analysis, then, the fact that *enough* governs a complement is not unexpected.

The present analysis also has a much more striking advantage. If *enough* is a Deg, it can be made to function also as an Art by eliminating the feature [−Subj] from its lexical representation. Under this assumption, it would resemble the bicategorial *this* and *that*. It could then enter into two possible structures within NP, (6.35) and (6.36).

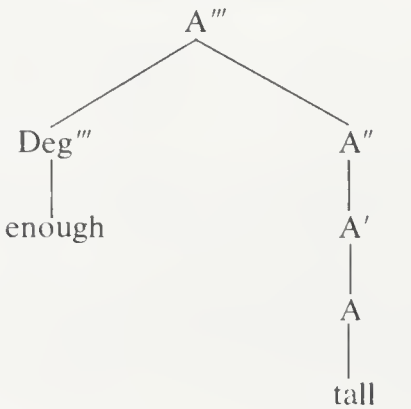
(6.36)



Now suppose that *Enough*-Shift, like other transformations, is subject to the condition of the strict cycle, i.e. it can apply only on the lowest possible X‴ domain in which *enough* appears. Then it will be unable to move *enough* around *pudding* in (6.35), because of the Q‴ dominating *enough*. On the other hand, in (6.36) the rule can apply to yield *pudding enough*. Hence the two possible NP structures, stemming from the categorial ambiguity of *enough*, interact with a well-motivated constraint on transformations to produce the variation between *enough pudding* and *pudding enough*. Thus *Enough*-Shift can be made obligatory in all contexts, an improvement on Bresnan's rule (6.32), which singles out NP for special treatment.

Now consider what happens in A‴. Under the present analysis of A‴, *tall enough* has the underlying structure (6.37).

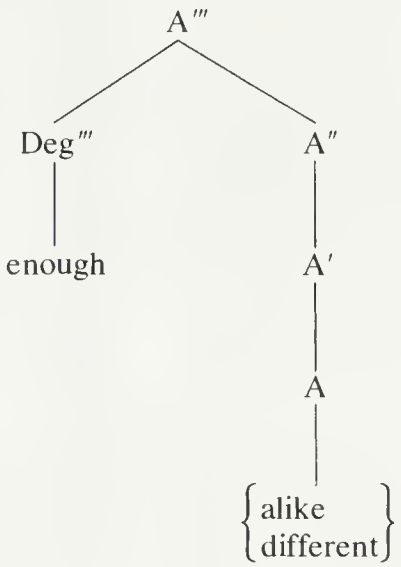
(6.37)



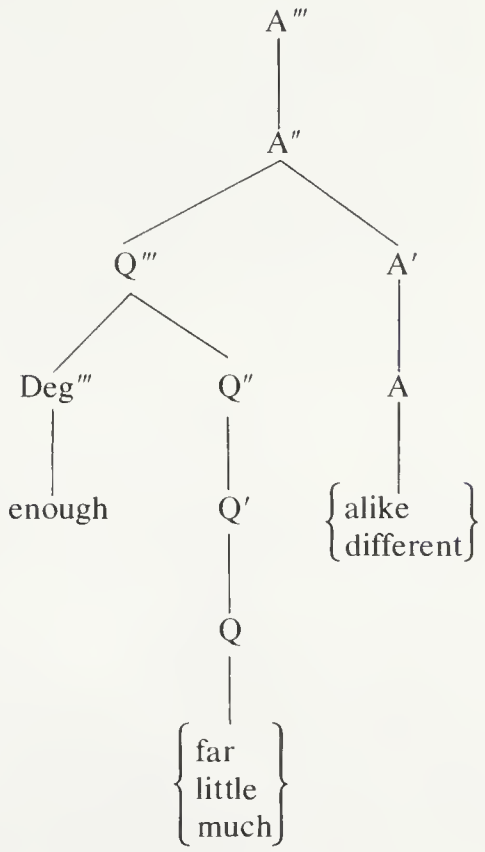
In (6.37), *enough* obligatorily permutes around *tall*, just as in *pudding enough* and *far enough*.

But observe how *enough* behaves with the exceptional adjectives *alike* and *different*. Superficially it looks as if *enough* permutes around them optionally rather than obligatorily, since we have both *alike enough* and *enough alike*. But there is a more interesting account: since *alike* and *different*, unlike most adjectives, allow a Q‴, they can occur in two distinct underlying forms with *enough*:

(6.38) a.



b.



In (6.38a), *enough* permutes around the head to give *alike enough* and *different enough*, parallel to (6.37). In (6.38b), *enough* permutes around the Qs *far* and *little* to give *little enough* {*alike* / *different*}, *far enough* {*alike* / *different*}. But with *much* in (6.38b), *Many–Much* Deletion takes place, and because of the strict cycle, *enough* does not permute around the adjective; the forms derived are thus *enough alike* and *enough different*. This is just like what happens in the nominal paradigm (6.35)–(6.36). What at the surface looks like an exceptional optionality in *Enough*-Shift with these adjectives turns out to be a consequence of the fact that they allow a QP in their specifier.

Thus the analysis of *enough* as a Deg confirms many details of the specifier system argued for in previous sections. It makes possible the use of *enough* with the quantifiers *little*, *few*, and *far*; it makes possible the generalization that degree words,

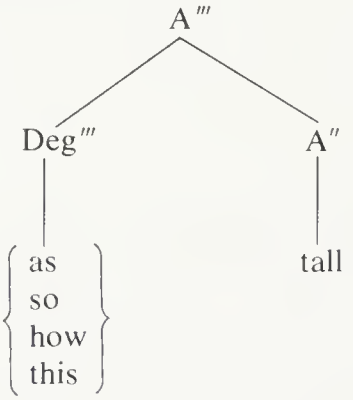
but not quantifiers, govern complements; and, most striking, it makes the apparent optionality of *Enough*-Shift in certain cases emerge as an automatic result of the base rules and the principle of the strict cycle. Furthermore, the exceptionality of *alike* and *different*, rather than being spread out over two transformations, *Much*-Deletion and *Enough*-Shift, is localized in a single phrase structure difference: unlike other adjectives, they allow a QP.

6.4.5. *Summary of Degree Words*

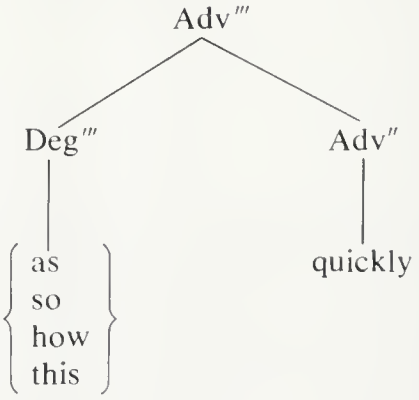
To recapitulate the argument of this section, we have claimed that the category Deg includes the words *as, so, too, how, more, less, most,* and *least*, plus the bicategorials *this, that,* and *enough,* which function also as articles.

The constituent Deg‴ is generated as an X‴ specifier in A‴, Adv‴, and Q‴, partially generalizing with the Art‴ position in the N‴ specifier.

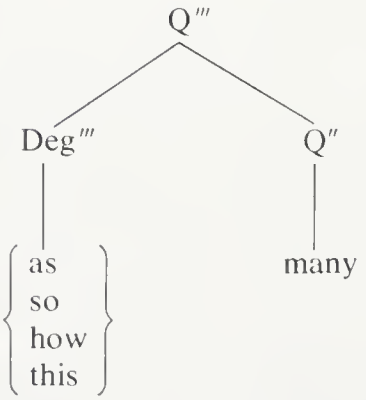
(6.39) a.



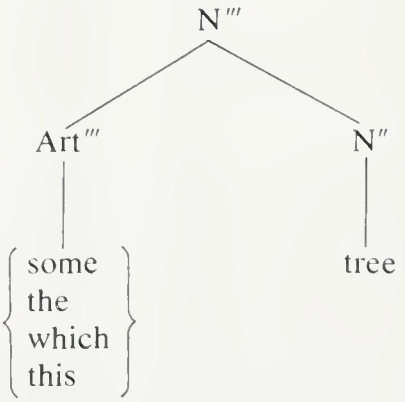
b.



c.



d.



The relevant phrase structure rules are these:

(6.26) a. A‴ → (Deg‴) - A″
b. Adv‴ → (Deg‴) - Adv″
(6.22) Q‴ → (Deg‴) - Q″
(6.23) N‴ → (Art‴) - N″

Using feature notation, these can be collapsed as (6.40).

$$(6.40)\quad \begin{bmatrix} X \\ \langle +\text{Subj} \\ +\text{Comp} \rangle \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''' \rightarrow \left(\begin{bmatrix} \langle +\text{Subj} \rangle \\ -\text{Obj} \\ -\text{Comp} \\ +\text{Det} \end{bmatrix}''' \right) - X''$$

The primary argument for the feature relationship of Deg and Art and for the X‴ attachment of Deg‴ has been the categorial ambiguity of *this*, *that*, and *enough*. This is admittedly not overwhelming evidence, but in the absence of any other basis for the decisions, we have followed the preferences dictated by the $\bar{X}$ Convention.

The irregularities of the degree system have been accounted for with various local rules. *More* and *most* trigger spelling rules which attach them to quantifiers and certain adjectives. Combinations of *less* and *least* with quantifiers (other than the combination *less far*) are ungrammatical and must be ruled out by some as yet unprincipled filter. *Enough* triggers two local rules, *Enough*-Shift and *Many–Much* Deletion.

Having determined the occurrence of Deg‴, we will now explore its internal structure.

6.5. Recursion in Degree Phrases

Quantifier phrases with certain Degs can be further modified by measure phrases and quantifier phrases.

(6.41) a. five times $\left\{\begin{matrix}\text{as}\\\text{that}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{many}\\\text{much}\\\text{far}\end{matrix}\right\}$

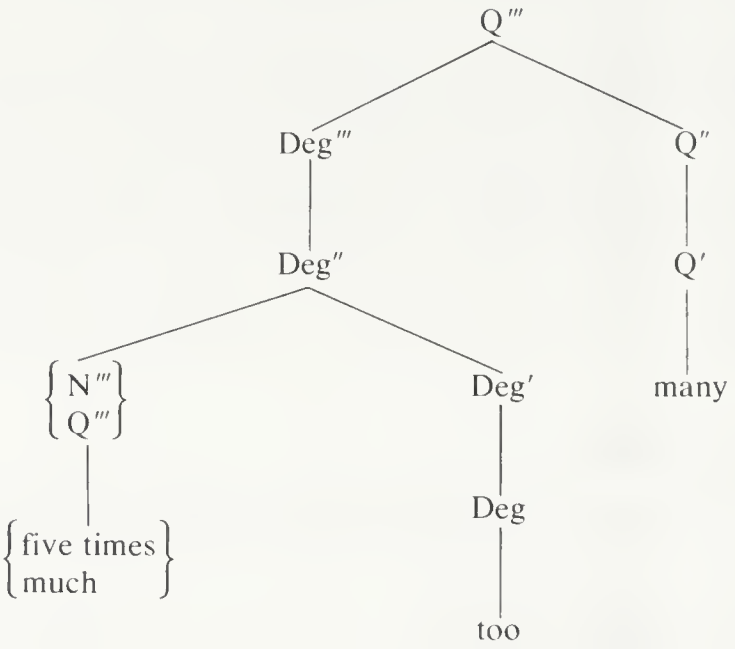
b. five times enough $\left\{\begin{matrix}\text{many}\\\text{much}\\\text{far}\end{matrix}\right\}$ ⇒ five times $\left\{\begin{matrix}\text{enough}\\\text{enough}\\\text{far enough}\end{matrix}\right\}$

c. $\left\{\begin{matrix}\text{five times}\\\text{five miles}\\\text{many}\\\text{much}\\\text{far}\\\text{a bit}\end{matrix}\right\}$ too $\left\{\begin{matrix}\text{many}\\\text{little}\\\text{much}\\\text{few}\\\text{far}\end{matrix}\right\}$

d. $\left\{\begin{matrix}\text{five times}\\\text{five miles}\\\text{much}\\\text{far}\\\text{a bit}\end{matrix}\right\}$ more $\left\{\begin{matrix}\text{many}\\\text{little}\\\text{much}\\\text{few}\\\text{far}\end{matrix}\right\}$ ⇒ $\left\{\begin{matrix}\text{five times}\\\text{five miles}\\\text{much}\\\text{far}\\\text{a bit}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{more}\\\text{less}\\\text{more}\\\text{fewer}\\\text{farther}\end{matrix}\right\}$

The presence of the measure phrases and quantifier phrases depends on the presence of the degree words, as can be seen by eliminating the degree words in (6.41). Furthermore, which measure phrases and quantifier phrases are possible depends on the choice of degree word, as can be seen by substituting other degree words into (6.41). This suggests that the measure phrases and quantifier phrases are in the specifiers of the degree words. Bowers (1968b) and Selkirk (1970) therefore argue that this recursion is through the Deg‴, as in (6.42).

(6.42)

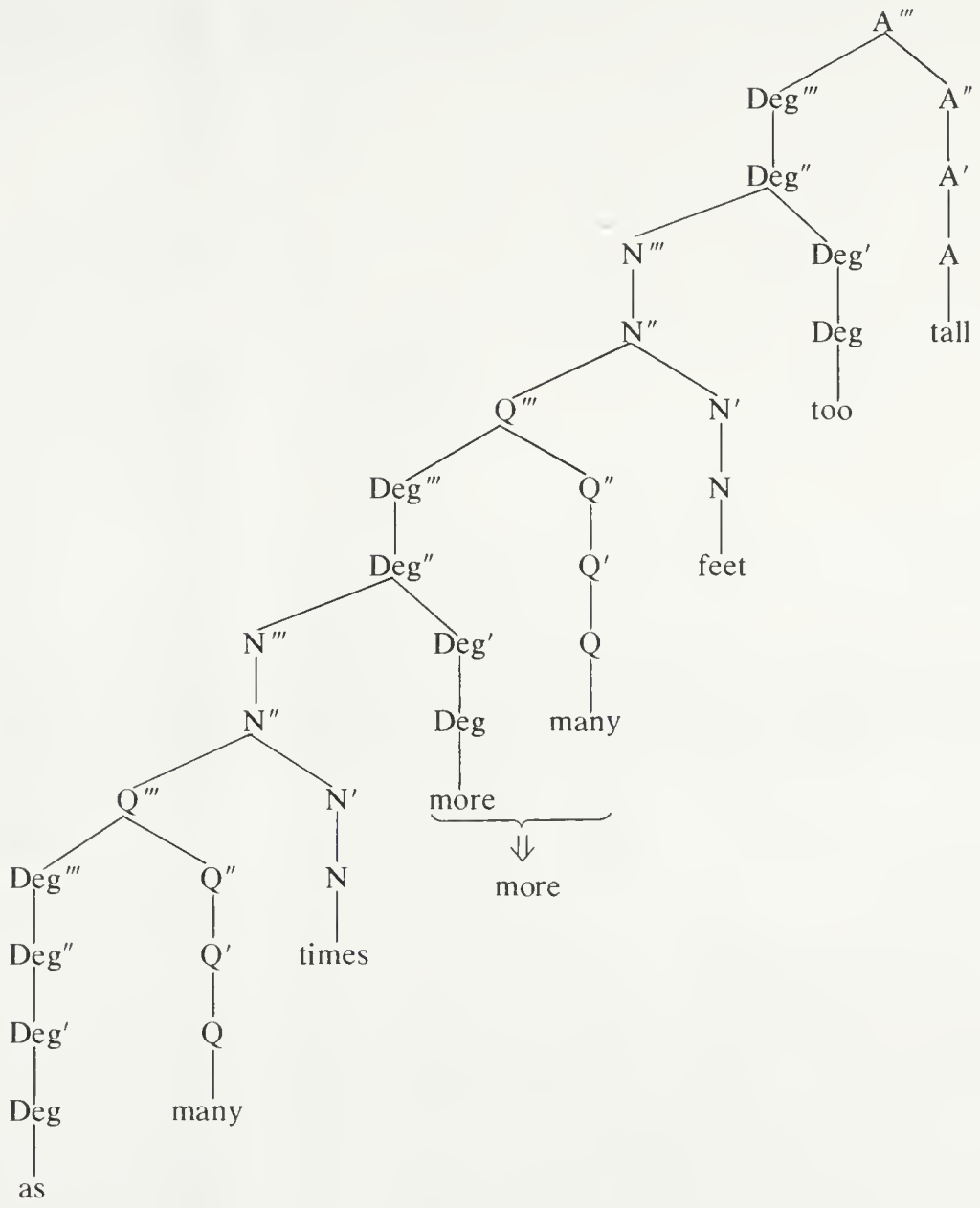


In (6.42) I have attached the new specifier to the Deg″ rather than to Deg′ or Deg‴. This choice enables us to generate these specifiers with the phrase structure rule (6.43), which generalizes with rules (6.15)–(6.19) of section 6.3. Furthermore, the definition (6.13) of the grammatical relation *measure phrase* can apply with complete generality. These generalizations argue that Deg is a full-fledged participant in the $\bar{X}$ system, as predicted by the Uniform Three-Level Hypothesis of phrase structure.

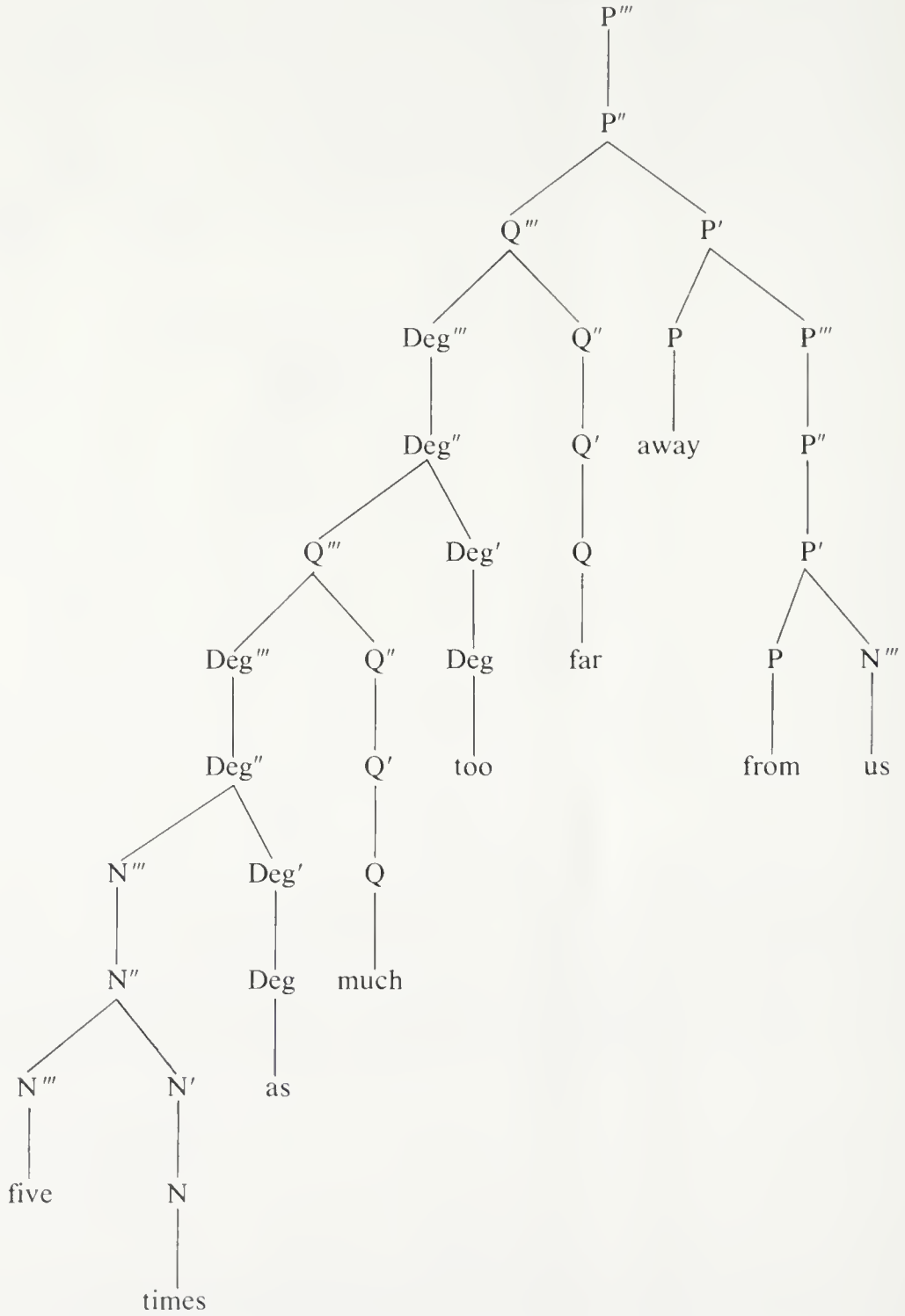
(6.43) $\text{Deg}'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - \text{Deg}'$

Given this measure phrase or QP in the Deg″ specifier, we are in a position to generate such recursive wonders as (6.44), for which it is left to the reader to find a context, if (s)he cares.

(6.44) a. as many times more feet too tall



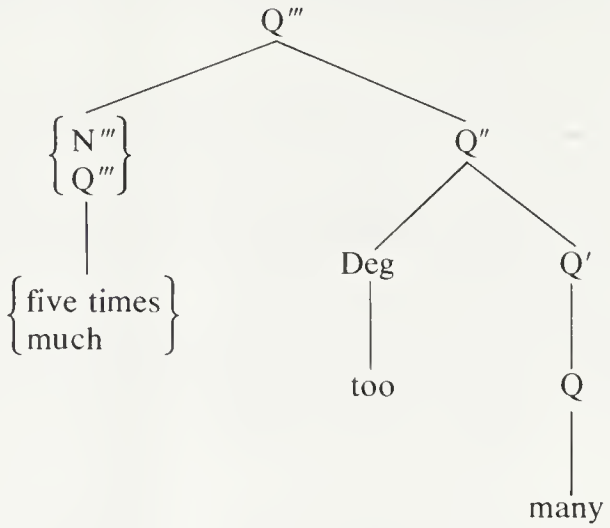
b. five times as much too far away from us



These structures are generated by the recursion of the two basic phrase structure rules developed in this chapter, the one generating a Deg‴ in various X‴ specifiers, and the other generating NP or QP in a wide range of X″ specifiers.

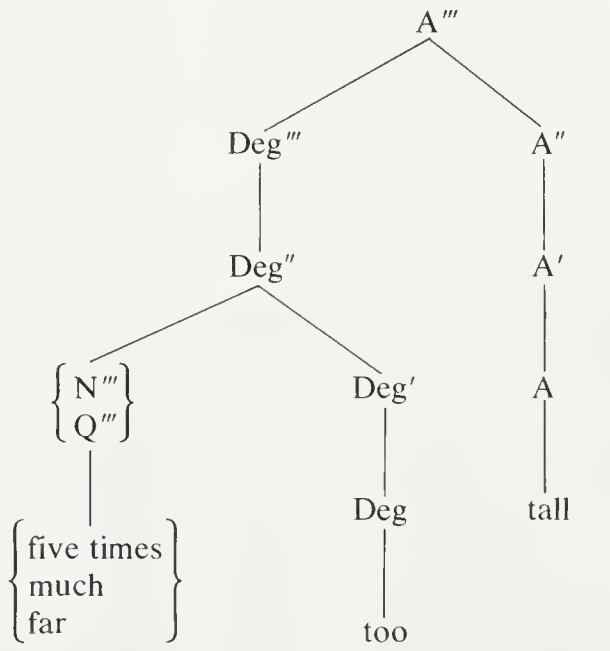
Bresnan's analysis of these constructions is different. She claims that the recursion of measure phrase and QP goes (in our terms) through the Q‴. Thus, instead of (6.42) she proposes the structure (6.45).

(6.45)



However, if we adopt the structure of A‴ argued for in sections 6.3 and 6.4.2, the X̄ Convention compels us to adopt (6.42): since most adjectives do not subcategorize a quantifier, the only structure available for *much too tall* is (6.46), parallel to (6.42). No structure can be constructed for *much too tall* that parallels (6.45).

(6.46)



Since Bresnan purports to eliminate (6.42) as a possible structure, it is incumbent on us to deal with her argument against it.

Bresnan points out that (6.45) describes *too many* as a constituent but (6.42) does

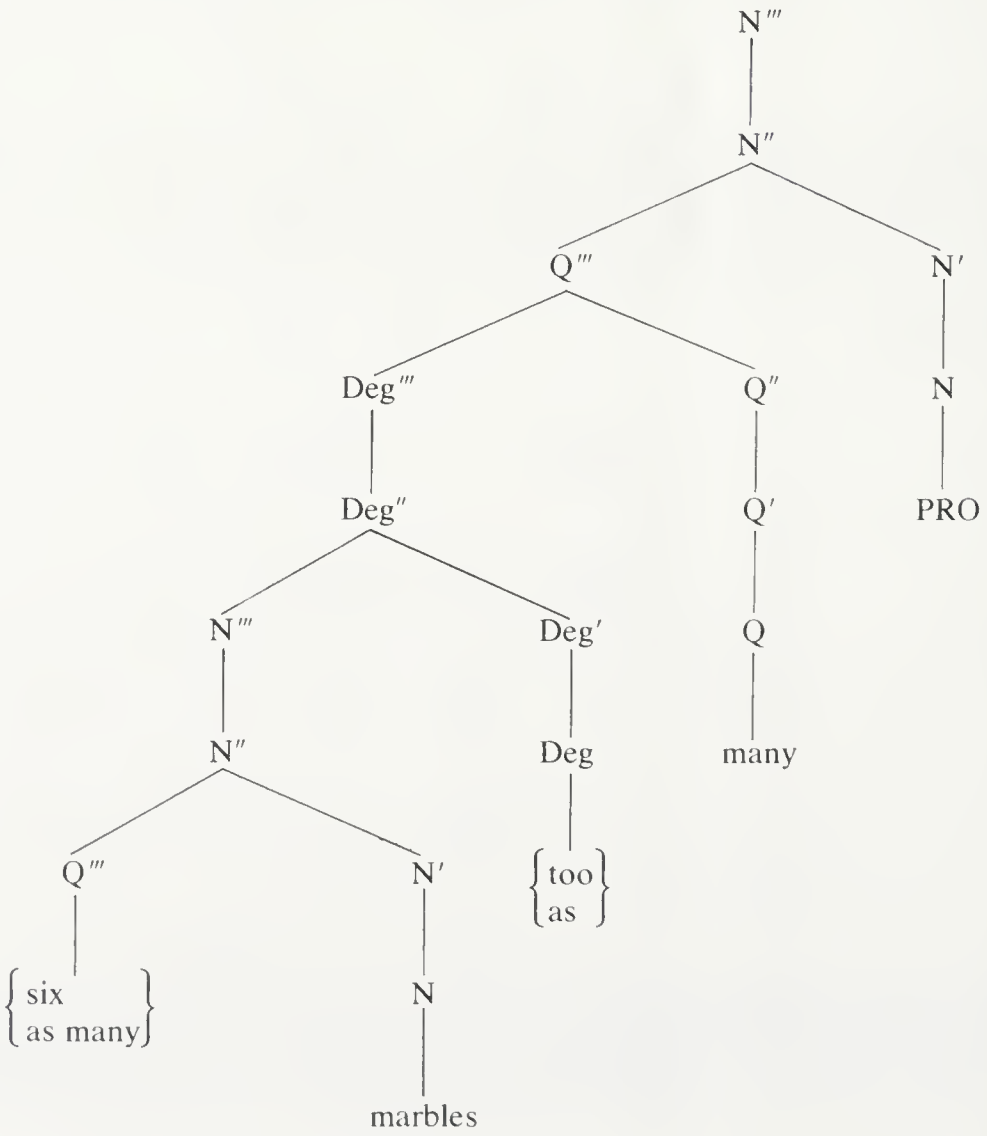
not. She then claims that there is a transformation that moves *too many*; hence it is a constituent, and (6.45) must be correct. The paradigm in question is (6.47).

(6.47) a. I have as many too many marbles as you.
b. I have as many marbles too many as you.
c. I have six too many marbles.
d. I have six marbles too many. (= Bresnan's (132)–(133))

(6.47a,c) contain recursive degree phrases like (6.41), i.e. either structure (6.42) or (6.45). Bresnan claims that (6.47b,d) are derived from (6.47a,c) by an optional reordering transformation which moves *too many* to the right of *marbles*, a rule called QP-Shift. Since only under hypothesis (6.45) is *too many* a constituent in (6.47a,c), the requirement that only constituents be moved favors (6.45) over (6.42).

I claim, however, that there is no such transformation as QP-Shift, but rather that (6.47b,d) are base-generated forms with the following structure:

(6.48)



According to this claim, $\left\{\begin{matrix}six\\as\ many\end{matrix}\right\}$ *marbles* in (6.47b,d) is functioning as a measure phrase in Deg''', just as *five times* is in (6.42). By contrast, the QP-Shift theory of (6.47b,d) claims that *marbles* is the head and does not form a constituent with $\left\{\begin{matrix}six\\as\ many\end{matrix}\right\}$.

To differentiate between the theories, observe first that the string *six marbles too many* is paralleled by *five days too long* (*after the accident*), *five miles too far* (*down the road*), and *five pounds too much,* in which measure phrases precede degree modifiers. The only question in establishing the parallelism is the unusual use of *marbles* as a unit of measure. But it is not implausible to say of a tray for marbles or an array of marbles that it is *six marbles wide; marbles* is a reasonable unit of measure if one happens to be counting marbles. Such is the case in (6.48): presumably the interpretation of PRO must be fixed as *marbles*. The PRO posited as head in (6.48) can furthermore be replaced by a lexical item in certain (albeit limited) cases where the unit of measure differs from what is being measured, as in (6.49).

(6.49) a. This book took five reams too many pages.
b. This cake took five pounds too much sugar.

Thus the relevant features of the structure (6.48) are independently motivated, and the appropriate interpretation is assigned to it by regular rules. Hence QP-Shift appears unnecessary in order to generate the paradigm (6.47).

Furthermore, the data which motivate constraints on QP-Shift can be explained immediately by structure (6.48). Bresnan claims that QP-Shift must follow *Of*-Insertion in order to prevent (6.50) (Bresnan's (136)–(137)).

(6.50) a. I have six more of them.
b. *I have six of them more.
c. I have half a dozen too many of these marbles.
d. *I have half a dozen of these marbles too many.

However, contra Bresnan, an *of* can appear in this construction:

(6.51) a. I have hundreds more marbles.
b. I have hundreds of marbles more.
c. I have five pounds too much sugar.
d. I have five pounds of sugar too much.

The real constraint seems to be that the noun cannot be definite. But this is a general constraint on measure phrases:

(6.52) a. *five of them too far
b. *half a dozen of those miles too long

Hence the constraint follows from the claim that $\left\{\begin{matrix} six \\ as\ many \end{matrix}\right\}$ *marbles* is a measure phrase in (6.47b,d).

Second, Bresnan claims that QP-Shift must be restricted to "count" Qs:

(6.53) a. much too much bread
b. *much bread too much

But this restriction too is not quite accurate, as *much* appears after *sugar* in (6.51d). Again, the real restriction is a general one on measure phrases, namely that *they* must be count. Since by the present hypothesis *much bread* would have to be a measure phrase in (6.53b), it is ungrammatical. But in (6.51d) *five pounds of sugar* is count, so the construction is acceptable.

Thus the QP-Shift theory of (6.47b,d) requires special constraints on an extra transformation. The base theory uses an independently motivated base construction to describe (6.47b,d); the constraints on the paradigm follow automatically from independent properties of measure phrases.

Returning to the issue of (6.42) vs. (6.45), we see that Bresnan has not made a case for the constituency of $\left\{\begin{matrix} as \\ too \end{matrix}\right\}$ *many* in (6.41); hence we are free to choose (6.42) rather than (6.45) as the proper structure. (6.46) is therefore a viable structure for *much too tall,* precisely parallel to (6.42), and it is in turn possible for us to continue to maintain that most adjectives cannot take quantifiers.

Having defended the recursion of measure phrase and QP through the Deg″, we have completed our analysis of the degree system. Though we have not dealt with the complete range of data discussed by Bresnan, in particular the relation of *so* and *such,* our analysis here has been successful in dealing with the major points of Bresnan's theory, eliminating such irregularities as the nonoccurrence of *less* before most adjectives, the optional application of *Much*-Deletion and *Enough*-Shift under certain conditions, and the suspect transformation QP-Shift. The reader will have noticed that at each step of the analysis, the issue of whether most adjectives take quantifier phrases has played a crucial role, for it led to differences in the treatment of the exceptional adjectives *alike* and *different,* the position of degree phrases in A‴, the categorial status of *more, less,* and *enough,* and finally the recursion of degree phrases.[5]

[5] Lest the part of Bresnan's theory we have not dealt with should be thought unassailable, it is worth mentioning that there are problems with her treatment of the *such/so* alternation, for which I have been unable to find a more satisfactory solution. The crux of the issue is her transformation AP-Shift, which transforms *taller a man* (parallel in the base to *as tall a man*) into *a taller man,* and *so tall a man* into *such a tall man.* She claims (p. 308) that AP-Shift is obligatory when the NP lacks a determiner, so that underlying **so tall ϕ men* becomes *such tall men.* However, this alternation is not general, since none of the other constructions of the form *Deg A a N* permit a plural alternant:

(i) $\left\{\begin{matrix} \text{as} \\ \text{how} \\ \text{too} \\ \text{that} \end{matrix}\right\}$ tall $\left\{\begin{matrix} \text{a man} \\ \text{*men} \end{matrix}\right\}$

The analysis here has been carried out within the strong hypothesis of phrase structure proposed in chapter 3. We have shown that the entire range of specifier possibilities we have discussed is the product of two major phrase structure expansions which apply to a large number of categories. In arriving at these rules, we made the fullest possible use of the $\bar{X}$ Convention to make up for the scarcity of data within any single category, under the assumption that cross-category generalization is an important part of the evaluation measure in syntactic theory.

As a final note, we should briefly mention a sizable group of specifiers we have not touched on, which have a wide range of cross-category application, such as *not, almost, just, quite, nearly, practically, hardly, scarcely,* and others. Some of these are patently adverbs, and in NPs are replaced by their adjectival parallels; but the rest are of some as yet undetermined category, perhaps Deg. *Almost* is a good example of cross-category applicability—it can modify even the articles *every* and *all,* though up to now we have had no need for internal structure in Art‴.

(6.54) a. John almost fell. (in V‴)
b. The pudding has become almost glue. (in N‴)
c. The men look almost alike. (in A‴)
d. We pushed John almost out the window. (in P‴)
e. Almost every groundhog saw his shadow. (in Art‴)
f. John left almost happily. (in Adv‴)
g. You talk almost too much. (in Deg‴)

Only the categories Q‴, M‴, and Prt‴ are missing. Such behavior is not atypical among this class of modifiers. Study of this class within the lexicalist framework would no doubt yield a rich source of evidence for feature analyses and phrase structure rules.

6.6. Summary of Rules

The following phrase structure rules were developed in this chapter:

X‴ rules

(6.23) $N''' \rightarrow (Art''') - N''$
(6.22) $Q''' \rightarrow (Deg''') - Q''$
(6.26) a. $A''' \rightarrow (Deg''') - A''$
b. $Adv''' \rightarrow (Deg''') - Adv''$

This suggests that *such tall men* may not be a transform of underlying [*so tall*] *men,* but has some other source, parallel to *what tall men.*

On the other hand, *enough,* like the comparative, does not share this restriction to singular heads, since *tall enough men* is grammatical. Thus AP-Shift must be obligatory for comparatives (**taller a man*) and optional for *A enough a man,* and there must be a constraint ruling out the plural cases in (i). In all, the system is highly idiosyncratic.

There is one semantic factor worth mentioning. Many of the constructions which require this indefinite article—(i) plus *more of a man, so much of a man,* and perhaps others—can be used only predicatively. A similar restriction appeared on NPs with measure phrase specifiers such as *two parts water.* One might speculate whether the analysis of (i) involves some bizarre expansion of the measure phrase in the N″ specifier. On the other hand, *hardly a man* and *scarcely a man* can be used as subjects. Since they count men, and the phrases in (i) do not, they are semantically more like QP in the N″ specifier.

The categories expanded here are the class [−Obj, −Det]‴. The generalization is evident. As in the rules collected in section 4.5, NP is different by one feature. In that case many categories expanded X″ with a prehead Adv‴, but N‴ used an A‴, a difference in the feature *Comp*. Here the difference between Deg and Art is in the feature *Subj*. The same problems of felicitous collapsing arise here as did in that case.

X″ rules

(6.16) $V'' \rightarrow V' - (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''')$

(6.15) $N'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \\ \langle +\text{Comp} \rangle \end{bmatrix}''' - \langle (\text{of}) \rangle) - (A''')^* - N' - \ldots$

(6.19) $A'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - A'$

(6.17) $P'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - P'$

(6.18) $P'' \rightarrow P' - (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''')$

(6.43) $\text{Deg}'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - \text{Deg}'$

To this we can add (6.55), since *much differently* is an Adv‴, and none of the adjectives that allow measure phrases have corresponding adverbs.

(6.55) $\text{Adv}'' \rightarrow (Q''') - \text{Adv}'$

(or, perhaps with use of the N‴ being vacuous for semantic reasons)

$\text{Adv}'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - \text{Adv}'$

Thus the rule creating this X″ specifier expansion applies to all [+Comp] categories and all [−Subj, −Obj] categories.

One serious problem in collapsing these rules, of course, is that of the order of the N‴/Q‴ term in V″ and P″. We discussed this in section 4.5, and I have nothing to add here.[6]

[6] We have not, however, touched on the problem of combining these X″ rules with the rules of chapter 4. The major difficulty is with the prehead Adv‴ in A″, P″, and Adv″. Consider A″, for example. (6.19) must

We have developed the following transformations.

(6.28) *More/most* – Q spelling rules
More/most – A spelling rules } generalizable

(6.32) Enough-*Shift* (revised)
enough – $X^0 \Rightarrow \phi - 2 + 1$
OBLIGATORY

(6.34) *Many–Much* Deletion
$enough - \begin{Bmatrix} many \\ much \end{Bmatrix} \Rightarrow 1 - \phi$
OBLIGATORY

The results of this chapter should be by now obvious enough that I need make no further polemic about how the specifier system demonstrates the virtues of the $\bar{X}$ Convention and the proposed set of syntactic distinctive features. It is worth pointing out again, though, that the inclusion of Q‴ and Deg″ in the generalized phrase structure rules is a strong argument that the $\bar{X}$ Convention and the Three-Level Hypothesis apply not only to the major categories N, V, A, and P, but to minor categories as well.

Finally, for convenience, here is a summary of the major differences between our theory and Bresnan's.

	Bresnan	*$\bar{X}$ Syntax*
1.	Adjectives subcategorize quantifiers	Most adjectives do not subcategorize quantifiers
2.	A few adjectives such as *alike* and *different* govern optional rather than obligatory *Much*-Deletion	A few adjectives such as *alike* and *different* subcategorize quantifiers
3.	*Enough* is a Q	*Enough* is a Deg and an Art
4.	*Much* deletes before adjectives	*Much* and *many* delete after *enough*

---

be combined with (4.52).

(4.52) A″ → (Adv‴) - A′ -. . .

The Adv‴ and the N‴/Q‴ cannot cooccur; for example **five miles incredibly wide* and **much surprisingly alike* are out. This would initially suggest that the rule collapses N‴, Q‴, and Adv‴ into a single term. Yet the feature configuration that picks out these three terms is the improbable $[-\text{Obj}, -\text{Det}, \begin{Bmatrix} +\text{Subj} \\ -\text{Comp} \end{Bmatrix}]$.

Consideration of the N″ rule suggests that this is the wrong way to collapse (6.19) and (4.86) anyhow. For in N″ the parallel N‴/Q‴ and A‴ terms are separate for good syntactic reasons, for example, *many tall men*. This suggests that for the sake of generality the A″ rule must be (i), in principle allowing the N‴/Q‴ and the Adv‴ to cooccur.

(i) $A'' \rightarrow (\begin{bmatrix} +\text{Subj} \\ -\text{Obj} \\ -\text{Det} \end{bmatrix}''') - (\text{Adv}''') - A'$

But since both modifiers have the semantic effect of designating a degree, there are semantic reasons for their noncooccurrence. We may thus invoke a principle similar to the Specifier Constraint of section 5.2 in justifying (i). Similar arguments apply in dealing with the rules for P″ and Adv″.

5.	*-er* is the comparative Deg	*More* and *less* are comparative Degs
6.	*Less far* cannot be generated	*Less much,* etc., must be ruled out
7.	Adjectives are compared by means of a Q'''	Adjectives are compared by means of a Deg'''
8.	Degree phrase recursion is through Q'''	Degree phrase recursion is through Deg'''

I would like to emphasize again here my debt to Bresnan's work. Despite the substantive difference between our analyses, the spirit of the work is in total agreement. Though I feel the analysis here is superior in several respects, it could not have been begun were it not for Bresnan's insight in organizing the mass of data in what I feel is so close to final form.

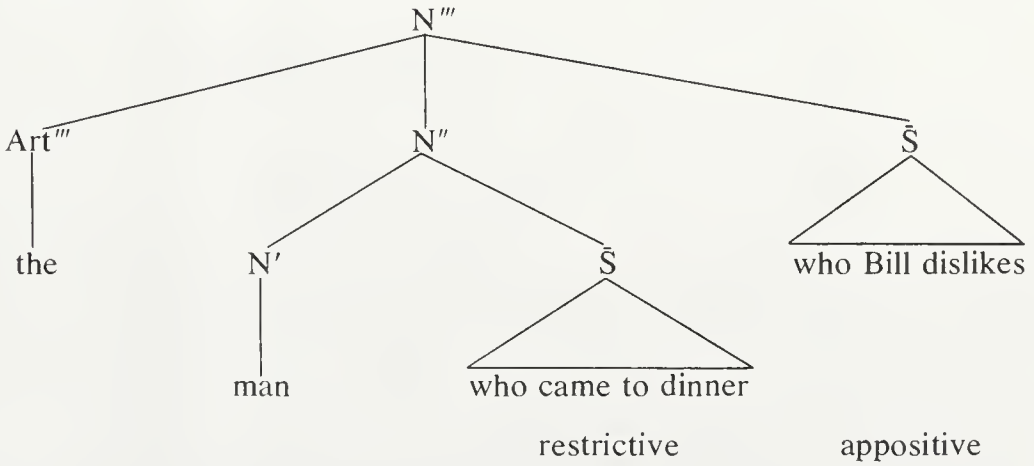
7: Relative Clauses

7.1. Introduction

This chapter is concerned with where relative clauses are attached in deep structure. It is concerned only secondarily with the internal structure of relative clauses and with the processes of relative pronoun formation and *wh*-preposing, since these have little bearing on the form of the base.

Chapter 4 claimed that restrictive relative clauses are daughters of N″ and that appositives are daughters of N‴, as in (7.1).

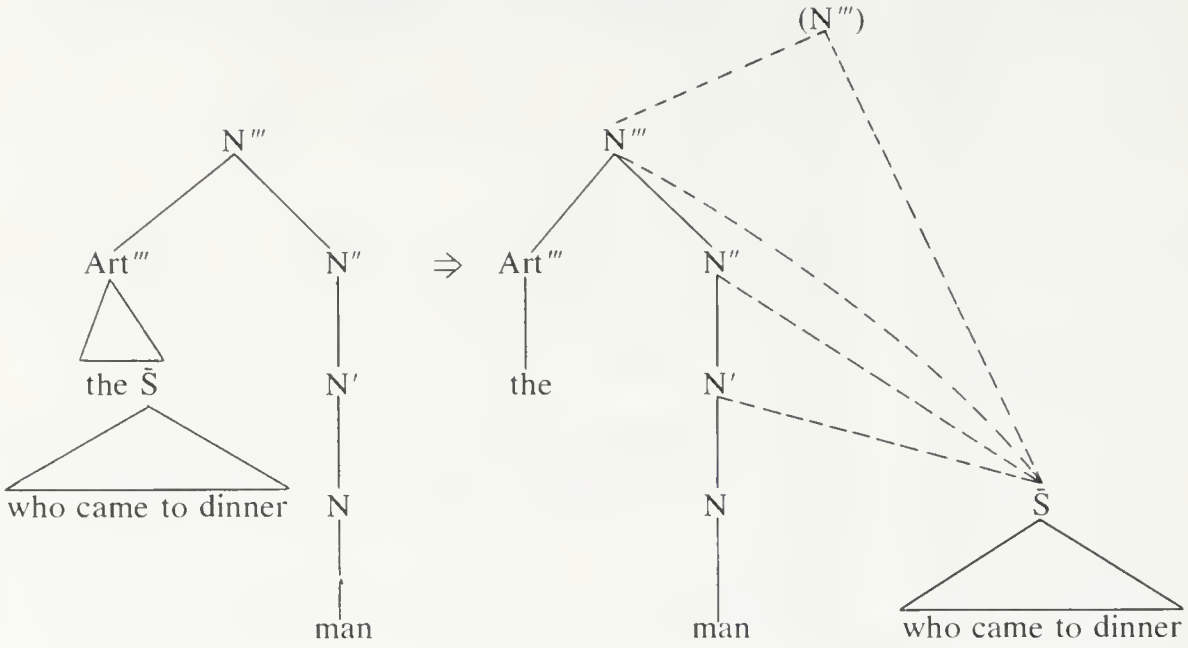
(7.1) the man who came to dinner, who Bill dislikes



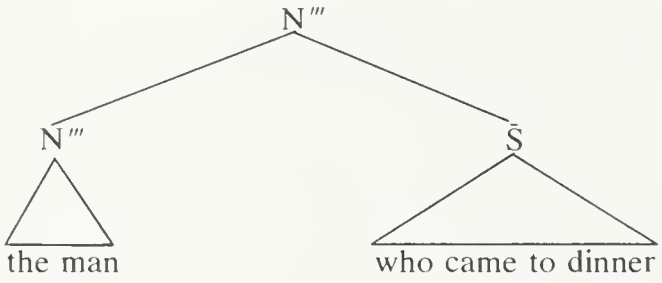
Though this claim fit in nicely with the general framework proposed in chapter 4, we postponed defending it in any detail until the present chapter, where we will show how it accounts for many well-known facts about relatives and compare it with other widely accepted analyses of relative clauses. We will show that our theory, the "NP-complement" theory, accommodates the data used as evidence for other theories at least as well as the other theories themselves do.

The major theories with which the NP-complement theory is in contention are the "determiner" theory, illustrated in (7.2), and the "Chomsky-adjoined" theory, illustrated in (7.3).

(7.2) the man who came to dinner



(7.3)



The determiner theory, advocated by Smith (1964) and Stockwell, Schachter, and Partee (1973), among others, is consistent with the constraints on phrase structure rules adopted in chapter 3. But since this theory requires an obligatory extraposition, the NP-complement theory is prima facie less complex. Hence any evidence that the NP-complement theory is equally adequate must be taken very seriously.

The popular Chomsky-adjoined theory, advocated by Ross (1967), is appealing because of the simplicity of stating the antecedent of the relative pronoun: it is just the entire lower NP. However, (7.3) is not a possible structure within the phrase structure rule schema of chapter 3, which requires every X^i to immediately dominate an X^{i-1}. If at all possible, we would like not to have to weaken the schema for the sake of relative clauses.

We will not consider the hypothesis in which restrictive relative clauses originate as sentences conjoined to some clause dominating the NP, as advocated by Thompson (1971), for example. Section 7.9 will advance arguments against Ross's (1967) well-known theory that appositives are so derived, and all these arguments apply with equal or greater force to restrictive relatives.

The next two sections deal with the relatively simple problems presented by

appositives, after which five sections are devoted to the much more complex issues involved in restrictives. In many cases, we will show that the theory of complements laid out in chapter 4 uncovers otherwise unexpected generalizations which competing theories cannot express—in other words, that what the $\bar{X}$ Convention leads us to believe is the simplest solution is in fact simplest.

7.2. Appositives

To start with, an adequate theory of appositives must account for numerous differences between restrictive relatives and appositives. Appositives must appear to the right of restrictives:

(7.4) a. The man that came to dinner, who was drunk, fainted.
b. *The man, who was drunk(,) that came to dinner fainted.

An NP may contain two or more concatenated ("stacked") restrictive relatives; but an NP may take multiple appositives only if they are conjoined:

(7.5) a. the man who came to dinner who hated lox
b. *the man, who came to dinner, who hated lox
c. the man, who came to dinner and who hated lox

Appositives, unlike restrictives, cannot be introduced by the complementizer *that* or ϕ instead of a *wh*-word:

(7.6) The man, $\left\{\begin{matrix}\text{*(that) Bill saw} \\ \text{who Bill saw}\end{matrix}\right\}$, sneezed.

Appositives, unlike restrictives, must be able to modify proper nouns:

(7.7) a. John, who came to dinner, sneezed.
b. *John that came to dinner sneezed.

As pointed out in chapter 4, appositives can occur in any $\begin{bmatrix} X''' \\ +\text{Comp}\end{bmatrix}$, whereas restrictives can occur only in NP:

(7.8) a. Relative clause formation is obligatory in NPs, $\left\{\begin{matrix}\text{which} \\ \text{*that}\end{matrix}\right\}$ accounts for the difference in surface shape. (in V‴)
b. That Sheila was beautiful, $\left\{\begin{matrix}\text{which} \\ \text{*that}\end{matrix}\right\}$ she was, was not realized until later. (in A‴)[1]
c. Solving this problem will take from now until doomsday, $\left\{\begin{matrix}\text{which} \\ \text{*that}\end{matrix}\right\}$ is more time than we've got. (in P‴)

[1] (7.8b) is taken from Ross (1969), where it is used as an argument that APs must be dominated by NP. The $\bar{X}$ Convention, of course, allows a different account, namely cross-category generalization.

The differences in intonation, complementizer, and distribution between restrictives and appositives argue rather strongly that the two kinds of clauses have different syntactic sources. The fact that appositives always follow restrictives is explained by the NP-complement theory as a consequence of restrictives being N″ complements and appositives being N‴ complements. Furthermore, the comma intonation before appositives is characteristic of all X‴ complements, since sentence-final sentence adverbs, parentheticals, and the like have similar intonation:

(7.9) This ice cream tastes like spaghetti, $\left\{\begin{array}{l}\text{of course}\\ \text{probably}\\ \text{I bet}\\ \text{no doubt}\end{array}\right\}$.

Since the two kinds of relatives arise from different base positions, the difference illustrated in (7.5) can be a consequence of the base rules. We will argue in section 7.6 that (7.5a) is generated by the phrase structure rule (7.10), which permits multiple clauses.

(7.10) $N'' \rightarrow \ldots N' \ldots (\bar{S})^*$

The rule for appositives, on the other hand, lacks the star, so multiple appositives are possible only through conjunction under the single $\bar{S}$.

(7.11) $N''' \rightarrow \ldots N'' \ldots (\bar{S})$

Note, by the way, that appositives do not concatenate in S, AP, or PP either; again only conjunction is possible:

(7.12) a. Relative clause formation is obligatory in NPs, which accounts for the difference in surface shape, $\left\{\begin{array}{l}*\phi\\ \text{and}\end{array}\right\}$ which only your theory predicts.

b. That Sheila was beautiful, which she was, $\left\{\begin{array}{l}*\phi\\ \text{and}\end{array}\right\}$ which she hadn't been two years previously, was not realized until later.

c. Solving this problem will take from now until doomsday, which is longer than most problems take, $\left\{\begin{array}{l}*\phi\\ \text{and}\end{array}\right\}$ which is more time than we've got.

I intend to provide no account of the differences in complementizer treatment between restrictives and appositives. However, the fact that the clauses arise in different base positions makes it relatively simple to distinguish between them in whatever rules insert or delete relative pronouns and complementizers. Thus the NP-complement theory is not inconsistent with the observed differences.

The difference in base position of the two kinds of relatives also implies that they have different grammatical relations in the sentence, that is, they are interpreted by different projection rules. Two different types of projection rules are involved: the general rules for N″ or N‴ complements, and the rules for interpreting relative pronouns in N″ relatives and in X‴ relatives. One consequence of the former was

mentioned in section 4.1: restrictives, like other X″ complements, can be focused and negated, but appositives, like other X‴ complements, cannot:

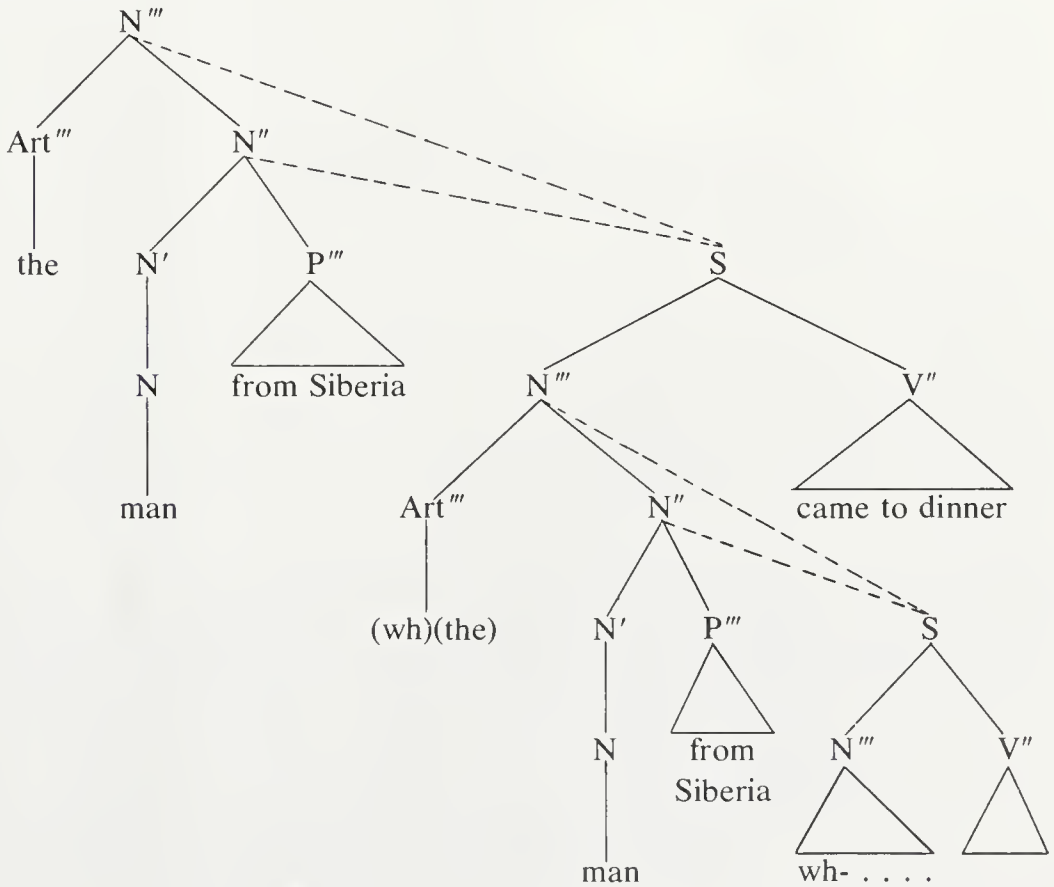
(7.13) a. We didn't talk to the man who married SUSAN. (we talked to the man who married JAYNE)
b. *We didn't talk to the man, who married SUSAN.

The second kind of projection rule deals with the relationship between the relative pronoun and the head. The rules for interpreting the relative pronoun in appositives (section 7.3) and restrictives (section 7.8) will account for the differences observed in (7.7) and (7.8), as well as a number of other differences to be pointed out in later sections. Section 7.9 will show that a competing hypothesis of appositives, deriving them from underlying conjunction, encounters serious difficulties.

7.3. Interpretation of the Appositive Relative Pronoun

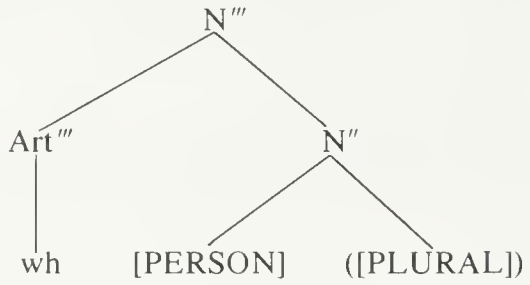
Before the development of interpretive theories of anaphora, the interpretation of the relative pronoun was a potential obstacle to the NP-complement theory. For example, in the analysis of Chomsky (1965), the deep structure of a relative pronoun is taken to be a literal copy of its antecedent; most work of that period makes a similar assumption. But if the relative clause is inside the NP, the relative pronoun cannot originate as a copy of the head, for, with either N″ or N‴ attachment of the head, an infinite regress results:

(7.14)

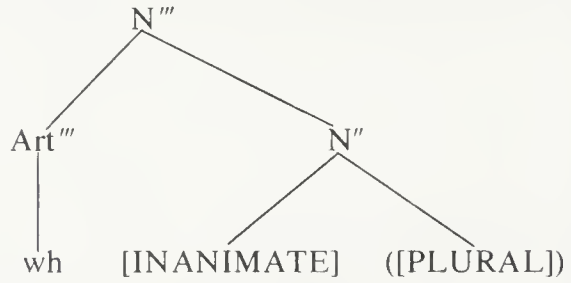


This problem for the NP-complement theory can be eliminated by adopting an interpretive theory of the relative pronoun, parallel to the by now familiar interpretive theory of definite pronouns of Jackendoff (1972), Wasow (1972), and others. In the interpretive theory, the relative pronoun is a relative pronoun in deep structure, not some more fully specified NP form. The only semantic features it needs to carry are those necessary for agreement with the head and with the verb of the relative clause. Its semantic structure, if it were composed of syntactically separable elements, would look like this:

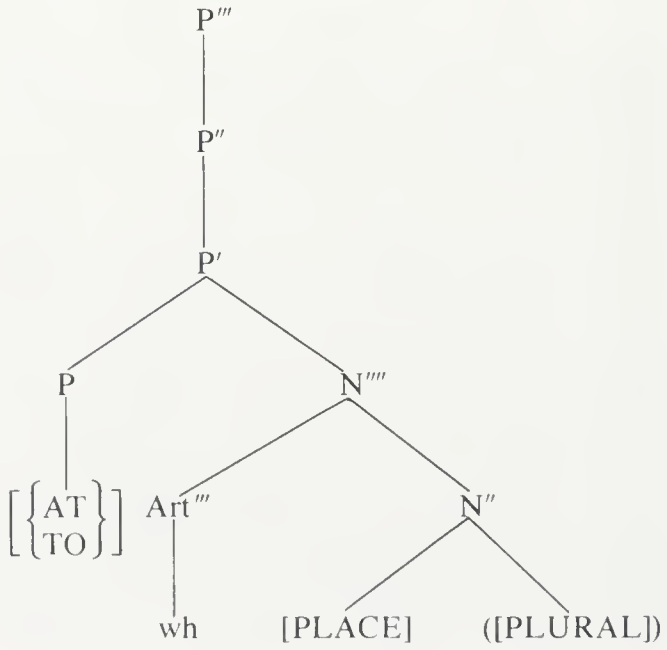
(7.15) a. who



b. which



c. where



The bracketed expressions in capitals represent semantic markers of the usual sort. The *wh* is a placeholder which triggers rules of semantic interpretation.

The simplest possible rule of anaphora is one which marks the relative pronoun anaphoric to the entire dominating NP. The infinite regress of (7.14) does not arise, since such a semantic rule does not provide the relative pronoun with internal syntactic structure. Rather, two NPs are simply designated as coreferential, and one happens to dominate the other. Such a rule seems appropriate for appositive relative pronouns, though not for restrictive relative pronouns, for which we will propose a different sort of rule in section 7.8.

The appositive rule can be stated like this:

(7.16) *Appositive* Wh-*Interpretation*

$\begin{bmatrix} X \\ +wh \end{bmatrix}'''$ is anaphoric to Y''', in the configuration

$[_{Y'''} \ldots Y''[_{\bar{S}}[_{Comp}. \ldots X''' \ldots] S]]$

(7.16) has been stated in its most general form, so as to account for appositives in V''', A''', and P''' as well as in N'''. The proper notion of anaphora for N''' appositives is obviously coreference; but in other Y''' the proper interpretation of the notion "anaphoric" is less clear. Whatever that notion may be, it is in any event independently motivated, since the demonstratives *it* and/or *that* exhibit exactly the same kind of anaphora as *which*. Compare the following pairs:

(7.17) a. John likes your idea, but it's crazy. (in N''')
John likes your idea, which is crazy.

b. Bill came late, and that bothered Susan. (in V''')
Bill came late, which bothered Susan.

c. Bill is drunk all the time—is that how you'd like to be? (in A''')
Bill is drunk all the time, which is probably how you'd like to be.

d. Bill went into the tree, and that's where I'd like to go too. (in P''')
Bill went into the tree, which is where I'd like to go too.

That is, relative pronouns in appositives can be anaphoric to the same constituents as ordinary demonstrative pronouns can. However such anaphora is defined for demonstratives, the definition will be equally appropriate for appositive relative pronouns as well.

A rule of anaphora like (7.16) makes it very natural that appositives can be applied to proper nouns; the burden is thus on the rule for restrictives to explain why restrictives cannot be applied to proper nouns. However, there is also a case in which restrictives can occur but appositives cannot, with certain quantifiers in the head:

(7.18) a. $\left\{\begin{matrix} \text{Any} \\ \text{Every} \end{matrix}\right\}$ man who drives a Cadillac is insane.

b. *$\left\{\begin{matrix} \text{Any} \\ \text{Every} \end{matrix}\right\}$ man, who drives a Cadillac, is insane.

Ross (1967) cites this construction as evidence for the conjunction source, since the putative conjoined source of (7.18b) is also ungrammatical.

(7.19) *{Any / Every} man is insane, and he drives a Cadillac.

The present theory provides an explanation as well. Observe that *everyone* and *no one* can be coreferential with a pronoun which is in a restrictive relative clause or in other subordinate clauses within the scope of the quantifier (understood anaphora indicated by italics):

(7.20) a. *Everyone* there had a wife who loved *him*.
b. *No one* wanted Sue to waken *him*.
c. *Everyone* got from Boston to a place *he* had been to before.

But such anaphora is impossible if the *him* is within an appositive.

(7.21) a. **Everyone* there had a wife, who loved *him*.
b. **No one* wanted Sue to leave, which suited *him*.
c. **Everyone* got from Boston to Cincinnati, which seemed to *him* like a long way.

The ungrammaticality of (7.21a,b,c) follows from the more general fact that appositives are immune to the scope of quantifiers and negation. (7.22) confirms this fact by demonstrating that a negation outside an appositive cannot condition *some–any* type alternations within the appositive, though *any* can appear in restrictive modifiers in parallel positions.

(7.22) a. I didn't see a man who had had any drinks.
I didn't see Bill, who had had {some / *any} drinks.
b. That Sheila isn't nervous to any appreciable extent surprises me.
That Sheila isn't nervous, which she {sometimes / *ever} is, surprises me.
c. We couldn't get from Boston to a place any of us had been.
We couldn't get from Boston to Cincinnati, which seemed to {some / *any} of us like a short way.

Since two NPs can be coreferential only if they are subject to the same logical operators, and since only the pronouns in (7.20) and not those in (7.21) can be quantified by the subject, coreference is possible only in (7.20).[2] But the anaphora established between an appositive relative pronoun and its head by (7.16) is claimed to be of the same nature as anaphora with ordinary pronouns. Therefore, in (7.18b), the relative pronoun, which is within the appositive and hence obligatorily unquantified, is marked anaphoric to a quantified head. Such a situation is semantically anomalous, just as in (7.21); hence (7.18b) is ungrammatical.

[2] Cf. the Modal Coreference Condition of Jackendoff (1972, chapter 7), for elaboration of this point.

These two sections have shown that the NP-complement theory of relatives provides a principled distinction between restrictive and appositive relatives, and that an interpretive theory of the appositive relative pronoun is not only statable but also revealing. We now turn in some detail to restrictive relatives.

7.4. Constraints on the Determiner and the Head in Restrictives

This section and the next three will deal with four different kinds of data which provide apparent counterevidence to the claim that restrictive relatives are attached to N″ in underlying structure.

The first case concerns interaction between restrictive relative clauses and the determiners of their heads. These interactions have often been taken as evidence that the underlying position of restrictives is in the determiner (see (7.2)), since in such a theory the necessary constraints can be stated as constraints on the form of a single constituent; in the NP-complement theory, on the other hand, they must be stated in some sort of discontinuous fashion. This allegedly favors the determiner theory. We will show, however, that the determiner theory must be stretched to absurdity and the discontinuous environments are independently necessary.[3]

Let us start with some cases in which *the – X – Rel* can appear but *the – X* alone is ungrammatical.

(7.23) a. *the Paris
the Paris that I love
b. *the manner/way
the manner/way in which he did it
c. *the four of the boys
the four of the boys that came to dinner

There are also some cases (pointed out by Perlmutter (1970)) in which the proper article is dependent on the content of the relative clause:

(7.24) He greeted me with $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{the} \\ \text{*a} \end{array} \right\} \text{warmth I expected} \\ \left\{ \begin{array}{l} \text{*the} \\ \text{a} \end{array} \right\} \text{warmth I had not expected} \end{array} \right\}$

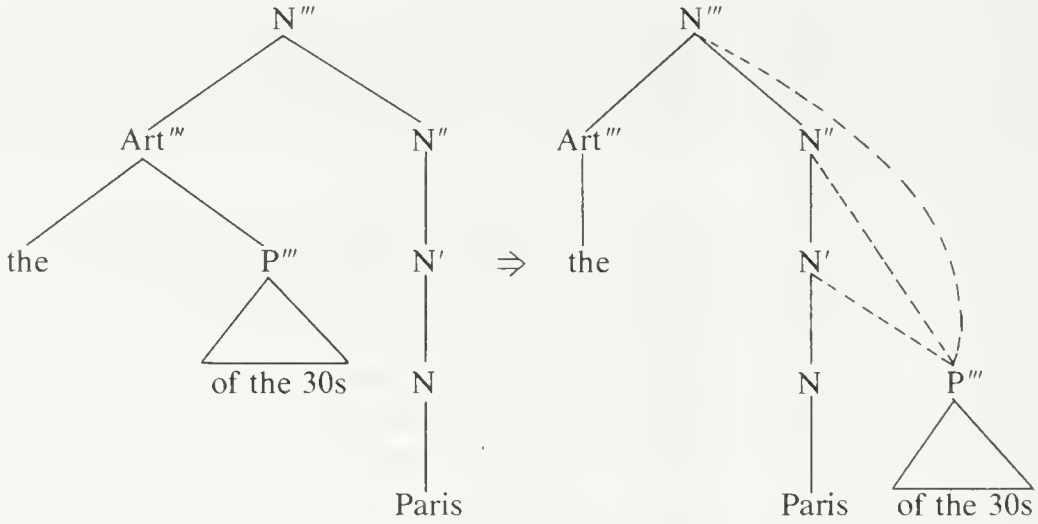
Consider first (7.23a). According to the determiner theory, the application of an Art‴ to a proper noun can be restricted to a particular kind of Art‴, namely *the* + $\bar{S}$. In the NP-complement theory, the restriction involves the cooccurrence of two distinct constituents, *the* before the proper noun and $\bar{S}$ after it; such a discontinuous constraint is more complex to state.

Vergnaud (1974), however, observes that there are phrases like *the Paris of the 30s* and *the Paris of my youth,* in which PPs satisfy the constraint on use of the definite article. These particular PPs, moreover, are not reduced relative clauses, since **the*

[3] I am particularly indebted in this section to the arguments of Reed (1974) and Vergnaud (1974).

Paris that $\left\{\begin{matrix}is\\was\end{matrix}\right\}$ $\left\{\begin{matrix}of\ the\ 30s\\of\ my\ youth\end{matrix}\right\}$ is ungrammatical. In order to prevent discontinuity in the statement of this constraint, the determiner analysis must propose that the PP also originates in the determiner and is extraposed:

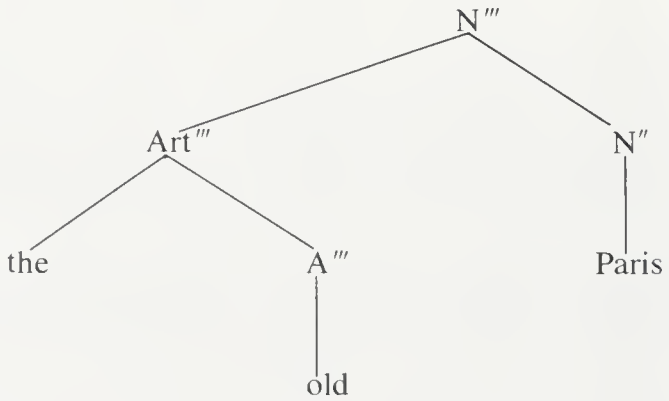
(7.25)



The deep structure in (7.25) of course entails adding an otherwise unmotivated option in the base rule for the determiner. The NP-complement theory, which countenances a discontinuous constraint anyway, can generate the PP in its surface position and simply generalize the constraint to a further class of modifiers.

Exploring further, we discover phrases like *the old Paris* and *the unknown Paris*, with prenominal adjectives satisfying the determiner constraint. Since the former of these means 'the part of Paris which is old' or 'the Paris of the old days', neither of its readings can be derived simply from the phrase *the Paris that* $\left\{\begin{matrix}is\\was\end{matrix}\right\}$ *old* by relative clause reduction. Thus a base-generated prenominal adjective is semantically as adequate as any other source. The determiner theory, in order to preserve the generality of the constraint for (7.23a), is forced to put the adjective in the determiner, thus:

(7.26)



There is of course no need for an extraposition rule this time, but the proposed structure looks even more suspicious than the one in (7.25).

Actually, a different generalization emerges from these examples: a certain class of restrictive modifiers (of any syntactic category) permits the use of the definite article with proper nouns. APs, PPs, and restrictive relative clauses can all be interpreted as restrictive modifiers, by whatever projection rules happen to apply to them. It follows then that there is no privilege in being attached to the determiner; hence (7.23a) is not evidence for the determiner theory of restrictive relatives.

Either of the theories, of course, can provide projection rules that interpret APs, PPs, and restrictive relatives all as restrictive modifiers. However, only the NP-complement theory permits the generalization, pointed out in chapter 4, that all restrictive modifiers are daughters of N″. This permits a single projection rule with a single generalized structural description to apply to all three categories. Such a generalization is impossible within the determiner analysis, since relative clauses originate in an entirely different position than prenominal adjectives or postnominal PPs do.[4]

As further evidence that relative clauses and PPs can perform the same semantic function, consider pairs such as these, pointed out by Vergnaud (1974):

(7.27) a. the reason $\left\{\begin{matrix}\text{for his disappearance}\\ \text{for which he disappeared}\end{matrix}\right\}$

b. the time $\left\{\begin{matrix}\text{of his departure}\\ \text{at which he left}\end{matrix}\right\}$

c. the manner $\left\{\begin{matrix}\text{of his arrival}\\ \text{in which he arrived}\end{matrix}\right\}$

Since in each case the PP and the relative are nearly synonymous, a grammar in which they are related to the NP by the same projection rule is to be preferred to one in which they are not.[5]

Summing up this argument, the supposed advantage of the determiner theory can be maintained only at the expense of positing deep structures like (7.25) and (7.26). The NP-complement theory permits a different sort of generalization, one that positions restrictive relatives so that the projection rule for restrictive modification can be of the simplest form.

A parallel argument can be constructed for (7.23b).

[4] The generalization is likewise impossible in Ross's theory of restrictives, though Ross allows such a variety of Chomsky-adjoined constituents and unrestricted transformations that it is difficult to tell what claims he could make here.

[5] Does the fact that these phrases cannot be *one*-pronominalized mean that both the PP and the relative are under N′ here?

(i) *the time $\left\{\begin{matrix}\text{of his departure}\\ \text{at which he left}\end{matrix}\right\}$ and the one $\left\{\begin{matrix}\text{of her arrival}\\ \text{at which she arrived}\end{matrix}\right\}$

(7.23) b. *the manner/way
the manner/way in which he did it

Here the head nouns cannot occur at all without the relative clause. The alleged advantage of the determiner analysis is that the constraint on the occurrence of these nouns is confined to a single constituent, the determiner. But again, the existence of PPs and adjectives that suit the same purpose calls into question the generality of such a solution:

(7.28) the way of the wise
the manner of his arrival
a new way
a pompous manner

The same argument applies: the generalization concerns the need for a restrictive modifier, of whatever category (or a possessive this time, as in *his manner*). Thus the only possible syntactic consideration is the generality of the configurations that can be interpreted as restrictive modifiers, and the advantage of the determiner analysis vanishes.

Similarly with (7.23c), where *the* is made acceptable not only by a relative clause, but by an appropriate PP or AP, as in (7.29).

(7.23) c. *the four of the boys
the four of the boys that came to dinner

(7.29) the four of the boys with green eyes
the usual four of the boys

And with (7.24):

(7.24) He greeted me with $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{the} \\ *\text{a} \end{array} \right\} \text{warmth I expected} \\ \left\{ \begin{array}{l} *\text{the} \\ \text{a} \end{array} \right\} \text{warmth I had not expected} \end{array} \right\}$.

(7.30) a. He greeted me with $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{the} \\ *\text{a} \end{array} \right\} \text{warmth of an old friend} \\ \left\{ \begin{array}{l} *\text{the} \\ \text{a} \end{array} \right\} \text{warmth of great intensity} \end{array} \right\}$.

b. He greeted me with $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{the} \\ *\text{a} \end{array} \right\} \text{usual warmth} \\ \left\{ \begin{array}{l} *\text{the} \\ \text{an} \end{array} \right\} \text{unusual warmth} \end{array} \right\}$.

Notice that at least some of these examples cannot be derived by reduction of relative clauses: **the four of the boys that was/were usual; *the four of the boys that were with green eyes; *the warmth that was of an old friend.* So again the constraints concern all restrictive modifiers, not just relative clauses.

Next consider the inability of proper nouns to take restrictive relative clauses:

(7.31) *John that came to dinner

The determiner analysis claims that this constraint is related to the inability of proper nouns to take determiners. But in fact, no restrictive modifiers at all are possible with proper nouns, so the generalization of the NP-complement theory accounts for this constraint adequately.

One might claim that a phrase like *poor John* is a counterexample to the generalization of the NP-complement theory. But *poor* here is not a restrictive modifier; it does not constrain the choice of possible referents of the phrase. Rather, something more akin to the projection rule for appositives must integrate *poor* into the interpretation of *poor John*. I leave open whether this entails a different structural position for such an adjective.

There remains one sort of argument for the determiner analysis, the inability of restrictive relatives to cooccur with a genitive NP in the determiner:

(7.32) *John's book that you stole

Here other restrictive modifiers are possible (e.g. *his old book, his book with a faded cover*), so the previous argument for the NP-complement theory does not apply. The determiner hypothesis predicts this ungrammaticality by claiming that when the determiner is filled by NP the expansion Art – S is unavailable.[6] Section 7.8 will show how this constraint can be explained as a consequence of the relative pronoun's interpretation, making the determiner theory's explanation unnecessary.

This section has thus shown that the motivations for the determiner analysis, namely the dependencies between the relative clause, the article, and the head, are (with one exception) consequences of more general constraints on all restrictive modifiers. Since restrictive relatives must be interpreted as restrictive modifiers no matter what syntactic analysis is assigned to them, there is in fact little ground for

[6] Many proponents of the determiner analysis, however, simply require the extraposition of the possessive to the form *the book of his that you stole*, a much more mechanical explanation, and one available to all three analyses of relatives. But a *syntactic* relation between the determiner genitive and this postposed genitive is very difficult to support, as argued in Jackendoff (1969b). For a different analysis of the postposed genitive, see section 5.3.4.

preferring the determiner analysis. Rather, the NP-complement analysis appears to allow the most general statement of the projection rule for restrictives, since the constraints show that restrictives in nearly all respects behave just like other daughters of N″.

7.5. Restrictive Relatives on Quantified NPs

Another problem in the analysis of restrictive relatives (to my knowledge, first pointed out by Dean (1966)) concerns the interpretation of relative clauses in NPs with quantifiers.

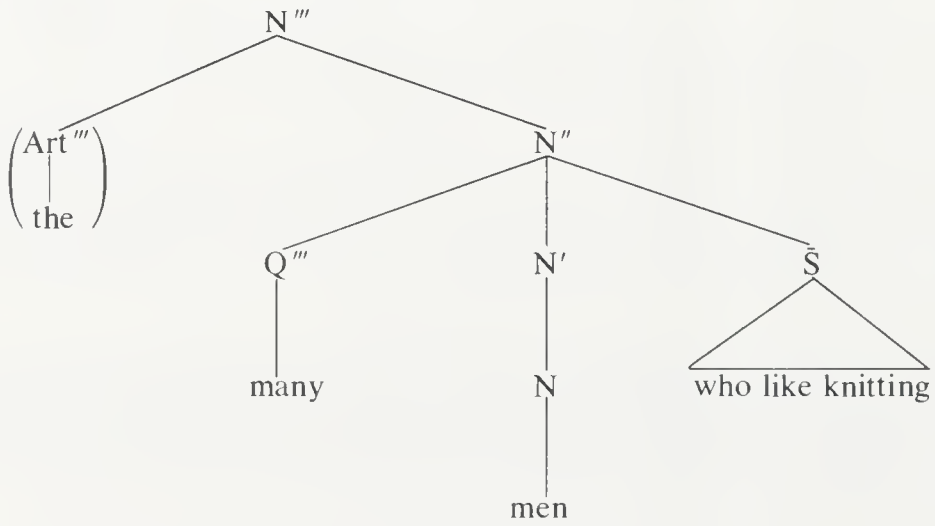
(7.33) Mary knows { a. many men who like knitting
b. many of the men who like knitting
c. the many of the men who like knitting }.

In (7.33a) the antecedent of the relative pronoun appears to be *men* or *many men*; in (7.33b) *the men*; in (7.33c), *the many*. The problem is to specify quantifier and relative clause structures which interact properly to create these interpretations.

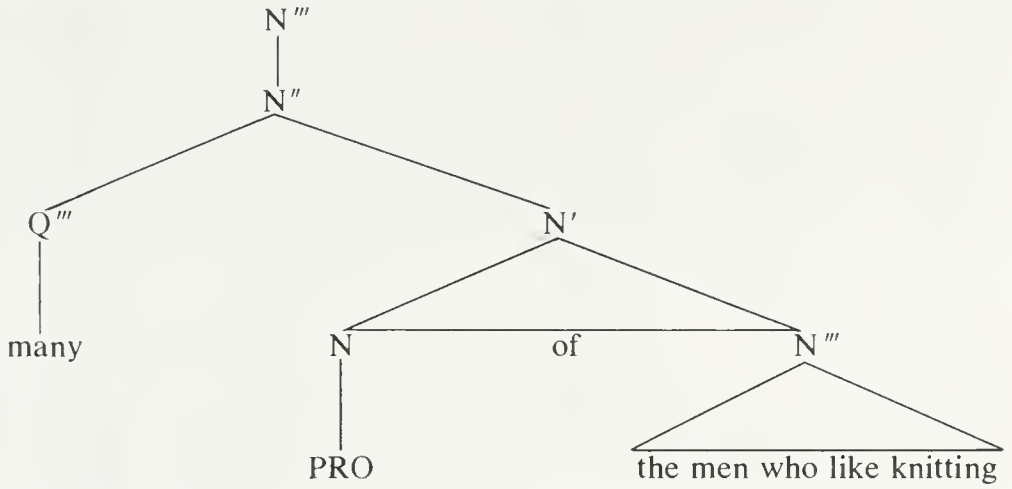
I shall not be concerned here with describing various attempts to account for these facts within other frameworks, for most writers seem to bog down hopelessly on this issue, mainly in accounting for the partitive structures. Rather I will be content to show that the analysis of partitives in section 5.3 interacts with the NP-complement theory to provide an adequate account of (7.33).

Here are the possible structures. When there is no partitive N′ complement, as in (7.34), the relative clause can be attached only in one position; when there is a partitive, as in (7.35), the relative can be attached to either the higher or the lower NP.

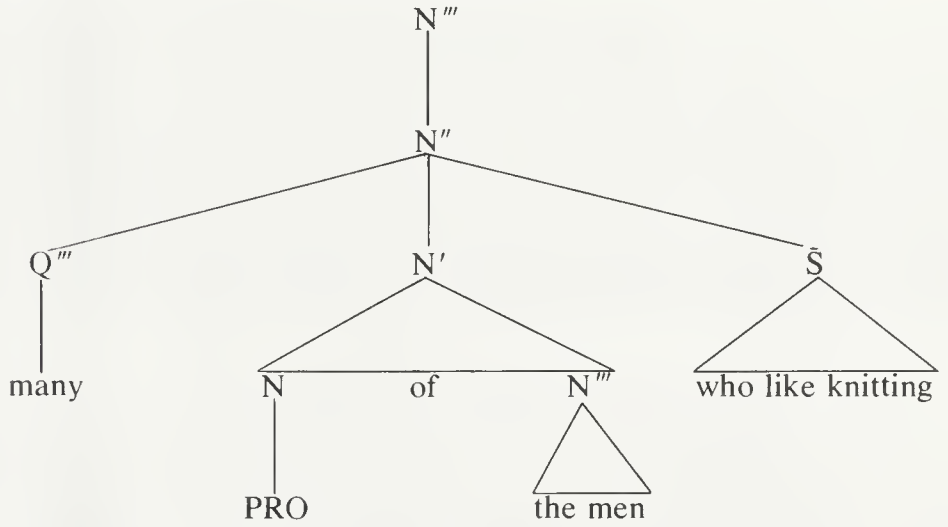
(7.34) (the) many men who like knitting



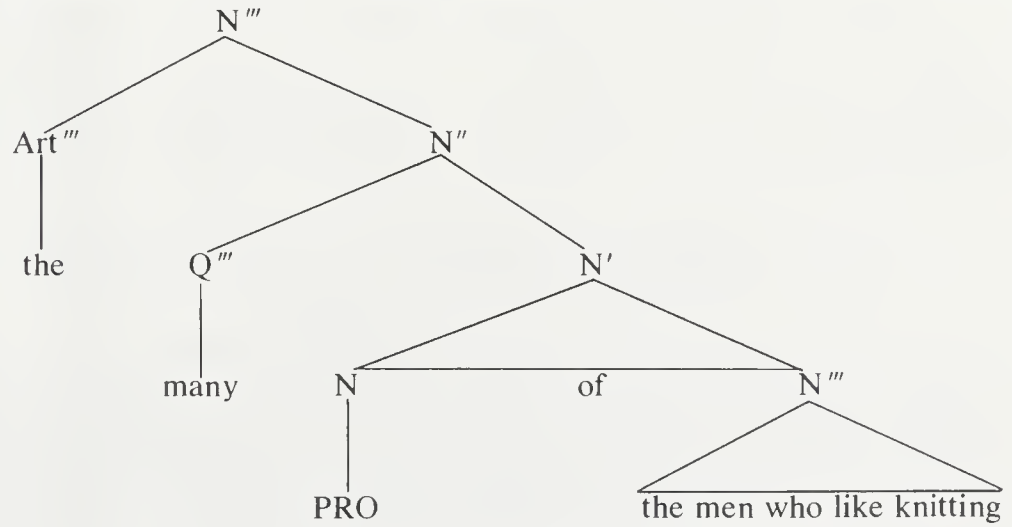
(7.35) a. many of the men who like knitting



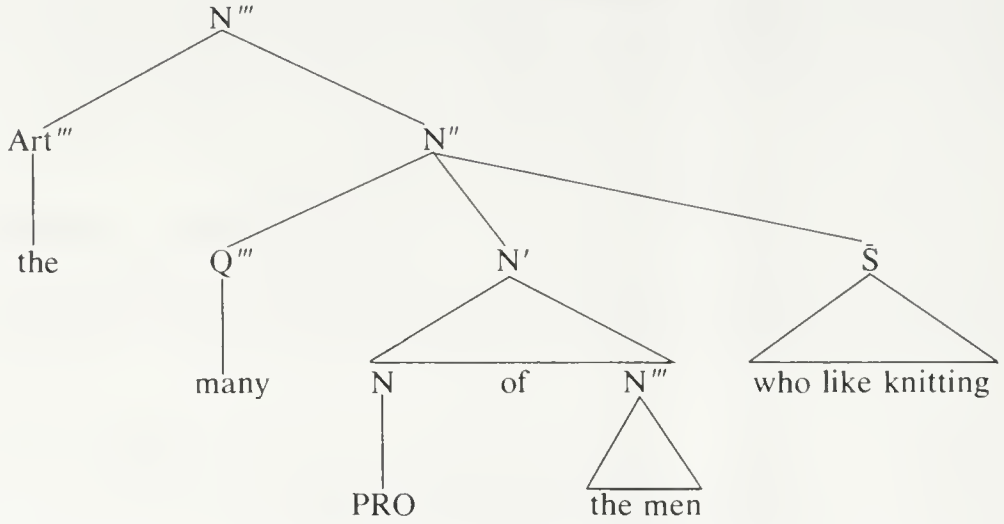
b. many of the men who like knitting



c. the many of the men who like knitting



d. the many of the men who like knitting



In (7.34) the relative clause modifies *many men*; there is no way to separate its modification into that of *many* vs. that of *men*. Such a prediction seems correct. *Many of the men who like knitting* is, however, ambiguous. Most typically it seems to have the structure (7.35a), with the relative attached to *the men*. But a phrase such as *many of them who like knitting*, in which the relative clause cannot be attached to the complement (**them who like knitting*), shows that structure (7.35b) is possible too. Further, if evidence from number agreement is at all reliable, (7.36a) must have structure (7.35a), and (7.36b) must have structure (7.35b).

(7.36) a. $\left\{\begin{matrix}\text{one}\\\text{each}\end{matrix}\right\}$ of the men who like knitting

b. $\left\{\begin{matrix}\text{one}\\\text{each}\end{matrix}\right\}$ of the men who likes knitting

Finally, of the pair (7.35c,d), only the latter is a correct structure for *the many of the men who like knitting*, since, as was shown in section 5.3, there is a constraint requiring a relative clause attached to the upper NP in this construction. This can be shown either by preposing, as in (7.37), or by number agreement in the relative clause, as in (7.38).

(7.37) a. Of the men, the many who like knitting are the happiest.
b. *Of the men who like knitting, the many are the happiest.

(7.38) the (only) one of the men who $\left\{\begin{matrix}\text{likes}\\\text{*like}\end{matrix}\right\}$ knitting

Of course, if a second relative clause is added, it may modify the lower NP. Compare the preposing and number agreement in (7.39) to (7.37)–(7.38) to see that one restrictive in (7.39) is attached to the upper NP, and one to the lower.

(7.39) a. Of the men who like bowling, the many who like knitting are the happiest.
b. the (only) one of the men who like bowling who likes knitting

We have shown, then, that the NP-complement theory of restrictives combined with the N′-complement theory of partitives provides the right structures and interpretations for constructions involving both quantifiers and relatives. Thus the evidence from quantification supports the NP-complement theory of relatives.

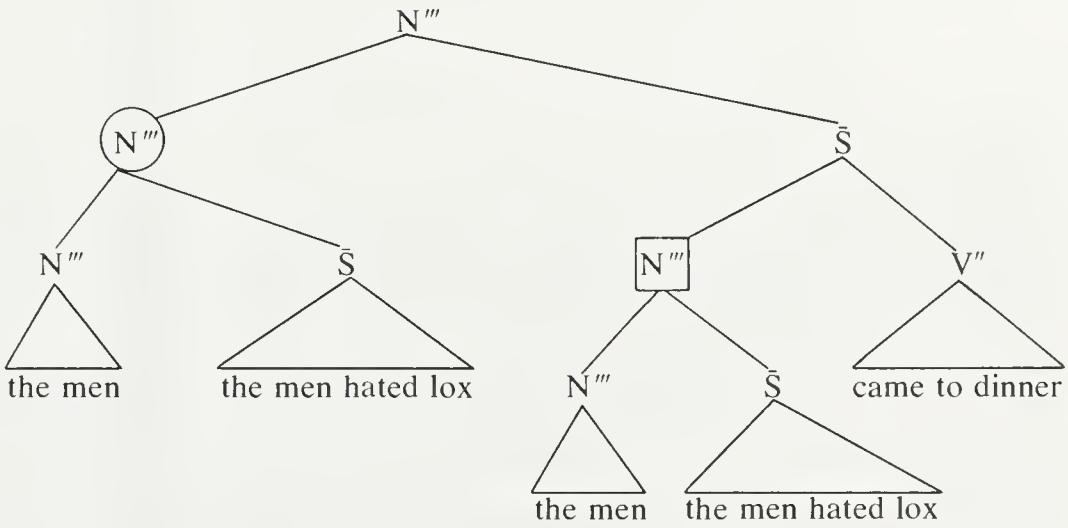
7.6. Stacking of Restrictive Relatives

Another kind of apparent counterevidence to the NP-complement theory concerns the phenomenon of "stacked" restrictives. To see what problems it raises, observe that (7.40a) and (7.40b) appear to differ in meaning.

(7.40) a. the men who hated lox who came to dinner
b. the men who came to dinner who hated lox

It is commonly claimed that the difference is due to the fact that the second relative clause modifies or restricts both the head and the first relative clause. Thus the antecedent of the first relative pronoun in (7.40a) is *the men*; of the second, *the men who hated lox*. The theory par excellence for expressing this difference is Ross's Chomsky-adjoined theory, since (7.40a), for example, can receive the deep structure (7.41).

(7.41)

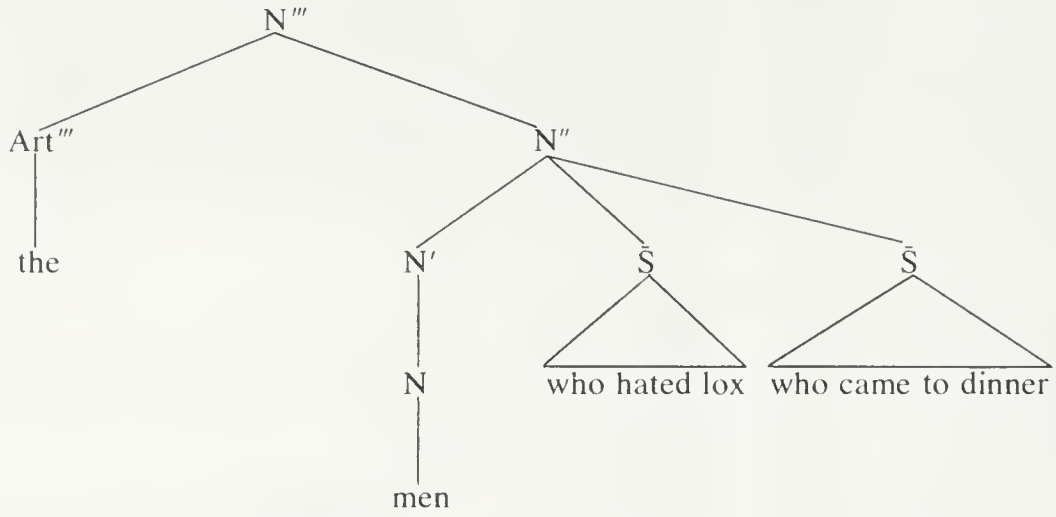


Within both the circled and boxed NPs, the subordinate clause can relativize, so each NP becomes *the men who hated lox*. Then the boxed NP is identical with the circled NP and can itself become a relative pronoun to yield the desired surface form (7.40a). Thus the intuitions about (7.40) are accounted for by "stacking" the relative clauses.

In the NP-complement theory, however, it appears difficult to account for the

"stacked" interpretation, since the most plausible structure with two restrictive relative clauses simply concatenates them:

(7.42)



The fact that the relative clauses are not stacked in (7.42) appears to be a serious drawback. We will show, however, that this is not a drawback but a virtue of the NP-complement theory.

Consider the kind of context in which the "stacked" interpretation emerges most clearly:

(7.43) a. The men who hated lox who came to dinner left early, but that wasn't true of $\left\{\begin{matrix}\text{the men who hated lox}\\ \text{those}\end{matrix}\right\}$ who arrived after 8:00.

b. ??The men who came to dinner who hated lox left early, but that wasn't true of the men who arrived after 8:00 who hated lox.

(7.43a) sets up a situation in which *the men who hated lox who came to dinner* contrasts with *the men who hated lox who arrived after 8:00*. According to the rules of focus and presupposition of Chomsky (1970b), Akmajian (1973), and Jackendoff (1972, chapter 6), the contrasting constituents must be foci and thus must receive prominent stress. The contrasting parts of the two subject NPs in (7.43a) are the final relative clauses, where highest stress normally appears. In (7.43b), though, the contrasting relative clauses are medial in the NP, and the contrast can be obtained felicitously only by placing a strong stress on the contrasting relatives and effacing stress on the final relatives:

(7.44) The men who came to DINNER who hated lox left early, but that wasn't true of the men who arrived after eight o'CLOCK who hated lox.

This suggests that the phenomenon of "stacking" is not to be accounted for in the syntax, but rather in the system of presupposition and focus. Two questions are raised by such a conjecture: first, are the rules of presupposition and focus capable of deriving interpretations which account for the intuitions of stacking? Second, if stacking has

nothing in particular to do with how relative clauses are attached, can other things produce stacked interpretations as well?

The rules for focus and presupposition derive (7.45a) as the interpretation of the first clause of (7.43a), (7.45b) for (7.43b), with normal stress, and (7.45c) for (7.44), with special stress.[7]

(7.45) a. Presupposition: some men who hated lox left early
Focus: who came to dinner
b. Presupposition: some men who came to dinner left early
Focus: who hated lox
c. Presupposition: some men who hated lox left early
Focus: who came to dinner

In (7.43a) and (7.44) the presuppositions can be matched in the two clauses, and the foci contrast properly. But in (7.43b) the presupposition of the second clause involves *men who arrived after 8:00* and the foci do not contrast, hence the observed infelicity.

Notice also that in (7.45a) the focused relative clause specifies members of the set of men who hated lox; in (7.45b) the focused relative clause specifies members of the set of men who came to dinner. These are exactly the properties required to satisfy the intuition of a "stacked" interpretation, and they are achieved without any syntactic counterpart of the semantic stacking.

"Stacked" interpretations can also be produced with pairs of PPs in the NP complement. As with relatives, the stacking can be reversed by changing stress.

(7.46) a. The weather in Kansas in 1875 was better than $\left\{\begin{matrix}\text{that}\\ \text{the weather in Kansas}\end{matrix}\right\}$ in 1936.
b. The weather in Kansas in 1875 was better than the weather in Australia in 1875.

(7.46a) is felicitous only with stress on *1875* and *1936*; (7.46b) is felicitous only with stress on *Kansas* and *Australia*. It is clear that the paradigm is parallel to (7.43)–(7.44). Furthermore, the same pattern appears in the case of a PP followed by a restrictive relative:

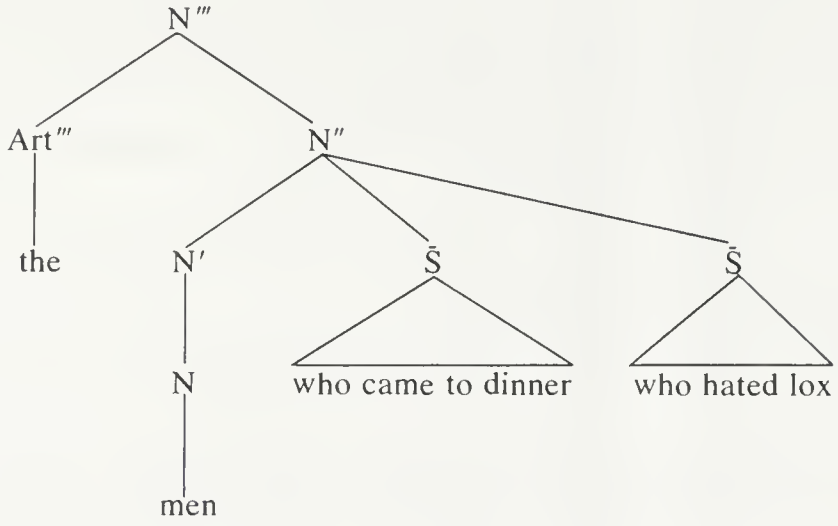
(7.47) a. The man with a scar who came to dinner hated $\left\{\begin{matrix}\text{the one}\\ \text{the man with a scar}\end{matrix}\right\}$ who cooked lunch.
b. The man with a scar who came to dinner was fond of the man with green eyes who came to dinner.

Again, (7.47a) must be read with stress on *dinner* and *lunch*, and (7.47b) with stress on *scar* and *eyes*.

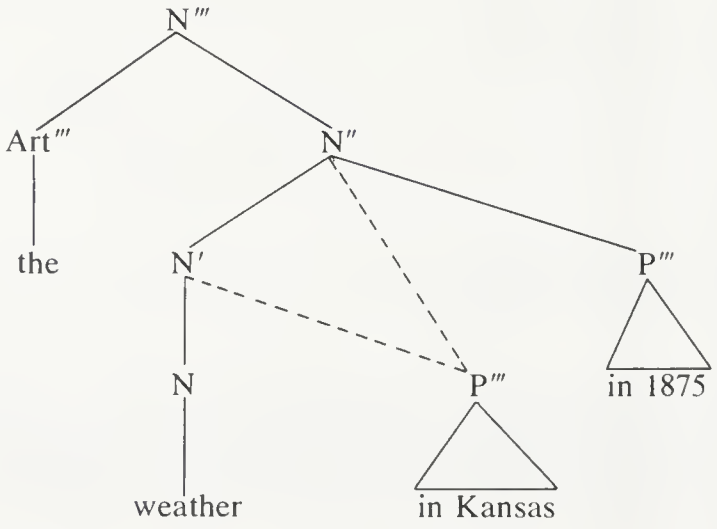
[7] I am oversimplifying here, assuming only one of a number of possible foci linked to the given stress pattern, namely the relative clause. The other possibilities are irrelevant to the context of (7.43) and (7.44). See the references cited above for details of the treatment of focus and presupposition assumed here.

Since the syntactic subordination has now been shown unnecessary, stacking is no longer a ground for preferring the Chomsky-adjoined analysis of restrictive relatives over the NP-complement theory. The latter assigns the relevant NPs in (7.43)–(7.47) the following structures:

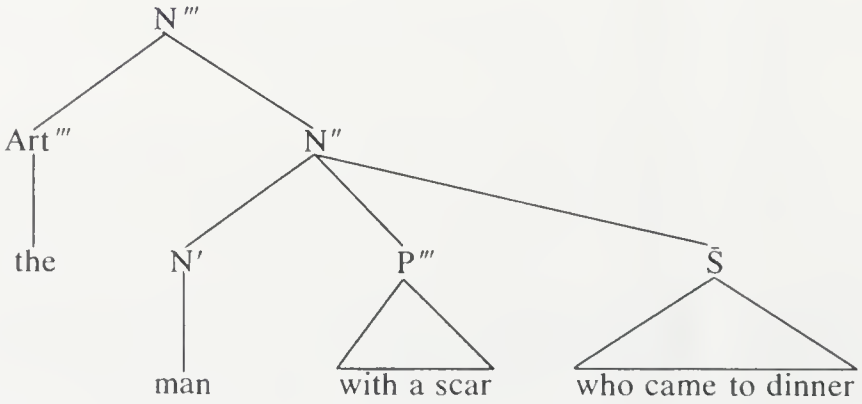
(7.48) a.



b.



c.



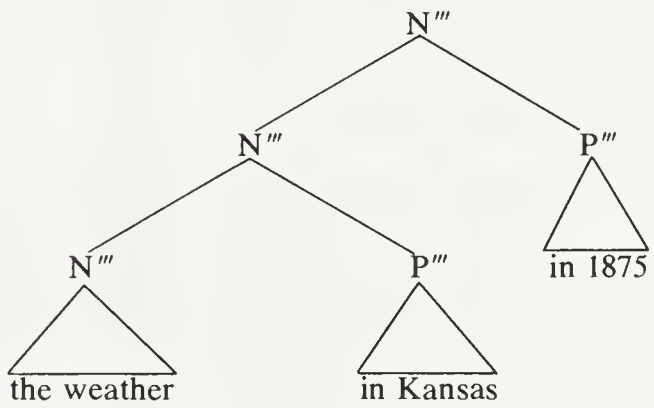
This analysis treats (7.43)–(7.47) as altogether comparable, since both PPs and relatives can be attached to N″.

Only one simple change must be made in the phrase structure rule for N″ to account for multiple relatives: the S̄ at the end must be replaced with S̄*.

(7.49) $N'' \rightarrow \ldots - N' - (P''')^* - (\bar{S})^*$

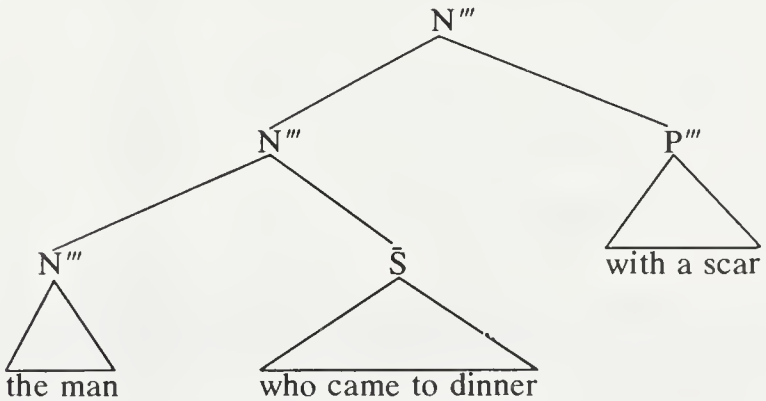
A moment's consideration shows in fact that our analysis of stacking in terms of focus vitiates the alleged advantage of the Chomsky-adjoined theory of restrictives—its claim that syntactic stacking reflects semantic stacking. A sentence with reversed stacking such as (7.44) will require a deep structure in which the order of the relative clauses is reversed, and an extraposition rule will be necessary to create the proper order. For parallelism, a sentence with stacked PPs such as (7.46a) will require a structure such as (7.50), for which there is little evidence, and the reversed stacking in (7.46b) creates the same problems as (7.44).

(7.50)



(7.47a) will not be a problem, but the reversed stacking in (7.47b) requires the deep structure (7.51), which appears at the surface only with great awkwardness, unless the S is extraposed.

(7.51)



(Note ??*the man who came to dinner with a scar*)[8]

[8] I must emphasize again that PPs like *with a scar* cannot in general be derived by reducing relative clauses.

Furthermore, the rules dealing with the interpretation of focus and presupposition on the basis of surface stress are independently necessary, and make it possible to produce "stacked" interpretations with normal stress patterns and no syntactic subordination of the "inside" relative clause.

The syntactic stacking of the Chomsky-adjoined theory is thus not only superfluous but even disadvantageous. Hence the phenomenon of stacking is not only not counterevidence to the NP-complement theory, but also provides important evidence in its favor. The generalization underlying this result is the same as in section 7.4: restrictive relatives work exactly like all other restrictive modifiers.

7.7. Conjoined Heads

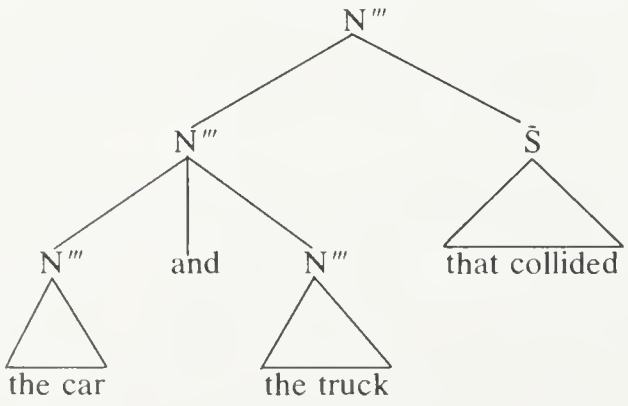
The fourth and final class of apparent counterexamples to the NP-complement theory involves data first pointed out (to my knowledge) by Ross; the argument in the form given here is due to Vergnaud (1974). Consider the NPs in (7.52).

(7.52) the car and the truck that collided
the boy and the girl who met in Vienna

These NPs cannot be derived by conjunction reduction, since the sources would have to be the ungrammatical **the car that collided and the truck that collided, *the boy who met in Vienna and the girl who met in Vienna.* The verbs of these relative clauses require plural or conjoined subjects; hence the relative pronoun's antecedent cannot be either of the nouns alone.[9]

Ross's analysis deals with such examples straightforwardly:[10]

(7.53)

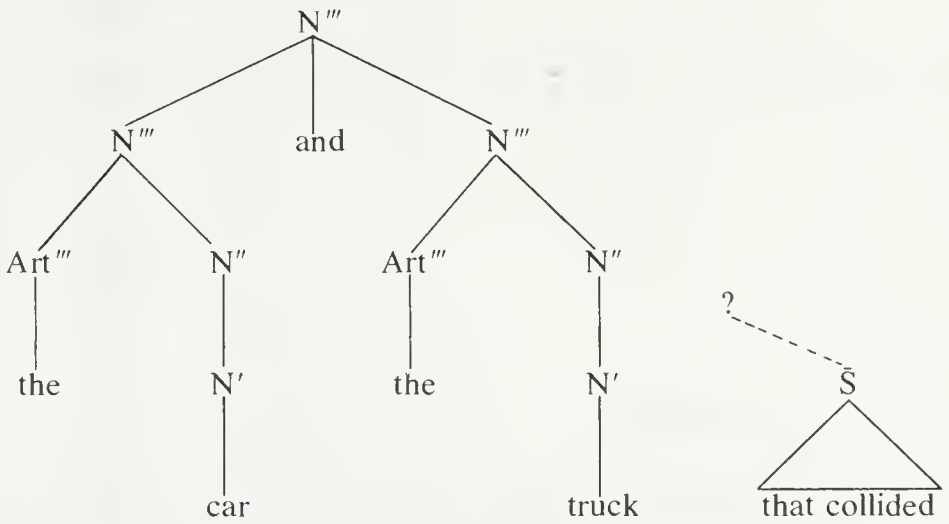


[9] Dougherty's (1970–1971) discussion of conjunction is, I believe, sufficient to eliminate such distant sources as *the car that collided with the truck and the truck that collided with the car.* A source once suggested by Ross (in lectures at MIT), *the car wh- [car and truck] collided and the truck wh- [car and truck] collided,* is semantically incomprehensible.

[10] Vergnaud is actually proposing a slightly more sophisticated theory than Ross's; however, with respect to this particular argument the two theories are equivalent. I will therefore couch the argument in terms of Ross's theory, which will probably be more familiar to the reader.

In this structure, the relative clause is sister of a constituent that dominates both NPs, and the antecedent of the relative pronoun can be the conjoined NP. On the other hand, the NP-complement theory provides no convincing structure, since the Art‴ is superior to the relative clause:

(7.54)



Since only a theory like Ross's can come even remotely close to a suitable structure for (7.52), and it does so with great ease, Vergnaud considers this construction overwhelming evidence for such a theory.

However, if Vergnaud's argument is correct, then PP restrictive modifiers must be outside the NP too, since they occur in the same sorts of constructions:

(7.55) the boy and the girl { with the same birthday / with mutual interests / with different-colored eyes / with a common background }

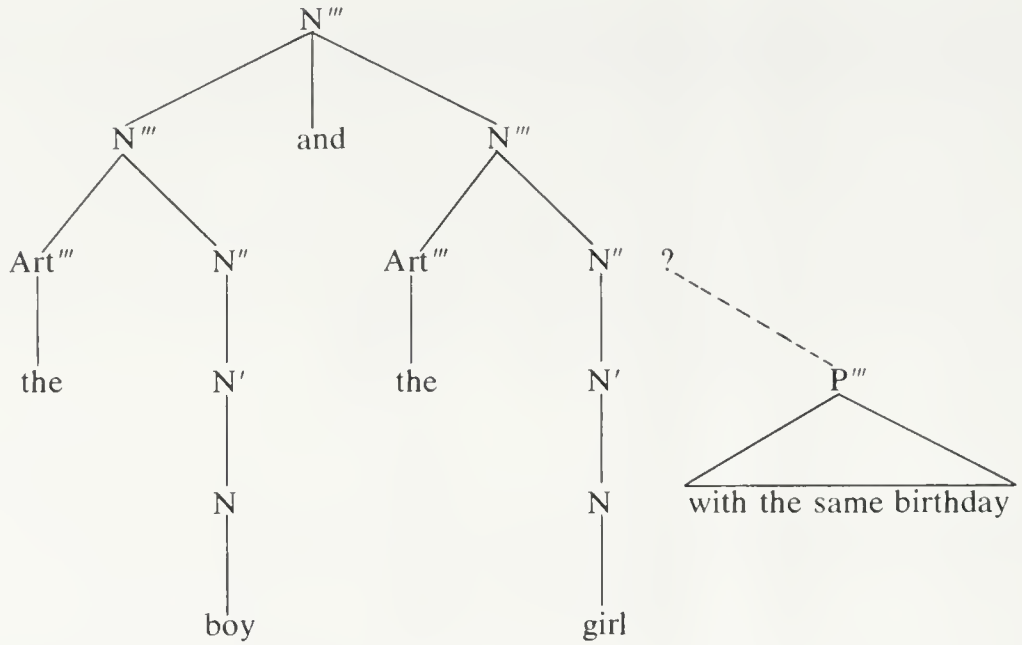
These examples can be derived neither by Relative Clause Reduction nor by Conjunction Reduction, since both (7.56a) and (7.56b) are ungrammatical.

(7.56) a. *the boy and the girl who were with the same birthday (Relative Clause Reduction source)

b. *the boy with the same birthday and the girl with the same birthday (Conjunction Reduction source)

If these PPs are in the NP complement, (7.55) raises the same problem in any theory as (7.52) does in the NP-complement theory of relatives: where is the PP attached?

(7.57)



Lest one should immediately yield to the temptation to attach these PPs outside the NP, we observe that N′ complements create this problem too.

(7.58) three students and two teachers of the same language
three members and two vice-chairmen of interlocking committees

Vergnaud's argument forces us to conclude that the proper bracketing is [[*three students*] *of French*] and [[*three members*] *of the committee*], not the otherwise highly motivated [*three* [*students of French*]] and [*three* [*members of the committee*]]. Things are beginning to look fishy.

To make life even worse, the $\bar{X}$ Convention as usual leads us to discover similar examples in sentences:

(7.59) a. John avoided and Bill ignored $\left\{\begin{matrix}\text{similar issues}\\ \text{the same man}\\ \text{men with the same birthday}\end{matrix}\right\}$.

b. John whistled and Mary hummed $\left\{\begin{matrix}\text{the same tune}\\ \text{at equal volumes}\\ \text{together}\end{matrix}\right\}$.

None of these sentences can arise from Conjunction Reduction or Right Node Raising, since the presumed sources would either have the wrong reading or be ungrammatical:

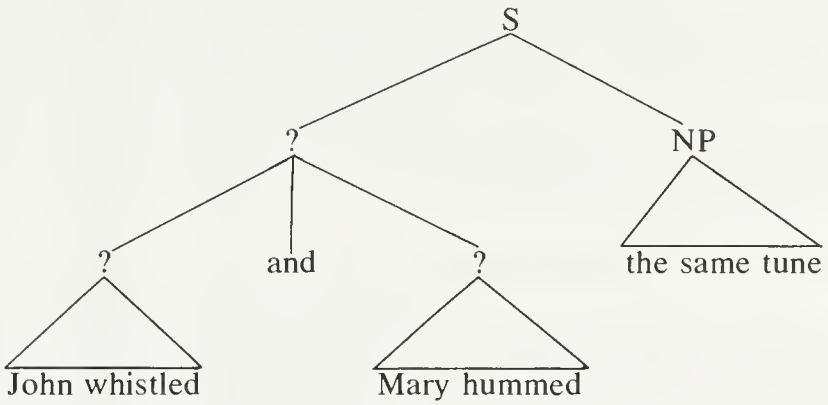
(7.60) a. John avoided $\left\{\begin{matrix}\text{similar issues}\\ \text{the same man}\\ \text{men with the same birthday}\end{matrix}\right\}$ and Bill ignored

$\left\{\begin{matrix}\text{similar issues}\\ \text{the same man}\\ \text{men with the same birthday}\end{matrix}\right\}$.

b. John whistled { the same tune / *at equal volumes / *together } and Mary hummed { the same tune / *at equal volumes / *together }.

By Vergnaud's argument, then, the phrases in braces must be right sisters of conjoined subject–verb pairs:

(7.61)



But this flies in the face of the presumably well-established fact that the phrases in braces are direct objects or other V′ and V″ complements.

It is really unfair to saddle Vergnaud with this argument, since such arguments based on conjunction have been considered strong evidence for structure throughout the history of transformational grammar. We have seen, however, that this type of argument leads ultimately to conclusions which contradict the most fundamental and best accepted notions of phrase structure.

The question is, which set of arguments should be given up? If we wish to retain the arguments based on conjunction, we give up all distinctions between various types of complements, since they must all reduce to the Chomsky-adjoined variety. All the arguments of chapter 4 distinguishing complement types must go by the wayside and all evidence for internal structure of NPs and Ss must be accounted for in some other way.

I prefer to believe that the error lies rather in the theory of conjunction, always a troubled area in transformational grammar. There are to my knowledge no remotely coherent formalizations of Right Node Raising; Gapping is ever problematical; and even Dougherty's (1970–1971) highly restricted version of Conjunction Reduction comes to grief in examples like these:

(7.62) The same man { got drunk and was arrested by the cops / is rarely easy to please and eager to please / praised you and seemed to hate you }.

Dougherty invokes a Conjunction Reduction transformation specifically in order to deal with conjoined derived verb phrases. But the underlying source he needs is ungram-

matical here:

(7.63) *The same man got drunk and the cops arrested the same man.
*It is rarely easy to please the same man and the same man is rarely eager to please.
*The same man praised you and it seemed [(for) the same man to hate you].

If I am correct, and the anomaly lies in the theory of conjunction, we can retain all the usual arguments for phrase structure; only the arguments from conjunction have to be rejected. I have no positive proposals to offer on the analysis of (7.52)–(7.62). But this choice defuses Vergnaud's otherwise powerful argument for the Chomsky-adjoined theory of relatives; thus the NP-complement theory of relatives, like the theory that there are direct objects in sentences, can be maintained. What is really needed is a thoroughly revised theory of conjunction.

7.8. Interpretation of the Restrictive Relative Pronoun

Having dealt with potential difficulties for the NP-complement theory of restrictives in the areas of determiner restrictions, quantification, stacking, and conjunction, we now turn to the problem of interpreting the restrictive relative pronoun. As in the case of appositive relative pronouns, we assume an interpretive theory, in which the relative pronoun is inserted into underlying structure and *wh*-preposed; a semantic rule applies to the surface structure to establish the interpretation of the pronoun. The problem is to find out exactly what this interpretive rule does. If it is properly defined, it can account for otherwise puzzling facts about relative clauses.

Section 7.3 showed that the appropriate interpretive rule for the appositive relative pronoun was one that established anaphora between the relative pronoun and the head. A similar rule of anaphora could be developed for restrictive relative pronouns, but there is reason to consider a different interpretation. To see this, let us look a bit more closely at the semantics of NPs.

The semantic functions of various parts of an NP down to the N″ level are of at least four types: demonstrative or deictic, classificatory, measuring, and commenting. The demonstrative function is taken by articles; they show whether the individual in question is old or new in the discourse, unique, pragmatically designated, generic, and so forth. Elements with the classificatory function place conditions on what sort of object the individual in question is; the N′ and the restrictive modifiers have this function. Measuring and commenting functions are taken by numerals (and perhaps N″ quantifiers) and by appositives, respectively.

Traditional logic has not had much to say about the latter two functions, but it makes a sharp distinction between the former two. All the classificatory functions have been represented as predicates; those demonstrative functions that have been treated at all emerge as operators such as quantifiers or the definite description operator. Thus,

for example, (7.64) might be represented as (7.65) in one version of logical notation, restricted quantification. For illustrative purposes, I will use the rather free translation characteristic of traditional logic.

(7.64) a. Some brown beaver died.
b. The picture of Bill from London fell down.
c. Every duck in Australia is pink.

(7.65) a. $\exists x_{x \text{ is brown, } x \text{ is a beaver}}$ (die(x))
b. fall down ($\iota x_{x \text{ is a picture of Bill, } x \text{ is from London}}$)
c. $\forall x_{x \text{ is a duck, } x \text{ is from Australia}}$ (pink(x))

In these translations, the article has become an operator; each classificatory constituent has been supplied with an "x is" and placed in the restriction on the variable controlled by the operator.

Restrictive relative clauses require a slightly different treatment. In this notation, the most appropriate translation of *The man who Bill saw died* is (7.66).

(7.66) die($\iota x_{x \text{ is a man, Bill saw } x}$)

Here the relative clause has become a restriction on the variable. The translation rules have not added the fixed material "x is", as in (7.64); rather, the relative pronoun has itself been replaced by the bound variable x. In other words, ordinary classificatory elements can be made to restrict the operator by adding fixed external material, but restrictive relative clauses must be tampered with internally, in the position corresponding to the relative pronoun. The tampering involves replacing the relative pronoun by the bound variable controlled by the operator.

I suggest that something parallel to these logical translations is going on in the natural language semantics of restrictive relative clauses. The relative pronoun is the syntactic marker of the position which must be bound by the demonstrative operator of the head. Since the syntactic correlate of the demonstrative is the Article position, the rule for interpreting the relative pronoun must establish a "binding" relationship between the relative pronoun and the article:

(7.67) *Restrictive* Wh-*Interpretation*
$\begin{bmatrix} X \\ +wh \end{bmatrix}'''$ is bound to Art$'''$, in the configuration
Art$'''$[$_{N''}$. . . [$_{\bar{S}}$[$_{Comp}$. . . X$'''$ . . .] S] . . .]

The exact characterization of the "binding" relationship depends on the ultimate form of semantic notation chosen; in the rough logical translation rules above, "is bound to" would mean "is replaced by the variable of".

The crucial mention of the Art$'''$ in (7.67) may be used to account for one of the differences between appositives and restrictives. Since N$'''$ is the only category which dominates Art$'''$, restrictives can be related only to NPs. But since appositive relative

pronouns are interpreted as anaphoric to the entire head, the internal structure of the head is not crucial, and appositives may thus be generalized to all major syntactic categories.

The dependence on Art‴ in (7.67) plays a role in accounting for certain ungrammaticalities as well. Section 7.4 mentioned one such example:

(7.32) *John's book that you stole

John's is not an Art‴, so the relative pronoun cannot be bound; since the example does not receive a complete interpretation, it is anomalous. Next consider the ungrammaticality of (7.68a), and compare it to (7.68b,c).

(7.68) a. *the destruction of the city that you observed
b. the destruction that you observed
c. the destruction of the city, which you observed

Recall the discussion of section 4.7.4. There it was argued that when *the* is the only specifier possible in a derived nominal, it arises not through expansion of Art‴, but through insertion into an empty N‴. *The destruction of the city* is such a case, and *the destruction* is not, since we have, for example, **much destruction of the city* but *much destruction*. Thus in (7.68a), *the* is dominated by N‴ and the structural description of (7.67) cannot be met, hence the phrase is ungrammatical; in (7.68b), *the* is dominated by Art‴, so the phrase is grammatical. Since the interpretation of appositive relative pronouns does not depend on the presence of an Art‴, (7.68c) too is grammatical, though its determiner is dominated by N‴.

This analysis of the relative pronoun eliminates the usual puzzles (cf. for example Kuroda (1969)) about whether relative pronouns are definite or indefinite: they are neither. In fact, the issue hardly makes sense in terms of the interpretation given by (7.67), since the bound variable is not referential at all, but simply has the effect of converting the proposition represented by the relative clause into a property. So, for example, *a man who you know* denotes an individual not completely identifiable from previous discourse (i.e. indefinite), who is male, adult, and known by the hearer. *The man who came to dinner* denotes an individual identifiable by this description plus previous discourse or general knowledge (i.e. definite), who is male and adult and who came to dinner. A relative clause may be "definitizing", i.e. render an otherwise indefinite NP definite, just in case it provides grounds for unique identification which were not present in the NP without the relative. Thus there is no transformation of the form that Perlmutter (1970) and others adopt, changing the article of an NP to *the* in the presence of a restrictive relative clause.

In cases where the demonstrative element is interpreted as picking out a sort or degree, it does so in the relative clause as well. For example, the object of *We drank the wine we always drink* denotes not a particular physical body of wine (barring

regurgitation), but a *sort* or *amount* of wine. If this interpretation is localized in the semantic markers of the determiner, it will correctly apply both to the head and to the relative clause.

In an effort to give a plausible underlying determiner to the relative pronoun in a generic sentence like (7.69a), Stockwell, Schachter, and Partee (1973) are forced to propose a source like (7.69b), with an attendant lowering transformation to derive the surface form.

(7.69) a. A beaver that eats worms builds leaky dams.
b. If a beaver eats worms, it builds leaky dams.

In the present analysis such a strategem is unnecessary. The generic character of the head NP is determined by the indefinite article *a*, and this binds both the head N *beaver* and the relative pronoun. Thus the conditional character of the relative clause in (7.69a) is a result not of some unusual underlying form, but of the binding by (7.67) of the relative pronoun to a generic determiner.

Similarly for NPs with quantifiers in the Art‴ such as *every man who owns a Cadillac*. There is no need to resort to an underlying form such as *if a man owns a Cadillac, he . . .* in order to capture the semantic connection between the relative clause and the main clause. Rather, binding takes place in the usual fashion, as shown in (7.65a,c), and the quantifier automatically affects the relative clause.

A more subtle analysis than we have given here is required to account for the fact that *any* can show up in quantified relative constructions such as *every man who has any sense*. However, the failure of the present analysis to predict this is not serious, since these peculiar *any*s turn up not only in relative clauses but also in restrictive PPs, for example in *every man with any sense*. Thus the possibility of *any* here is not to be explained in the semantics of restrictive relative clauses, but in a more general account of restrictive modification, which we have not attempted to provide.

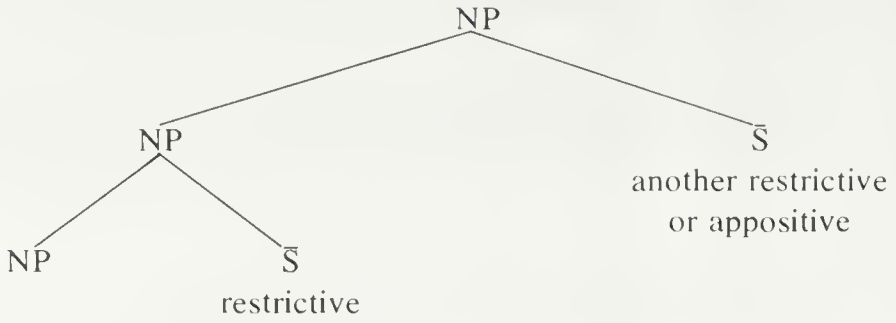
This completes our account of restrictive clauses. We have shown that the NP-complement theory of restrictive relatives is capable of dealing with a wide range of apparent counterevidence, and that in fact it permits a more general account of restrictive modification than either the determiner source or the Chomsky-adjoined source. Moreover, it requires neither an extraposition transformation, like the determiner theory, nor a weakening of the fundamental base rule schema, like the Chomsky-adjoined theory. This last consideration is most important to the general goals of this study, and it is the reason we have devoted so much space to the problems of relative clauses.

7.9. Appendix: Against the Conjoined Source for Appositives

In the Chomsky-adjoined theory of restrictives, restrictives are attached as high in the NP as they can possibly be. Thus there is no remaining position in NP to attach an

appositive that distinguishes it from a "stacked" restrictive:

(7.70)



Thus there can be no surface structural significance associated with the intonation break characteristic of appositives, already a serious problem.

To provide a deep structure difference between restrictives and appositives, Ross (1967) proposes deriving appositives from underlying conjoined clauses. For example, (7.71a) is to be derived from something more or less like (7.71b).

(7.71) a. That man, who came to dinner, sneezed.
b. That man sneezed, and he came to dinner.

Thus appositives have a different underlying source from restrictives, and many of the differences between them can be captured in terms of the underlying forms or the derivations of the two kinds of clauses.

Ross himself (section 6.2.4.1) points out a serious difficulty with the conjunction source: appositives can occur within imperatives and questions, though the sentences (7.73a,b), putatively more like the underlying form, are ungrammatical.

(7.72) a. Go to Cincinnati, which is on the Ohio River.
b. Are we landing in Washington, which is on the Potomac?
(7.73) a. *Go to Cincinnati, and it is on the Ohio River.
*Cincinnati is on the Ohio River, and go there.
b. *Are we landing in Washington, and it is on the Potomac?
*Washington is on the Potomac, and are we landing there?

Ross therefore halfheartedly suggests that appositives be derived from pairs of independent sentences, in effect by a *Syntactic Structures*-style "generalized transformation". Thompson (1971) rightly objects to this. But the only alternative she offers, without argument, is that structures like (7.73) indeed do underlie (7.72), but if appositive formation does not take place, the conjunction is expediently deleted to form two independent sentences. In effect she circumvents the counterexamples by brute force, explaining nothing.

Were this not difficulty enough, the derivation of the surface form requires a transformation lowering the appositive into the NP. Even if lowering rules are countenanced (as they are not in most lexicalist theories), one should have serious

reservations about lowering appositives into such strong islands as conjoined NPs (7.74a) and complex NPs (7.74b).

(7.74) a. Haj, who likes beer, and Sandy, who hates it, couldn't agree on one restaurant.
b. Paul's book about the transgalactic lexicon, which is the hottest topic of the decade, is eagerly awaited.

In defense of the conjunction source, Ross points out a sentence type which he claims is an intermediate step in the derivation from normal conjunction to appositives.

(7.75) Bill, and I've known him for many years, isn't as smart as he makes himself out to be.

Here a conjoined sentence has been inserted after the subject, supposedly halfway to becoming the appositive *who I've known for many years*. Unfortunately for this argument, though, the presumed intermediate step ranges from dubiously grammatical to impossible when substituted into examples such as (7.72) and (7.74):

(7.76) a. *Go to Cincinnati, and it's on the Ohio River.
b. *Are we landing in Washington, and it's on the Potomac?
c. *Haj, and he likes beer, and Sandy, and she likes pizza, couldn't agree on one restaurant.
d. ?*Paul's book about the transgalactic lexicon, and it's the hottest topic of the decade, is eagerly awaited.

The transformation producing (7.75), in other words, has all the constraints expected of a rule combining two conjoined sentences; the rule producing appositives does not. Hence the construction in (7.75) cannot be an intermediate step on the way to appositives; in fact its constraints constitute evidence *against* Ross's theory of appositives.

We see then that Ross's theory still lacks a viable underlying form for appositives. Of course, one way to avoid the problems of the conjunction source is to generate appositives in their surface position, as part of NP. But since restrictives already occupy the highest possible position within NP, there is no room in NP for a higher base position for appositives. Thus the disposition of appositives in Ross's theory is left enigmatic.

8: Degree Clauses

8.1. Introduction

Section 2.2 mentioned a parallelism, pointed out by Bowers (1968b), between NPs with restrictive relatives and APs with degree clauses associated with the degree words *so, too, enough, as, more,* and *less*, such as those in (8.1).

(8.1) a. He's *so* tired *that he'll never get up in time.*
b. He's *too* tired *to get up.*
c. He's tired *enough to sleep 14 hours.*
d. He's *as* tired *as he's ever been.*
e. He's $\left\{\begin{matrix} more \\ less \end{matrix}\right\}$ tired *than you are.*

Bowers was the first (to my knowledge) to express this parallelism in terms of the $\bar{X}$ Convention. Selkirk (1970), Milner (1973), and Reed (1974) develop the parallelisms further. Postal (1974) has an extensive discussion of their semantic similarities, giving a nonlexicalist formalization.

If this parallelism is genuine, and if restrictive relatives are attached to N″ in deep and surface structure, as argued in chapter 7, the $\bar{X}$ Convention compels us to adopt a theory of degree clauses in which they too are attached to X″ in deep and surface structure. The present chapter compares such a view to the prevalent view that degree clauses are extraposed from degree phrases. It will be shown that the X″-complement theory is in fact quite plausible, though not yet entirely compelling. Thus the framework proposed here is at least consistent with the facts of degree clauses, even if it cannot yet be shown that it predicts them.

8.2. The Extraposition Theory and its Evidence for the $\bar{X}$ Convention

Degree clauses can appear at the surface in any of the major syntactic categories A‴, N‴, V‴, and P‴. Bowers points out that they always follow strictly subcategorized complements in surface structure. Furthermore, since they always precede appositives,

the most likely surface position for them is in the X″ complement. No rearrangement of the complements in (8.2)–(8.5) is possible.[1]

(8.2) In A‴:

a. he's [so [[afraid to lose $_{A'}$] that he's freaking out, $_{A''}$] which you'll never be $_{A'''}$]

b. he's [too [[proud that he's won $_{A'}$] to admit it wasn't such a big deal $_{A''}$]$_{A'''}$]

he's [too [[sick$_{A'}$] to admit it wasn't such a big deal,$_{A''}$] which you'll never be$_{A'''}$]

c. he's [[[afraid enough $\left\{\begin{matrix}\text{of Bill}\\ \text{that he'll lose}\end{matrix}\right\}_{A'}$] to freak out,$_{A''}$] which you'll never be$_{A'''}$]

d. that she was [$\left\{\begin{matrix}\text{more}\\ \text{less}\end{matrix}\right\}$ [[eager to win$_{A'}$] than Fred was,$_{A''}$] which she hadn't been expected to be,$_{A'''}$] surprised me.

(8.3) In N‴:

a. we heard [[so many [rumors that Dewey had won$_{N'}$] that we started believing them$_{N''}$]$_{N'''}$]

b. we heard [[too many [rumors that Dewey had won$_{N'}$] to question the papers$_{N''}$]$_{N'''}$]

c. we heard [[as many [rumors that Wallace had won$_{N'}$] as we could stomach,$_{N''}$] which wasn't many$_{N'''}$]

(8.4) In V‴:

a. [he [[gave so little to the fund$_{V'}$] that he looked stingy,$_{V''}$] which surprised no one$_{V'''}$]

b. [he didn't [[say enough to us about how he did it$_{V'}$] to justify himself,$_{V''}$] which was a shame$_{V'''}$]

c. [this county [has [elected $\left\{\begin{matrix}\text{more}\\ \text{fewer}\end{matrix}\right\}$ Democrats to the Senate$_{V'}$] than any other in Missouri,$_{V''}$] which is hard to explain$_{V'''}$]

(8.5) In P‴:

a. we waited [[long enough [after the earth stopped shaking$_{P'}$] for things to settle down$_{P''}$]$_{P'''}$]

b. [[as much [out of place$_{P'}$] as he is$_{P''}$]$_{P'''}$], you're worse

c. we pushed him [[farther [into the room$_{P'}$] than Harry had been,$_{P''}$] which still wasn't very far$_{P'''}$]

It is rather obvious that there is a connection between the degree words and the

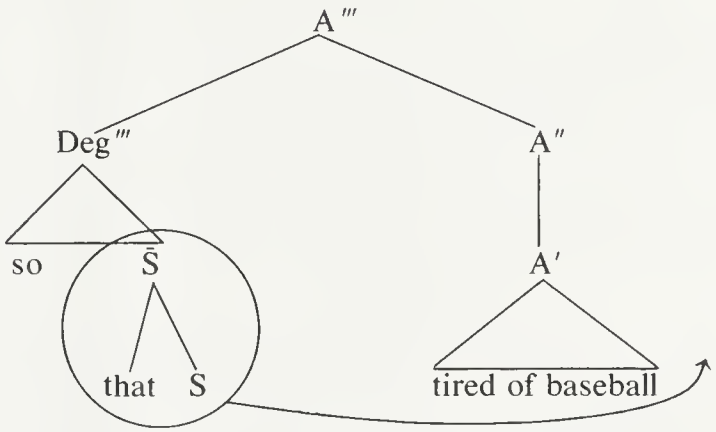
[1] The status of appositives in some of these examples is rather curious. The appositives in (8.3c) and (8.5c) seem to comment on the quantifier rather than the head, and no other type of appositive seems possible in N‴ or P‴ of this sort. I have no explanation.

associated degree clauses. If the degree words are omitted in (8.2)–(8.5), the sentences become ungrammatical unless the clauses in X″ are omitted too. Furthermore, each degree word governs characteristic complementizers:[2]

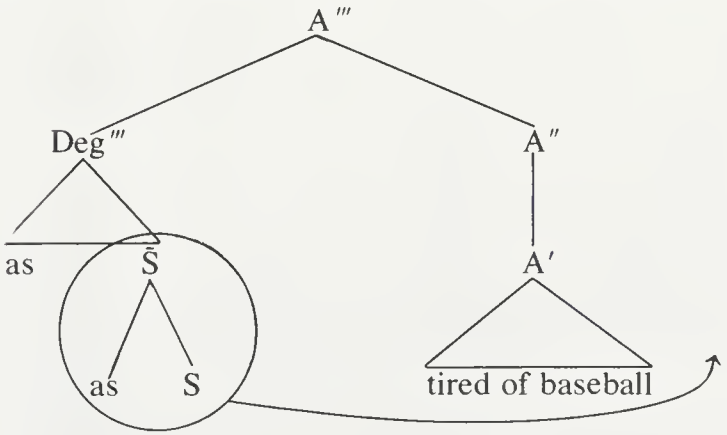
(8.6) so . . . that S
too . . . for-to S
enough . . . {that / for-to} S
as . . . as S
{more / less} . . . than S

This strongly suggests that degree clauses are strictly subcategorized complements of Deg, despite their surface position. Bowers (1968b) and Selkirk (1970) take this view, generating deep structures like (8.7) and extraposing the degree clause to its surface position.

(8.7) a.

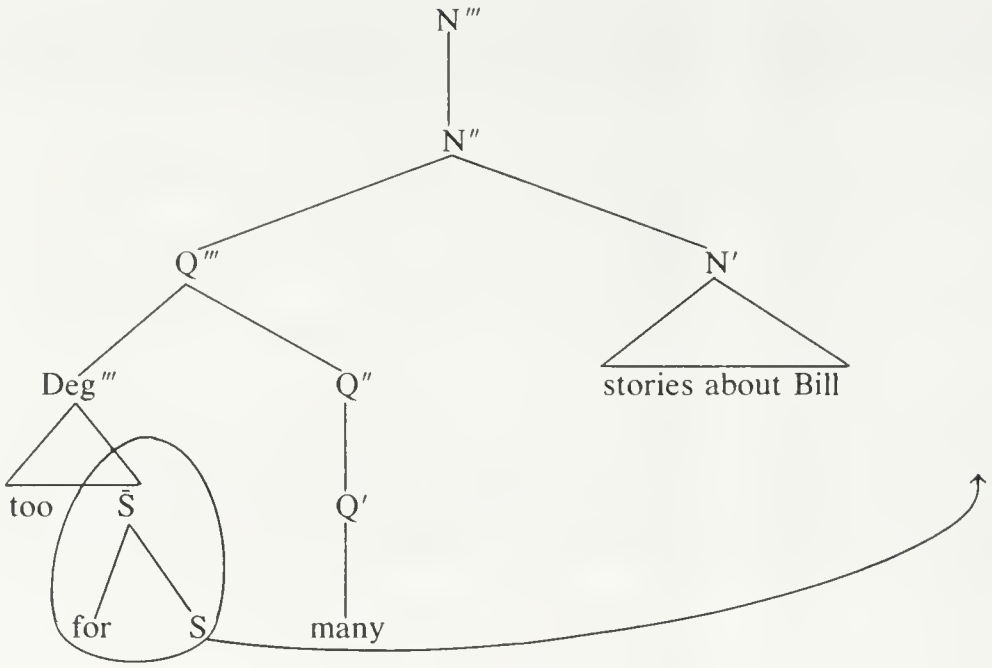


b.

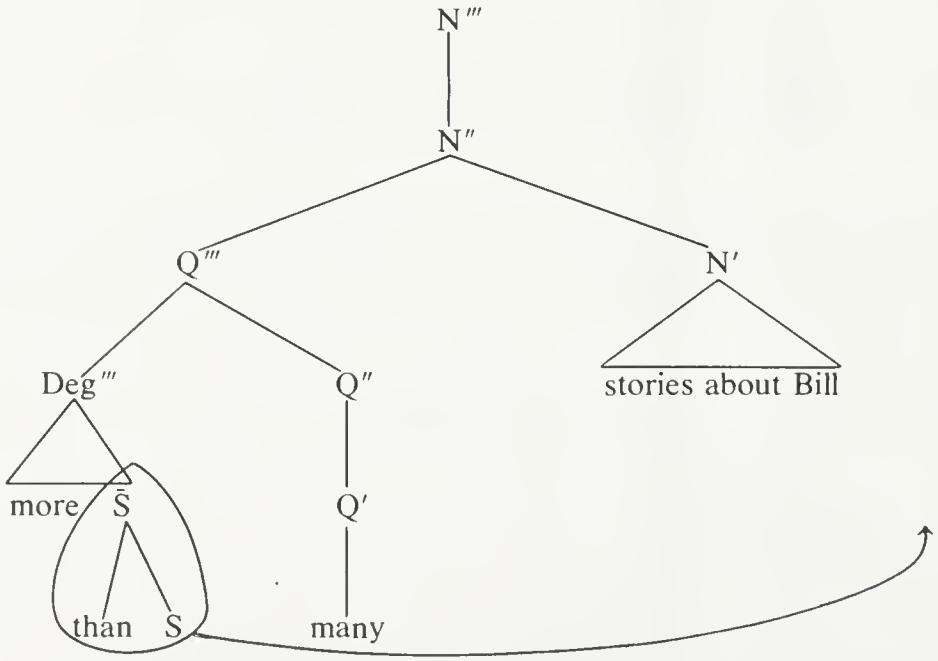


[2] For the complementizer status of *as* and *than*, see Bresnan (1970; 1972; 1974).

c.

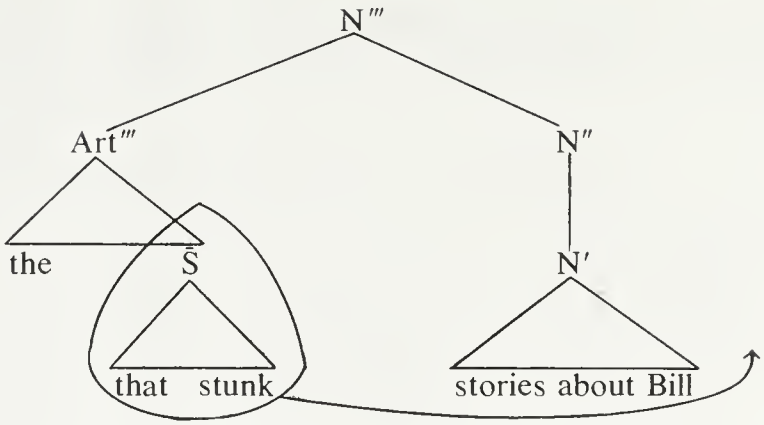


d.



The similarity of (8.7) to the determiner theory of restrictive relatives is evident:

(8.8)

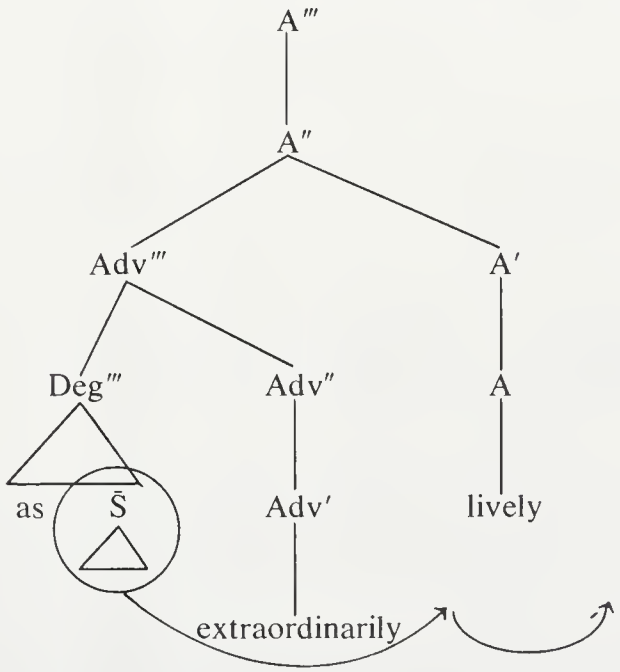


As Bowers and Selkirk point out, the extraposition operations in (8.7a,b) and (8.8) are structurally parallel; only by use of the $\bar{X}$ Convention can this parallelism be expressed. Furthermore, the deep structures are parallel, in that Deg and Art are both X‴ specifiers; in the present theory of syntactic features, they differ by the same feature, ±*Subj*, as their respective heads, A and N.

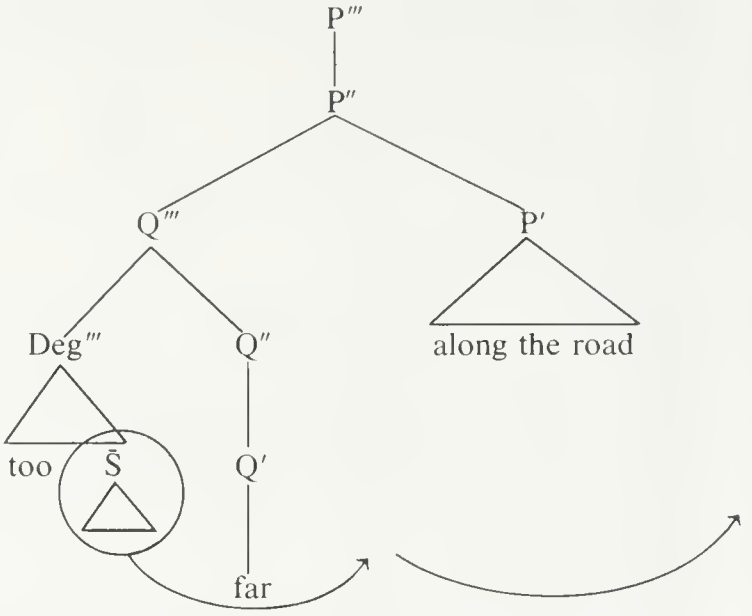
(8.7c,d) are more complex, but essentially the same as (8.7a,b), particularly if one assumes the $\bar{S}$ to be extraposed in two stages, first to the end of Q‴, then to the end of N‴. Under this assumption, the extraposition transformation used in (8.7a,b) and (8.8) needs only to be generalized to Q‴ via the $\bar{X}$ Convention; the double extraposition will be carried out cyclically, and no additional rules are necessary.

Similar double extrapositions take place in (8.9), where the categories dominating Deg‴ are different.

(8.9) a. as extraordinarily lively as he used to be

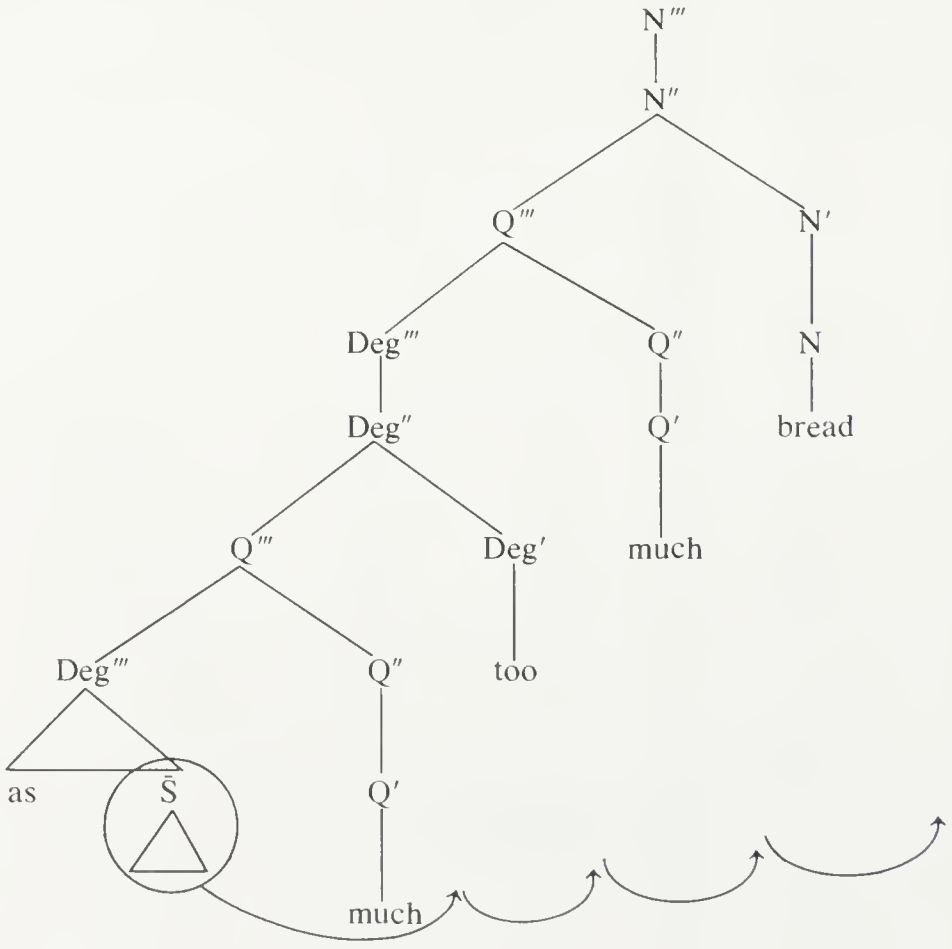


b. too far along the road for us to see him



With more complex degree specifiers, the number of extrapositions increases:

(8.10) as much too much bread as I could stand



It is clear that the rule accounting for surface position of degree clauses will have to make liberal use of the $\bar{X}$ Convention, whether or not it generalizes with the rules for restrictive relatives.

8.3. Arguments against Extraposition of Degree Clauses

In chapter 7 we rejected the view that restrictive relatives are generated under Art‴, in favor of generating them under N″. Does this mean that they are in fact not parallel in deep structure to degree clauses? Or should we seek arguments similar to those of section 7.4, to show that degree clauses too should be generated in their surface position under X″, restoring the parallelism?

Reed (1974) gives arguments relevant to this question. First, she points out that the expression *the same* governs a degree clause with the complementizer *as*, in which the permissible deletions are comparable to those in other comparatives.

(8.11) a. He looks the same as { you / you do / I think he used to (look) } . (in A‴)

b. He's reading the same book as { you / you are (reading) / I think you used to } . (in N‴)

c. He spoke the same as { you / you do / I think you used to } . (in Adv‴)

But in N‴ it also occurs with the complementizer *that* and the deletion pattern characteristic of restrictive relative clauses:

(8.12) He's reading the same book that { you're reading / *you (are) / I think you used to read / ?I think you used to } .

Thus *the same* must be allowed to govern either a degree clause or a relative clause; since the latter is permissible only in N‴, it will of course not crop up in other categories. This option is stated most simply if degree clauses and restrictive relatives are derived from the same underlying position, either in the specifier or in the X″ complement.[3]

One of the arguments adduced against the determiner theory of restrictives was that there are PPs that serve the same semantic function as restrictives. Reed observes that there are similar examples for some degree clauses:

[3] One possible drawback of this argument is that *which* cannot be substituted for *that* in (8.12), as would be expected in an ordinary relative clause. This may indicate that *the same* does not govern a genuine relative clause. On the other hand, it may indicate that we have been mistaken in regarding relative clauses with *that* and those with *which* as deriving from precisely the same source. If there were an underlying difference, *the same* could be constrained to govern only the former type.

(8.13) a. He's too fat {to wear shorts / for shorts}.

b. He's tall enough {to be a butler / for a butler}.

c. He's as sly {as you are / as a fox}.

d. He's taller {than you are / than six feet}.

(Notice that in the latter two, the PP case is unlikely to be a reduction from the questionable *?as sly as a fox is, *taller than six feet is.*[4] Rather, *as* and *than*, like *for,* must double as prepositions and complementizers.) Thus if degree clauses are underlying complements of Deg, extraposed to their surface position, these PPs must be too.

Reed next points out that the degree clauses and PPs of (8.13a,b) occur with a number of adverbial specifiers too, with similar meaning:

(8.14) a. He's quite young {to be a President / for a President}.

b. He's awfully young {to be in school / for school}.

c. This book is sufficiently boring {for us to use (it) / for a textbook}.

d. He's a pretty careful scholar {to have made that mistake / for a mistake like that}.

And in fact they can occur without any specifier:

(8.15) He's young {to be in school / for school}.

Since at least these particular degree clauses and PPs seem to have an existence independent of the degree words, Reed argues that they must be generated in deep structure to the right of the adjective, parallel to restrictive relatives, and that their

[4] Bresnan (1973) says the underlying clause in *taller than six feet* is *than six feet = x much,* invoking an underlying verb "=" which never appears on the surface and which has a reading suspiciously similar to the copula.

interaction with the degree words is purely semantic. The base rule for A″, therefore, must be of the form (8.16).

(8.16) $A'' \rightarrow \ldots A' - (\begin{Bmatrix} P''' \\ \bar{S} \end{Bmatrix})$

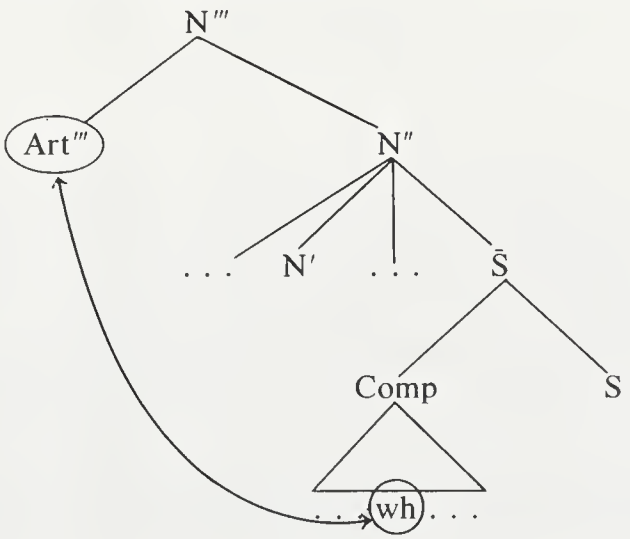
I have no further empirical arguments as to whether other degree clauses are generated in their surface position or whether they are extraposed to their surface position from the Deg‴ complement. However, note that the extraposition theory certainly goes against the spirit of the Extended Standard Theory, in that it claims an underlying structure for which there is no purely syntactic evidence. Only cooccurrence restrictions suggest the relationship of degree clauses to Deg‴, and furthermore there are no well-formed surface structures which contain a true complement of Deg. If cooccurrence restrictions could be stated as a function of the surface configuration, the grammar would be spared both an otherwise unmotivated phrase structure rule and the transformation needed to eliminate from the surface that phrase structure rule's effects.

We therefore will explore the alternative theory that degree clauses are generated directly in their surface position in X″, attempting to state an interpretive rule that accounts for the semantic dependency between degree clauses and degree phrases.

8.4. A Generalization in the X″-Complement Theory

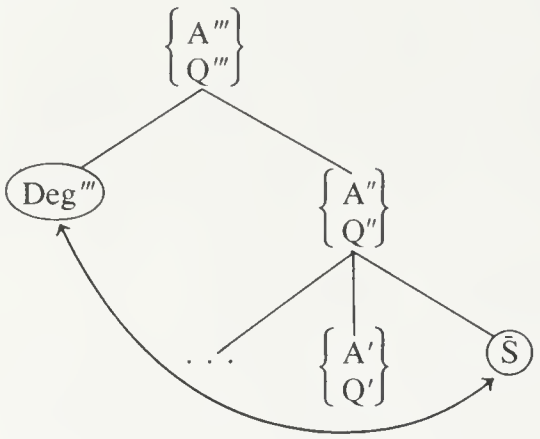
We first observe that the X″-complement theory of degree clauses restores the parallelism between degree clauses and restrictive relatives pointed out by Bowers, Selkirk, and others. Not only is there a structural parallel, since both types of clauses are generated as X″ complements, but there is a parallelism in their interpretation rules as well. Recall that the rule for interpreting restrictive relative pronouns developed in section 7.8 involves establishing a binding relation between the Article and an element of a clause in the following configuration:

(8.17)



Similarly, the simplest cases of degree clauses, such as (8.7a,b), are interpreted by establishing a semantic relationship between an X‴ specifier and part or all of a clause in the X″ complement:

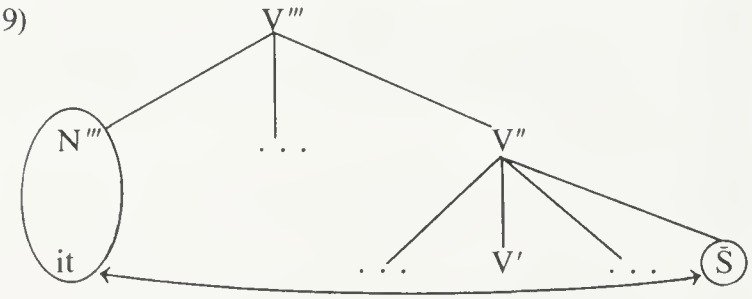
(8.18)



More complex cases have a less direct syntactic relation between the Deg‴ and the S̄; it is interesting, however, that at least the simplest cases so closely parallel restrictive relatives.

As some indication that this parallelism is not just an accident, recall the discussion in section 4.8 of sentences with "subject complements", for example *It bothers me that Tom isn't here*. It was argued that such a complement clause is a V″ complement, interpreted by establishing an anaphoric relation between it and the dummy subject *it*. The configuration is precisely parallel to (8.17) and (8.18):

(8.19)



Thus it seems possible to claim that there is a general rule of anaphora between X‴ specifiers and X″ complement clauses, with three special cases. If the X‴ specifier is an NP, the X″ complement is interpreted as a subject clause; if the X‴ specifier is an Art‴, the X″ complement is interpreted as a restrictive relative clause; if the X‴ specifier is a Deg‴, the X″ complement is interpreted as a degree clause. Notice that this generalization is more or less exhaustive, since N‴, Art‴, and Deg‴ are the only major syntactic categories other than M‴ that occur in any X‴ specifier.[5] Furthermore, both constitu-

[5] Since sentence adverbs, parentheticals, etc., are interpreted appositively, they are grouped semantically with complements, not with specifiers, even when they precede the head.

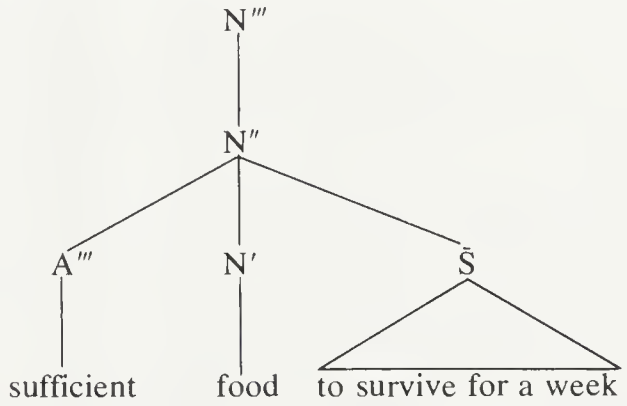
ents that take N‴ in the X‴ specifier, namely V‴ and N‴, take subject clauses; the only constituent with an Art‴, namely N‴, is the only one with restrictive relatives; all constituents that take Deg‴, namely A‴, Adv‴, and Q‴, also allow degree clauses. Thus an interpretive rule for degree clauses generated in their surface position appears to fit into a general cross-categorial class of anaphoric rules for which there is some independent motivation. Furthermore, a grammar incorporating such a rule would express the observed similarities between restrictive relatives and degree clauses, while maintaining the NP-complement theory of relatives argued for in chapter 7.

8.5. The Location of the Binding Degree Word

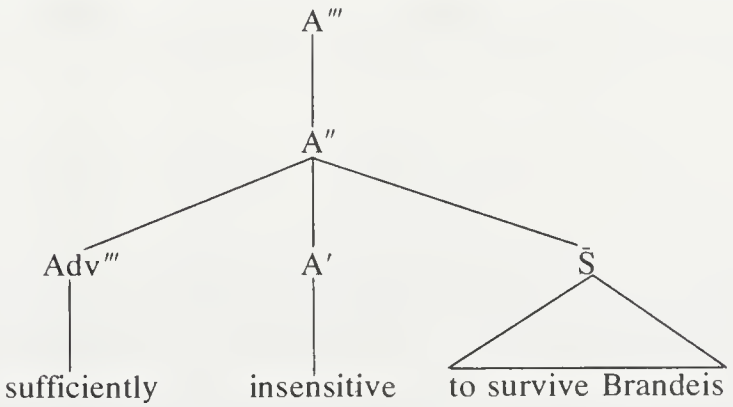
The configuration (8.18) is, however, only the simplest in which degree clauses can be interpreted. This section will deal with more complex situations.

Besides cases like (8.18) in which the degree clause is bound to the X‴ specifier of the immediately dominating phrase, it can be bound to an X″ specifier:

(8.20) a. sufficient food to survive for a week

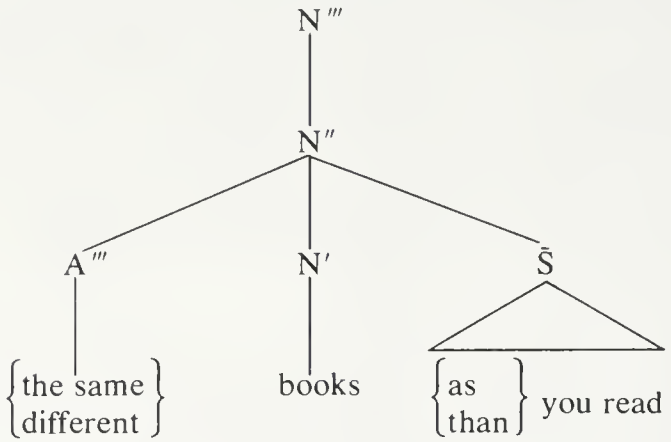


b. sufficiently insensitive to survive Brandeis



This configuration is also the one that appears with *same* and *different*:

(8.21) the same books as you read
different books than you read



Also, as seen in section 8.2, the relationship can pass through a (probably indefinite) number of embedded X‴ and X″ specifiers:

(8.22) a. too many stories about Bill for us to bear (*too* is Q‴ specifier of N″ specifier)
b. as extraordinarily lively as Bill is (*as* is Adv‴ specifier of A″ specifier)
c. sufficiently many men for us to stop counting (*sufficiently* is Q″ specifier of N″ specifier)
d. as much too tall (for the job) as Bill is (*as* is Q‴ specifier of Deg″ specifier of A‴ specifier)
e. as much too much bread as I could stand (*as* is Q‴ specifier of Deg″ specifier of Q‴ specifier of N″ specifier)

The most general condition I have been able to discover for the position of the specifier is this:

(8.23) The specifier binding a degree clause must be the head of a phrase on the leftmost extremity of that X‴ which most directly dominates the degree clause.

This condition certainly holds in all the examples given so far, except for those in V‴ (8.4).

(8.4) a. He gave so little to the fund that he looked stingy, which surprised no one.
b. He didn't say enough to us about how he did it to justify himself, which was a shame.
c. This county has elected {more / fewer} Democrats to the Senate than any other in Missouri, which is hard to explain.

As evidence of the essential correctness of (8.23), consider (8.24). *The* cannot cooccur with a degree clause. (8.23) predicts this: since *the* is the leftmost specifier, the degree clause is bound to *the* instead of to the degree word, producing semantic anomaly.

(8.24) a. the too many stories about Bill (*for us to bear)
b. ?the sufficiently many men (*for us to stop counting)

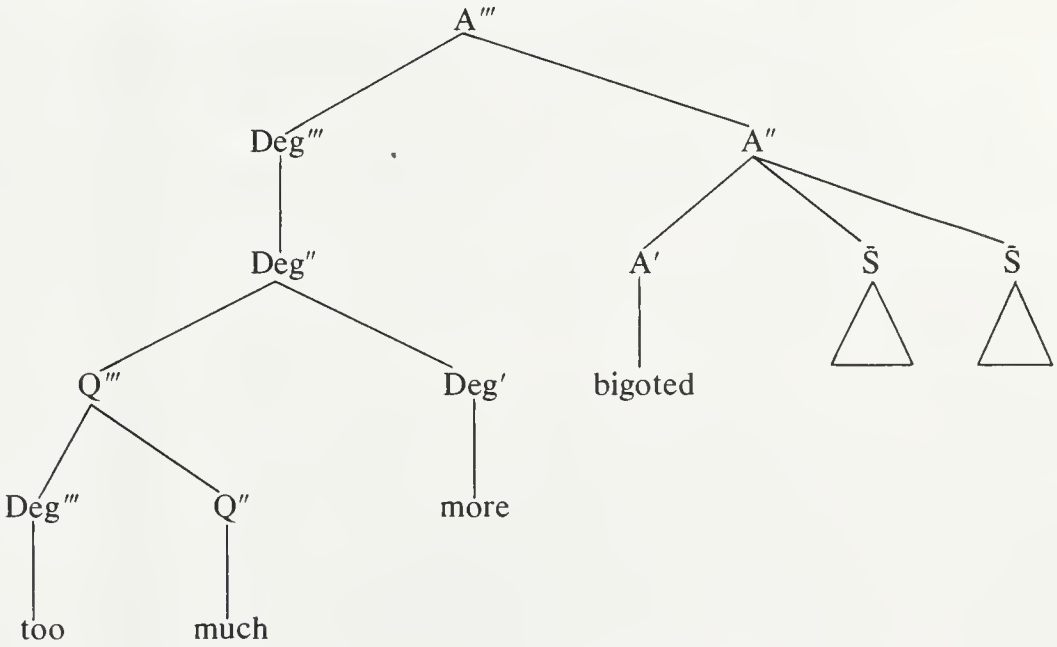
Since there seem to be very few other cases at all in which a degree specifier can be preceded by another independent specifier phrase, the generalization stated in (8.23) appears to be correct.

Note that the condition speaks of the degree word being not leftmost, but the head of a leftmost phrase. This permits it to take its own specifiers:

(8.25) a. [five times too] many men for us to catch
b. [twice as] much soup as we wanted
c. [three pounds more] sugar than was necessary

In fact, this statement of the condition permits two different degree words to bind clauses in the same phrase:

(8.26) too much more bigoted than Vera is for us to like him



Since the lower Deg‴, of which *too* is the head, and the upper Deg‴, of which *more* is the head, are both at the leftmost extremity of the A‴, both *too* and *more* can bind degree clauses. The fact that the clauses are in the order they are in will be accounted for in section 8.7.

It thus appears that (8.23) is reasonably close to the correct relationship between degree clauses and their binding specifiers, though the cases in V‴ still remain to be accounted for.

8.6. Why an Interpretive Theory Is Needed

Once the degree clause is related properly to its associated specifier, the interpretations of these two elements must be related to the rest of the sentence. In the cases of the clauses governed by *enough, so,* and *too*, this presents no further syntactic problem, just one of understanding the semantics. Since the semantic problems must be surmounted no matter where the degree clause is attached in underlying structure, their solution does not affect the decision between the extraposition and the X″-complement theories.

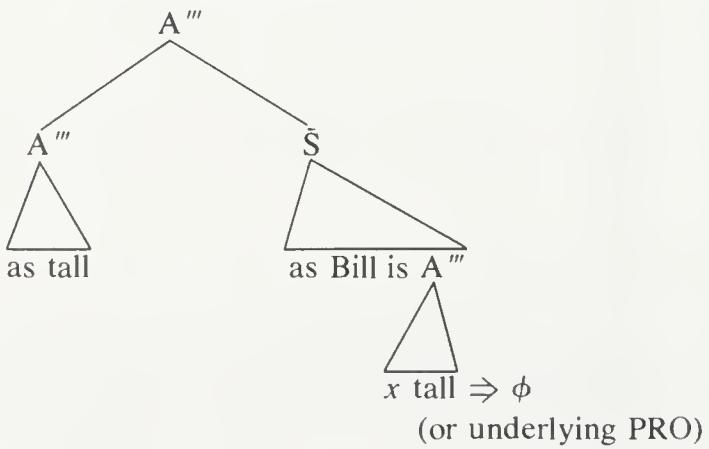
In comparative clauses, though, there is a missing element whose reading bears some relation to the phrase dominating the clause. For example, in *John read more books than Fred read,* the object of the comparative clause is missing. Defining the interpretation of this missing element involves syntactic problems as well as semantic ones, as it turns out.

Selkirk (1970) and Bresnan (1973) argue that this deleted (or PRO) string begins with a *wh*-like dummy degree word *x* and is otherwise identical to (or anaphoric to) the string between the governing degree word and the clause.[6] In (8.27) the bracketed string represents the material missing from the surface.

(8.27) a. John is as tall as Bill is [*x* tall]
b. John has as many books as Fred has [*x* many books]
c. Susan is more incredibly smart than Moe is reputed to be [*x* incredibly smart]
d. we bought as many too many books as you bought [*x* many too many books]

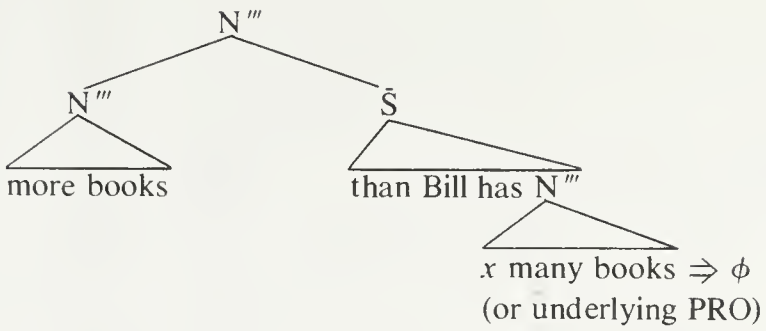
In Bresnan's and Reed's (1974) analyses, the comparative clause is attached outside the constituent dominating the degree word:

(8.28) a.



[6] In Lees (1961), *x* is identified with the Deg *that*, just as *wh* at the time was identified with the Art *the*.

b.

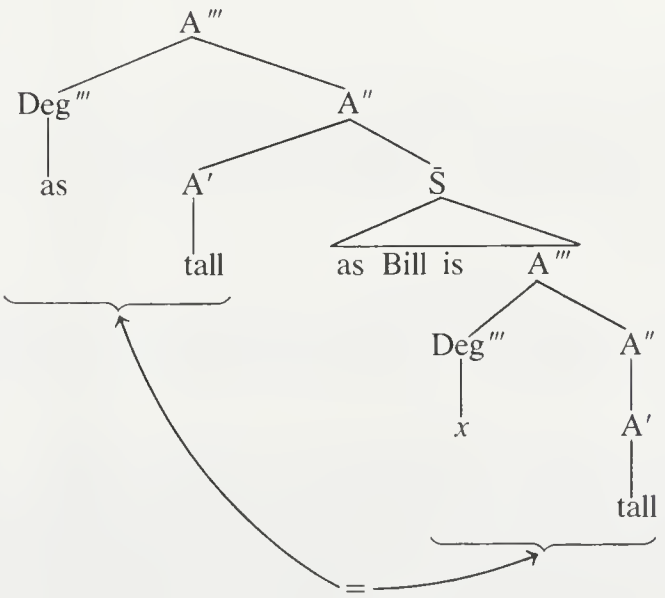


Under this analysis of the surface structure, the deletion or anaphora can be based on a relationship between full constituents. In this it is structurally similar to the Chomsky-adjoined theory of relative clauses, in which the relative pronoun's antecedent is taken to be the entire head NP.

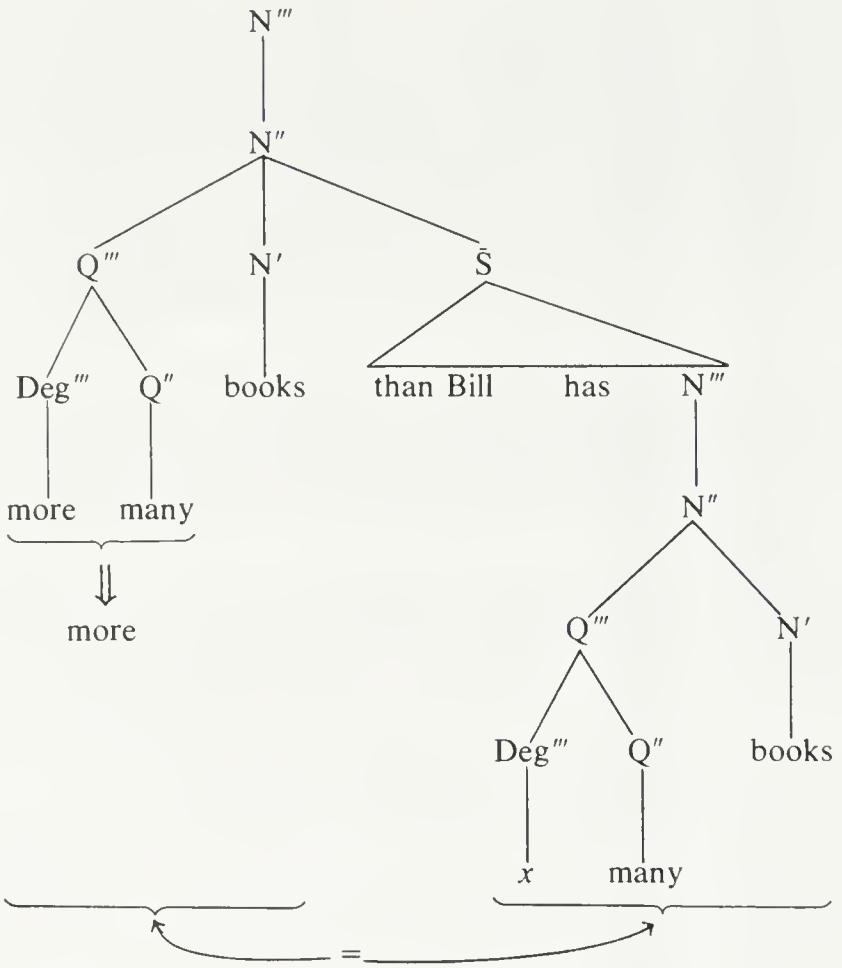
But this analysis is also parallel to the Chomsky-adjoined theory in another way: it violates the theory of phrase structure of chapter 3, having an X‴ whose head is an X‴ instead of an X″. Besides threatening the general theory of phrase structure, it contradicts the conclusion of section 8.2 that degree clauses are attached to X″. That conclusion, in turn, was based on the theory that appositives are attached to X‴, one of the strong points of the present theory of phrase structure.

In the X″-complement theory of degree clauses, though, the constituent relationship of (8.28) is impossible. The upper and lower X‴s cannot be identical, since one contains an S̄ and the other does not. The most one can say is that a substring of the upper X‴ is identical with the whole lower X‴ aside from the binding Deg‴.

(8.29) a.



b.



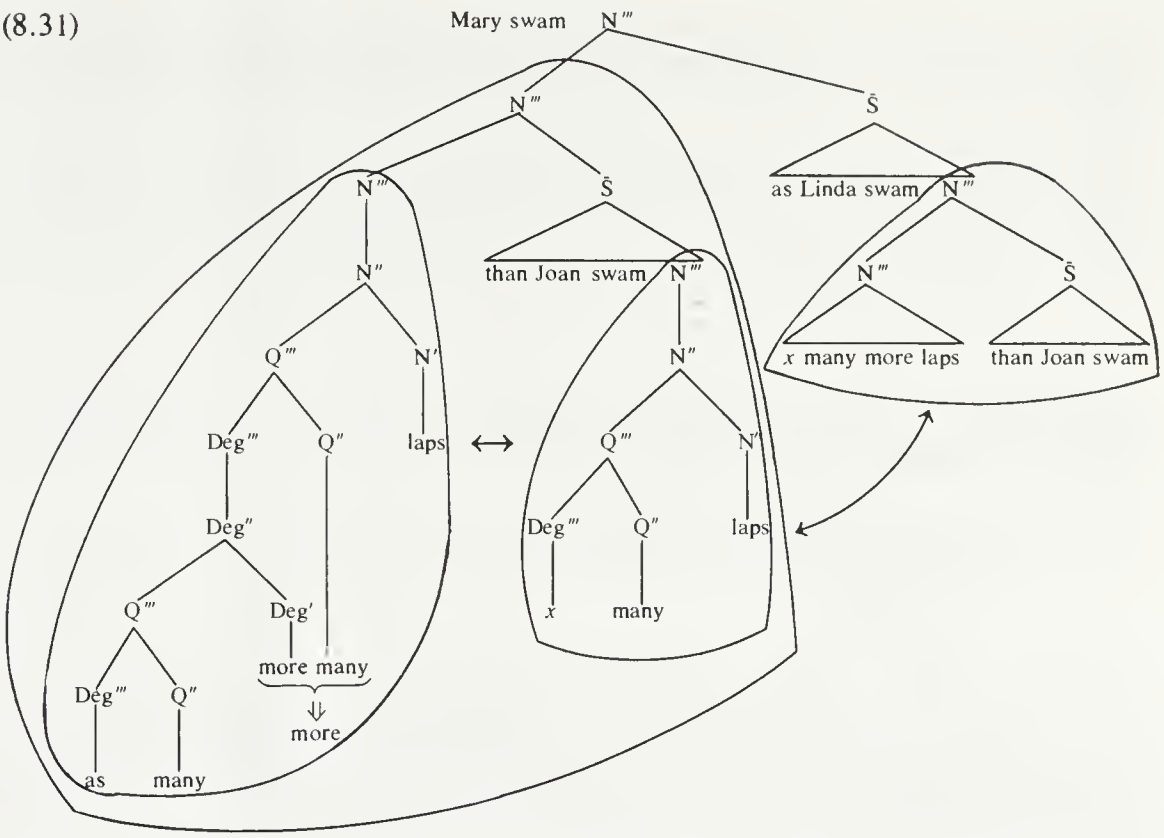
Thus the X″-complement theory of degree clauses has the same apparent difficulty of interpretation as the NP-complement theory of relatives: nonconstituent identity.

The problem is exacerbated by stacked degree clauses such as the ones in Bresnan's example (8.30).

(8.30) Mary swam as many more laps than Joan swam as Linda swam.

The degree words and clauses are related from the inside out: the inner degree clause goes with the second degree word and the outer clause goes with the first degree word. The outer degree clause, furthermore, measures not simply a number of laps, but a number of laps more than Joan swam. In Bresnan's outside-X‴ theory, a stacked structure can be constructed which accounts for this, preserving constituent identity in the anaphoric process:

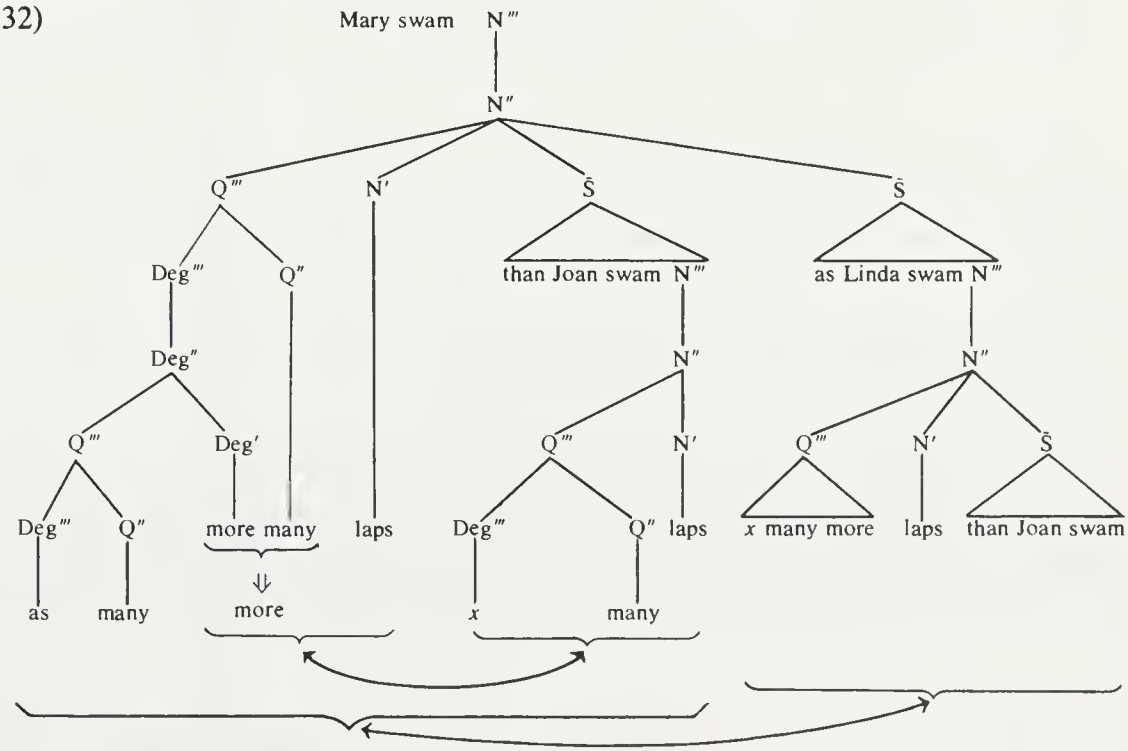
(8.31)



This is precisely parallel to the Chomsky-adjoined theory's account of the anaphoric relations in stacked restrictive relatives (cf. example (7.41), p. 185).

In the X″-complement theory of degree clauses, the anaphoric relations in this example can be established only stringwise, not as structural identity between single full constituents:

(8.32)



In a deletion theory of comparatives this structure would be untenable, since (at least under the usual assumptions) constituent identity, not string identity, is necessary for deletion.

Thus, if we are to adopt the X″-complement theory of comparative clauses, in keeping with the strong version of the $\bar{X}$ Convention, we are forced to develop an interpretive theory for the missing element of comparatives, a theory that is based on identity of strings, not structure. The next section proposes such a theory.

8.7. Interpreting PRO

In an interpretive theory of comparatives, the righthand bracketed elements in (8.29) and (8.32) are replaced by deep structure PRO:

(8.33) a. as tall as Bill is [$_{A'''}$PRO]
b. more many books than Bill has [$_{N'''}$PRO]
c. as many more many laps than Joan swam [$_{N'''}$PRO] as Linda swam [$_{N'''}$PRO]

There are three steps involved in interpreting such structures. The first step is to locate the degree word which binds it, i.e. which satisfies condition (8.23). This part of the process is common to the deletion and interpretive theories. For concreteness, suppose that the degree word is *more*, that the complementizer is *than*, and that the clause is dominated by N‴, as in (8.33b).

The next step of interpretation is to give content to PRO within the clause. In the cases we have been considering, PRO must be of the same category as the X‴ most immediately dominating the clause.[7] Assuming Bresnan's account of the semantics, PRO should be interpreted in the following way. Designate the string between *more* and *than* as *Z*. Replace the Deg‴ dominating *more* by a Deg‴ dominating *x*, and assign PRO the interpretation of the string *x* – *Z*. This string is equivalent to Bresnan's underlying structure.

It must be shown that *x* – *Z* has a coherent interpretation. *Z* contains some string of specifiers (for example *many* in (8.33b)), followed by the N′ and possibly some daughters of N″ to the right of the head. Since *more* is the head of a Deg‴ at the leftmost extremity of N‴, replacing this Deg‴ by *x* yields a complete specifier for *x* –

[7] We are not considering any cases of what Bresnan (1976a) calls Sub-Deletion, for example (i) (elements missing in the surface are bracketed).

(i) She is less beautiful than she is [*x*] funny.
We bought more books than we sold [*x* many] cars.
He has as many too many books as we have [*x* many] too few pencils.

Here the deleted elements are identical only to part or all of the *specifier* of the phrase dominating the degree clause. These cases do not constitute a potential threat to the X″-complement theory even with deletion, since the parallel strings are full single constituents, not partial constituents as in (8.29) and (8.32). I leave for future research how Sub-Deletion should be stated interpretively, though I assume it will retain many of the features of Bresnan's elegant deletion analysis.

Z. The N′ of course has a coherent interpretation. Any N″ complement within Z is either an ordinary restrictive modifier, which can be integrated into the interpretation independent of the rest of the contents of the N″ complement, or else it is of the type which must be bound by a specifier. If it is of the latter type, the binding specifier must also be contained within $x - Z$ for the interpretation of $x - Z$ to be coherent. But this is precisely what is found in the case of stacked degree clauses: any degree clause between a degree-word–degree-clause pair is matched by its own corresponding degree word. For example, in (8.33c), the first PRO is in a clause governed by *more*, so $x - Z$ is *x many laps*. The second PRO is in a clause governed by *as*, so $x - Z$ is *x many more many laps than Joan swam PRO*, exactly the desired interpretation.

The final step in interpreting the degree clause is to assign an interpretation to x itself. Presumably this involves logically binding it in some way to the governing degree words, by a process similar to that for restrictive relative pronouns. But since an appropriate logical formalization of degree words is not yet extant, this step must be left for future research. Again, this problem is common to the deletion and interpretive theories.

Thus, of the three steps involved in an interpretive theory of degree clauses, the first and last are independently necessary in the deletion theory. The middle step, the interpretation of PRO, has been shown to provide coherent interpretations of the desired sort; this step can be carried out with the comparative clause generated as a daughter of X″ instead of as a daughter of a higher X‴, despite the nonmatching of constituent structure. Thus, like relative clauses, comparative clauses need not be exceptions to the general phrase structure rule schema.

In turn, appositives must be attached higher than degree clauses, as shown in (8.2)–(8.5). If we had needed to generate comparative clauses at a higher node than X″, this would have forced appositives into a position which was also an exception to the phrase structure rule schema. Generating comparative clauses in X″, however, permits appositives to remain in X‴, in conformity with the arguments of sections 4.1 and 7.2.

8.8. Recapitulation

The argument has gone as follows. First, whatever the specific details of one's analysis of degree clauses, a theory of grammar is necessary in which cross-category generalizations can be made, since degree clauses occur in such a wide variety of positions. Second, the theory that degree clauses are generated in their surface position has certain syntactic advantages over the theory that they are extraposed from Deg‴. But this requires the more complex semantic rule (8.23) to establish the binding of the degree clause to the Deg‴. Finally, in order to preserve the X″-complement theory of comparative clauses, thus avoiding a weakening of the general theory of phrase structure, we had to face the problem of string rather than constituent identity. An interpretive rule was developed which was shown to provide coherent interpretations even to the complex stacked comparatives.

Along the way, we have observed somewhat distant but still striking generalizations among degree clauses, restrictive relatives, and possibly even subject complements. It has been particularly instructive to see how the determiner, Chomsky-adjoined, and NP-complement theories of relatives discussed in chapter 7 are paralleled by corresponding theories of degree clauses, how much alike the arguments for and against the three theories are in the two domains, and how in both cases an interpretive theory of the missing constituent in the clause is necessary to save the X″-complement theory. If nothing else, these similarities constitute an argument on the metalevel in favor of a close, though abstract, relationship between the two constructions.

9: Deverbalizing Rules

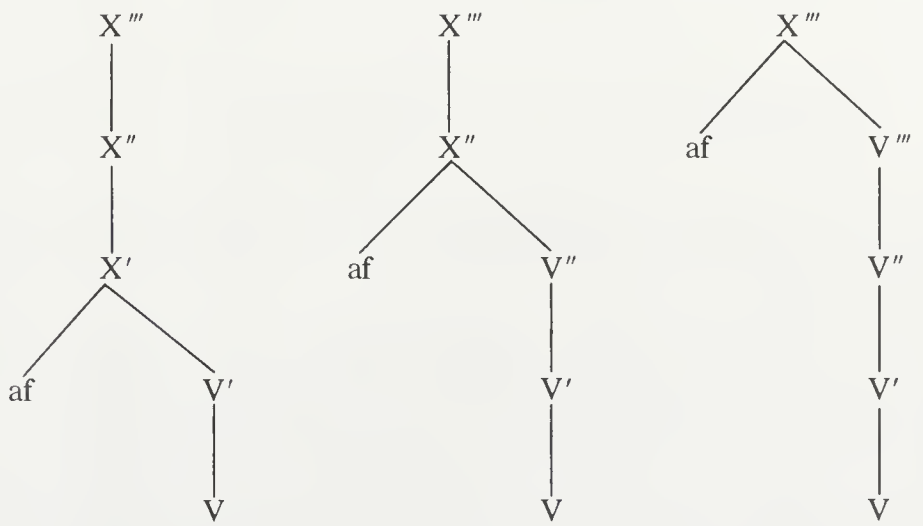
9.1. The Deverbalizing Rule Schema

Section 3.6 pointed out that gerundive nominals are an exception to the normal hierarchical arrangement of syntactic categories predicted by the $\bar{X}$ Convention: they are dominated by NP but contain no head N. It was suggested that the phrase structure rule that describes them is one of a class of systematic exceptions to the Uniform Three-Level Hypothesis, characterized by the following phrase structure rule schema:

(9.1) *Deverbalizing Rule Schema*
$X^i \rightarrow af - V^i$

(9.1) expands a supercategory of the lexical category X as a supercategory of verb of the same level, plus a grammatical formative or affix. This chapter will motivate several rules of this type in English. The overall effect of applying a deverbalizing rule will be an X''' whose head is not an X but a V; such an X''' will display some properties of X'''s and some properties of sentences. (9.2) shows the three possible skeletal structures, ignoring complements and specifiers.

(9.2) a. $i = 1$ b. $i = 2$ c. $i = 3$



The structures grow internally less X-like and more sentencelike, as i goes from 1 to 3 and as the complements and specifiers are determined by more supercategories of V

and fewer supercategories of X. Externally, however, they are all still X‴s and occur in positions characteristic of X‴, not S.

Note that (9.1) strongly restricts the possible exceptions to the major rule schema in several important respects. First, the only category that can appear on the right is V: this seems to be true of all category-switching rules I have encountered. Second, even when the category of the head of a phrase is switched to V by the application of one of these rules, the hierarchy of primes is undisturbed: there is no instance, say, of A′ expanded as an affix plus V″ or V‴. Third, at the level where the category is switched by a deverbalizing rule, there are no further complements or specifiers other than the grammatical formative that marks the change of category. In each of these three respects, it is clear what kind of evidence would entail weakening the restrictions; for the moment, though, (9.1) seems like a plausible initial hypothesis.

9.2. Gerundive Nominals

Gerundive nominals are the most NP-like of the three complement types studied by Rosenbaum (1967). As Emonds (1970) shows, they differ from *that* and *for-to* complements in that they occur in all NP positions, including subject position in questions, relative clauses, and *for-to* complements.

(9.3) a. What would { John's leaving / *that John left / *for John to leave } reveal about him?

b. a man who { John's leaving / *that John left / *for John to leave } would irritate

c. It would be disgraceful for { John's leaving / *that John left / *for John to leave } to bother us.

Thus gerundive nominals are apparently NPs, though they lack a noun head. To find the level at which they change from N^i to V^i, we investigate what complements and specifiers they take.

The level which explains the most about the structure of gerundive nominals is the X″ level. To recapitulate the discussion of section 2.1, gerundive nominals permit within them all the transformations such as Dative Shift and Particle Movement characteristic of V′ but not N′. They are preceded by preverbal adverbs, characteristic of V″, not by adjectives, characteristic of N″. They may contain perfect and progressive aspect, which we assigned to V″ in section 3.5. They may contain *because* and result clauses characteristic of V″ and not of N″.

On the other hand, gerundive nominals do not contain tense and modal, character-

istic of V‴ and absent in N‴. They contain a possessive NP in the subject, characteristic of N‴, not of V‴. As Schachter (1976) points out, they occur with a limited range of articles, as in *There is no enjoying this world without thee,* thus behaving like N‴, not V‴. Like N‴ and unlike V‴, they do not permit sentence adverbial PPs; (9.4) compares gerundive nominals to *that*-clauses to show that the impossibility of these particular sentence adverbials is not a restriction of all subordinate clauses:

(9.4) *John's { to our delight / in his haste / for some reason } leaving so early didn't distress Sue.

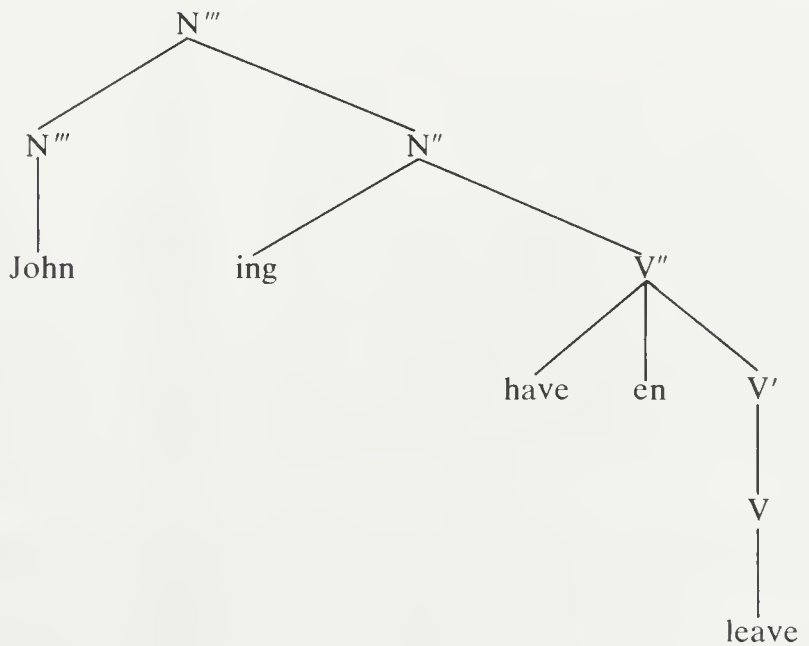
That John(,) { to our delight / in his haste / for some reason }(,) left so early didn't distress Sue.

Further arguments of this sort appear in Schachter (1976) and Horn (1975), and, though advocating a different theory, in Williams (1975). The point is clear. Gerundives have the constituent structure of sentences up to the X″ level and that of NPs above that. Thus the appropriate skeletal structure is (9.2b), and the phrase structure rule is (9.5).

(9.5) *Gerundive Nominals*
$N'' \rightarrow$ ing – V″

Ing, like other affixes, undergoes Affix Hopping to attach to the first verbal form. Thus the underlying form of *John's having left* is (9.6), exactly its surface form except for the insertion of *poss* and the attachment of verbal affixes.

(9.6)



9.3. Tensed Complements and Free Relatives

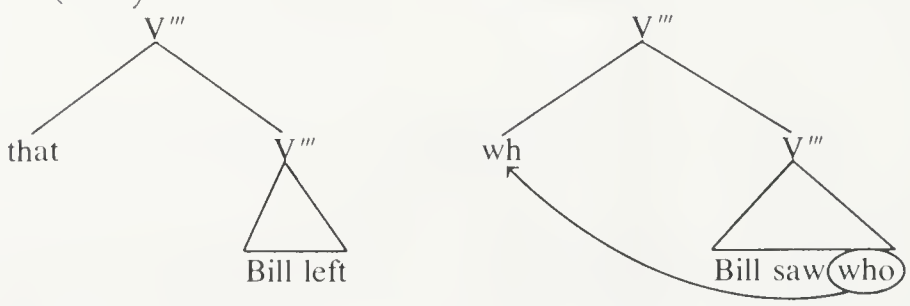
Let us turn now to tensed subordinate clauses. As Bresnan (1970) shows, these include not only *that*-clauses but also comparative clauses with the complementizers *as* and *than*, and indirect questions with the complementizer *wh*. All of these are indubitably V‴s, since they contain, in addition to all the V″ structure, tenses, modals, and nongenitive subjects; what sentence adverbials they lack can plausibly be excluded for semantic reasons. The discussion of chapter 3 showed that *that*-clauses, because their movement generalizes with that of other X‴, must be X‴ of some sort. Furthermore, it was shown that the complementizer is not within the V‴ that dominates the subject and the auxiliary. Thus a structure like (9.2c) appears most appropriate for tensed complement clauses. The only question is what category the dominating X‴ should be.

Chapter 4 alluded to the dispute over whether *that*-clauses are dominated by NP in underlying structure. We have taken the position that they are not, that their position at the end of X′ and X″ is base-generated, and that their appearance in the subject position of certain clauses results from the application of the root transformation Intraposition. Thus we claim that the X‴ in (9.2c) is V‴ for tensed complement clauses, and that there is no separate category S̄ as Bresnan (1970) and Chomsky (1973) assume. Under this theory, the phrase structure rule generating tensed complements is (9.7), and typical structures are given in (9.8).[1]

(9.7) *Tensed Complements*

$$V''' \rightarrow \left\{ \begin{matrix} \text{that} \\ \text{as} \\ \text{than} \\ \text{wh} \end{matrix} \right\} - V'''$$

(9.8)



Of course, if it were to turn out that *that*-clauses are dominated by NP, a possible realization of (9.1) to describe this would be (9.9).

(9.9) $N''' \rightarrow$ that - V'''

[1] In principle this rule could incorrectly apply to its own output to yield sequences of complementizers. I assume, perhaps unjustly, that either semantic considerations or a surface constraint prevent such an eventuality.

But there is another kind of tensed complement for which a dominating N‴ looks a great deal more plausible: free relatives.

Free relatives occur in all NP positions, including those in (9.3) from which *that* and *for-to* complements are excluded:

(9.10) a. What would what the FBI found out reveal about John?
b. a man to whom what you found out would be a nuisance
c. It would be disgraceful for what you found out to be revealed.

Furthermore, free relatives, like gerundive nominals and unlike *that* and *for-to* complements, do not have an alternate extraposed position:

(9.11) a. { What John saw / John's seeing a mouse / That John saw a mouse / For John to see a mouse } would surprise no one.

b. It would surprise no one { *what John saw / *John's seeing a mouse / that John saw a mouse / for John to see a mouse }.

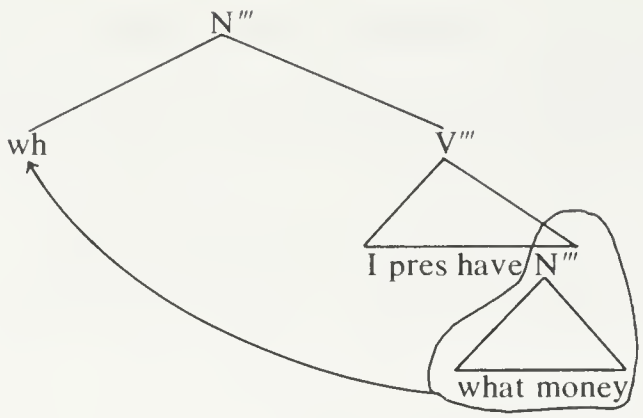
Thus free relatives behave externally much more like NPs than *that* and *for-to* complements do.

It is generally assumed that a free relative is derived from a noun phrase with a relative clause and some sort of dummy head, such as *the thing which John saw* or *PRO which John saw*. However, as Jespersen (1909–1949, Vol. III) points out, such a derivation would not provide an immediately plausible source for free relatives with *wh-ever*, as in *Whatever John saw interested me, We were interviewed by whoever we met*. Still worse are free relatives of the form *what money I have, what help you can offer him*, etc., in which the *wh*-word is followed by a noun also fronted from the relative clause. Underlying sources like *the thing* [$_S$*I have wh-money*] or *PRO* [$_S$*you can offer him wh-help*] do not make sense as relative constructions—particularly if we adopt the interpretive view of relative pronouns argued for in chapter 7.

Jespersen argues that the clause itself should be considered as the entire NP. In the present theory, this can be expressed as the claim that free relatives are generated with phrase structure rule (9.12), another deverbalizing rule; then *what money I have* is derived from (9.13).

(9.12) *Free Relatives*
$N''' \rightarrow wh - V'''$

(9.13)



This hypothesis enables us to make a structural distinction between indirect questions and free relatives, while still claiming that their great similarity is no accident. Quite simply, free relatives are N''' dominating *wh* - V''', and indirect questions are V''' dominating *wh* - V'''. This would be the structural difference, for example, in the well-known old minimal pair *What lay on the table was the tissue* (free relative) and *What lay on the table was the issue* (indirect question).[2]

Thus if there is any construction in which an N''' immediately dominates an S, it is much more likely to be free relatives than *that*-complements, since free relatives occur in all NP positions and are most easily derived by a rule like (9.12). Hence we now have three different applications of rule schema (9.1). (See also note 6, p. 238.)

9.4. Gerundive PPs

There are some *-ing* complements which, unlike most gerundives, alternate with PPs rather than NPs:

(9.14) a. John kept Bill {running / at a run}.

b. Bill remained {working / at work}.

c. Moe went on {working / with his work}.

There is no possibility of a subject in these *-ing* complements; further, like the corresponding PPs and unlike most gerunds, they cannot cleft: **It was running that John kept Bill.* The subcategorization of the verbs in (9.14) would be simplified if in fact these complements were not NPs but PPs. Jackendoff (1976) shows that they behave semantically as expressions of location, so PP constituency is not semantically implausible.

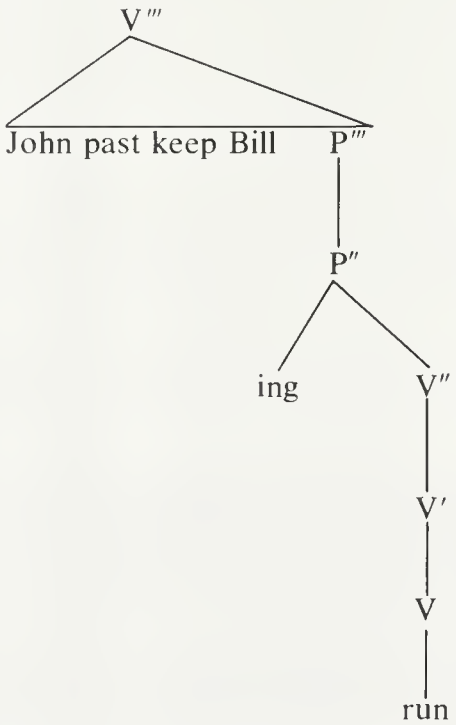
[2] The latter of these examples thus has a V''' in surface subject position. I believe a stylistic inversion has taken place—notice that the highest stress in this sentence is on *table*, in the middle of the sentence, not on *issue*, the expected position. By contrast, in the free relative sentence, the highest stress is at the end, as expected.

This class of gerunds allows certain V″ complements (e.g. *Moe went on doggedly leaving for work at 6 a.m.*), though others are excluded, probably for semantic reasons. Thus if they are generated by a deverbalizing rule, they must be of structure (9.2b) or (9.2c). If they are dominated by PP, choosing (9.2b) would automatically explain the absence of subject, tense, and modals, since P‴ contains none of these. The phrase structure rule would be (9.15), and the structure of (9.14a) would be (9.16).

(9.15) *Gerundive PPs*

$P'' \rightarrow$ ing - V''

(9.16)



Emonds (1970) suggests that the progressive *be* belongs to the class of verbs that takes this sort of *-ing* complement. The main attraction of Chomsky's (1957) analysis of the progressive auxiliary was that it explained the inability of progressive to be followed by modals, perfect, or another progressive. But Emonds shows that this may be part of a more general set of semantic restrictions, since the verbs of (9.14) cannot be followed by perfect or progressive either.

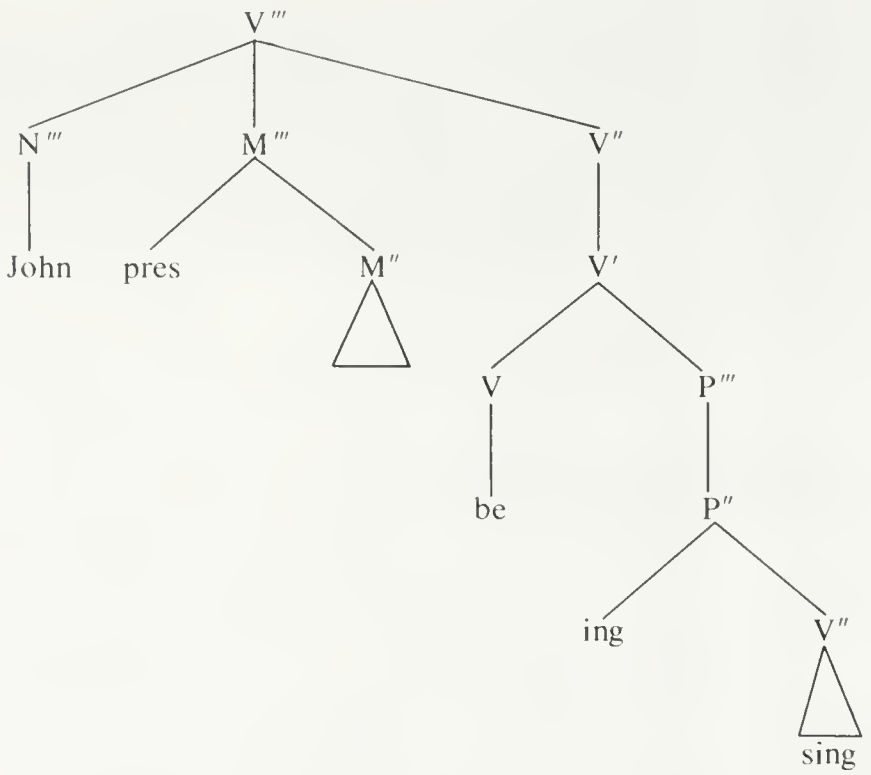
(9.17) a. *John kept Bill $\left\{\begin{matrix}\text{having run} \\ \text{being running}\end{matrix}\right\}$.

b. *John remained $\left\{\begin{matrix}\text{having worked} \\ \text{being working}\end{matrix}\right\}$.

c. *Moe went on $\left\{\begin{matrix}\text{having sung} \\ \text{being singing}\end{matrix}\right\}$.

Notice also that *be* takes a PP semantically parallel to the progressive: compare *John is working* to *John is at work*. Hence, however these restrictions are accounted for, they can be generalized to progressive *be* if *John is singing* is assigned the structure (9.18).

(9.18)



As further support for this analysis, Jackendoff (1976) shows that the verb *be* fills a gap in the semantic paradigm given by the verbs in (9.14), and that the interpretation of the progressive predicted by this semantic analysis is not implausible, at least in these cases. Hence we have some reason to consider progressive *be* as a main verb which takes a gerundive PP complement.

Other than the absence of subjects and the alternation with PPs, I have given little evidence that these gerunds are dominated by PP. One further argument concerns the transformation which inverts a simple verb with the subject in the presence of a preposed locational PP (see Emonds (1976, 37–40) for a formulation of this rule):

(9.19) Into the opera house raced Harpo.
Up the stairs ran the criminals.
On Cambridge Common stands a statue.

When preposed, certain gerunds cause such an inversion too:

(9.20) Buried here lies the producer of *Gone with the Wind*.
Bouncing out of the room came the garbagemen.
Screaming down the hall ran two famous linguists.

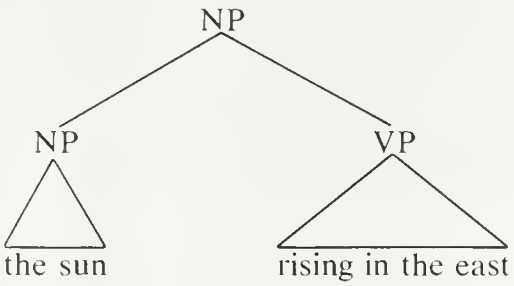
Note that we cannot appeal to a cross-categorial formulation of the inversion rule in this case, since no preposed verbal constructions other than these particular subjectless gerunds cause the inversion. If, however, they are dominated by PP, as in (9.16), the inversion rule can be stated simply in terms of PP.

Another use for rule (9.15) may be to account for the gerundive constructions in (9.21), discussed by Akmajian (1977).

(9.21) The sun rising in the east is a beautiful sight.

Akmajian argues that *the sun rising in the east* is not a clause, but rather that *the sun* is the head of the construction. He proposes the following structure:

(9.22)



Unfortunately, this structure contains two configurations ruled out by our theory of phrase structure: NP over NP, and VP dominated by some category other than S or af-X″.

In order to settle the first problem, we observe that these complements occur after restrictive relatives (9.23a), appear with proper nouns (9.23b), and exclude appositives (9.23c,d).

(9.23) a. The man you met falling on his face was a sad sight.
b. Bill falling on his face was a sad sight.
c. ?*Bill, who we all love, falling on his face(,) was a sad sight.
d. *Bill falling on his face, who we all love, was a sad sight.

This suggests that the gerundive may be treated as a new type of N‴ complement, in alternation with appositives.[3]

[3] This proposal is substantively different from Akmajian's analysis in two ways. First, in object position, these gerundive constructions permit appositives: *We saw Bill, who we all loved, driving away.* Akmajian shows that in object position there are two distinct structures for the same string: one with the object dominating the gerund, yielding passives like *Bill driving away could not be seen by anyone,* and one with the main VP dominating the gerund, yielding passives like *Bill was seen driving away.* Akmajian claims the latter is derived from the former by an extraposition rule. However, this would not predict the difference in the possibility of appositives in subject vs. object position. As an alternative, we may claim that if the gerundive is generated in the NP, it always excludes an appositive, but that these gerundives can also be base-generated in VP, in which case the object is permitted an appositive. We thus reject Akmajian's extraposition rule in favor of use of the base. In fact, if these gerunds are PPs, as I have claimed, the necessary phrase structure rule already exists, since it generates *Bill was seen on TV,* with a simple PP, parallel to *Bill was seen driving away,* with a gerundive PP.

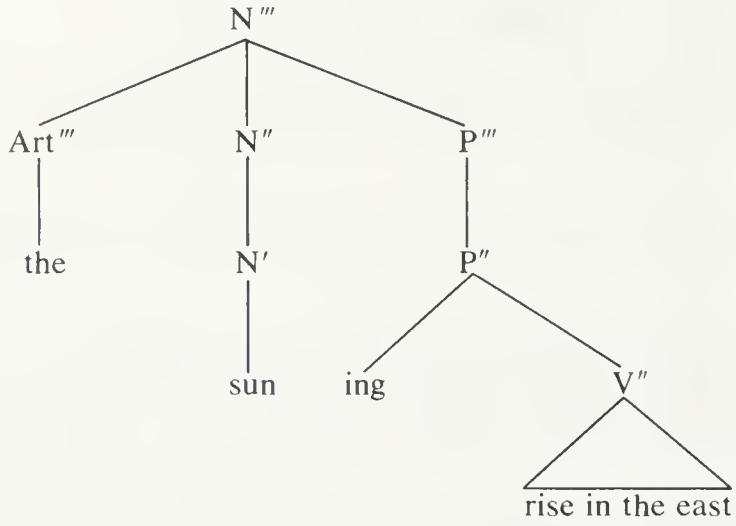
A more serious problem for the present analysis is the existence of passive gerunds of this sort, for instance *Bill being struck by lightning was a sad sight.* In Akmajian's structure the Passive can take place under the domination of the upper NP; but in the present analysis, there is no complete NP to the left of the gerund that could be moved there by the Passive transformation. In view of current doubts about the existence of the Passive transformation (cf. Freidin (1975), Schachter (1976), Bresnan (1976b), for example), this argument does not seem anywhere near as damning as it would have a few years ago. One can at least leave the issue in abeyance instead of giving up altogether.

As for the second problem with (9.22), the bare VP, we observe that these gerundives alternate with PPs of approximately the same semantic import and with the same syntactic distribution:

(9.24) a. Gretchen at the spinning wheel is lovely.
b. A man we like so much in such a mess is a disheartening sight.
c. ?*Bill, who we all love, in such a mess(,) was a sad sight.
d. *Bill in such a mess, who we all love, was a sad sight.

This suggests that the lack of subject, tense, and modal is explained not, as Akmajian suggests, by generating a bare VP, but rather by generating V″ under P″, again using phrase structure rule (9.15). Thus our alternative to (9.22) is (9.25).[4]

(9.25)



As a final example of phrase structure rule (9.15), consider what happens if we move the P‴ in (9.25) down one level to N″. Since in N″, P‴ precedes relative clauses, we should get a class of subjectless gerunds which follow N′ complements and precede restrictive relative clauses and appositives. And indeed there is such a class, the so-called gerundive relatives:

(9.26) a. A man *owning a Cadillac* is not to be envied these days.
b. A father of three children *begging for free time* should not be sneered at.
c. The man *paddling the canoe*(,) who is wearing a red hat(,) is your cousin.

Gerundive relatives are often thought to be derived from ordinary relatives by deletion

[4] A more traditional transformational analysis, of course, would probably suggest the reverse: that the PP cases are derived from the gerundives by a *being* deletion. Though I suspect there is such a transformation involved in cases like *John angry is a scary sight*, it is not so plausible for *John with a scar like that is a sad sight*, since *with* in this use does not occur after *be*.

of the relative pronoun plus *be*, so-called Whiz-Deletion. (9.26a) shows that this is impossible, since **a man who is owning a Cadillac* is ungrammatical. Williams (1975) gives a number of other arguments against Whiz-Deletion, which I will not reproduce here.

(9.26b) shows that gerundive relatives follow N′ complements; (9.26c) shows that they precede other restrictive and nonrestrictive relatives. Hence they are in the position of PPs under N″. Since they are without subjects, tenses, and modals, and since there is no NP position under N″ after the head, it seems most plausible on all grounds to invoke (9.15) again.

This section has shown that P″ quite plausibly expands as V″ in at least three different circumstances in the grammar. In each case the verb appears as a present participle; in each case a subject is impossible; in each case PPs appear in identical positions with parallel interpretations. We thus have evidence for a heretofore unsuspected phrase structure rule, another instance of phrase structure schema (9.1).

I have said nothing about infinitival complements so far in this chapter, and perhaps a word is in order at this point. The problems posed by infinitival complements are in many ways like those of gerundives; they share the obligatory absence of tense and modal, and the sometimes optional, sometimes obligatory absence of a subject. But infinitival complements are far more widespread and varied than gerundives, and thus less easy to categorize.

The standard approach to infinitival complements, of course, is that they are all S̄s with *for-to* or (following Bresnan (1972) and Chomsky (1973)) *ϕ-to* complementizers. It has occasionally been suggested that the "obligatory Equi" complements, such as those with *try*, are generated with bare VP complements. The approach to complementation proposed in this chapter suggests that a more enlightened use of phrase structure rules might help sort out some of the complications of infinitivals. For example, the *for-to* clauses which alternate with *that*-clauses, such as the complements of *prefer, wish,* etc., are probably best analyzed as V‴, as usual. But the many infinitivals which alternate with PPs of parallel interpretation and not with *that*-clauses, for example those in (9.27), may well be expansions of PP similar to gerundive PPs.[5]

(9.27) a. He tried {for a new job / to get a new job}.

b. We forced Bill {into shutting up / to shut up}.

c. He's too fat {for shorts / to wear shorts}.

d. They chose Max {for the job / to get the job}.

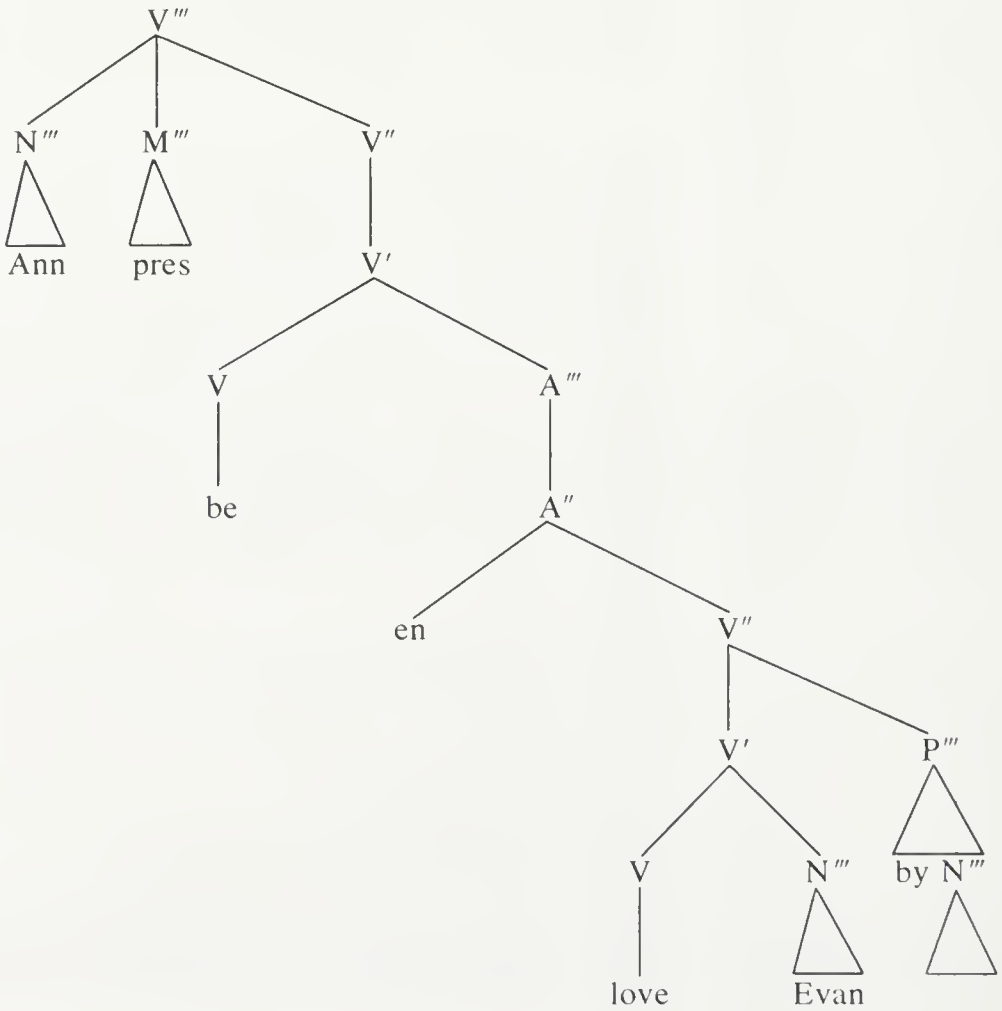
[5] Faraci (1974) presents a wide range of examples of this type, suggesting a phrase structure rule PP → P – S

It is not clear to me, however, what sort of phrase structure rule should be used to generate infinitivals, particularly *for-to* infinitivals. The fact that parallels in other languages to full *for-to* complements are relatively rare, by comparison with parallels to *that*-clauses, simple *to*-complements, and gerundives, may indicate that they are some particularly exotic variation on the phrase structure rule schema for which there is no other evidence in English. In any event, I leave the issue for future research.

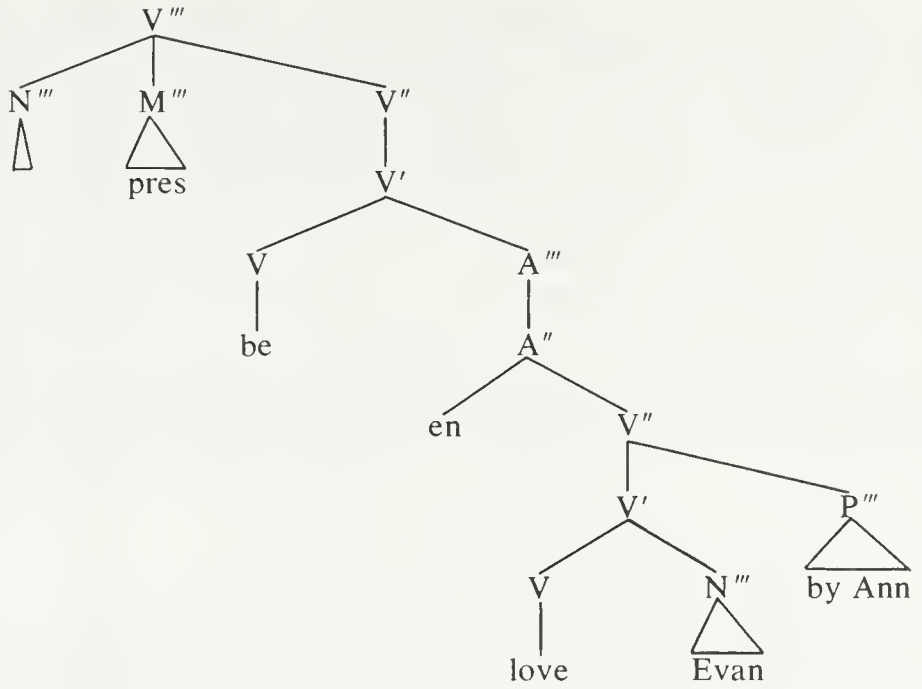
9.5. Passive VPs

A final tentative example of a deverbalizing rule comes from Fiengo (1974), who revives a proposal of Chomsky (1955), in which the passive verb phrase is an AP. The passive auxiliary is present in deep structure, and the necessity of preposing the direct object is guaranteed by semantic constraints (which, however, Fiengo does not spell out in much detail). The deep structure of a passive is (9.28a) or (9.28b); its surface structure is (9.28c) in the present framework.

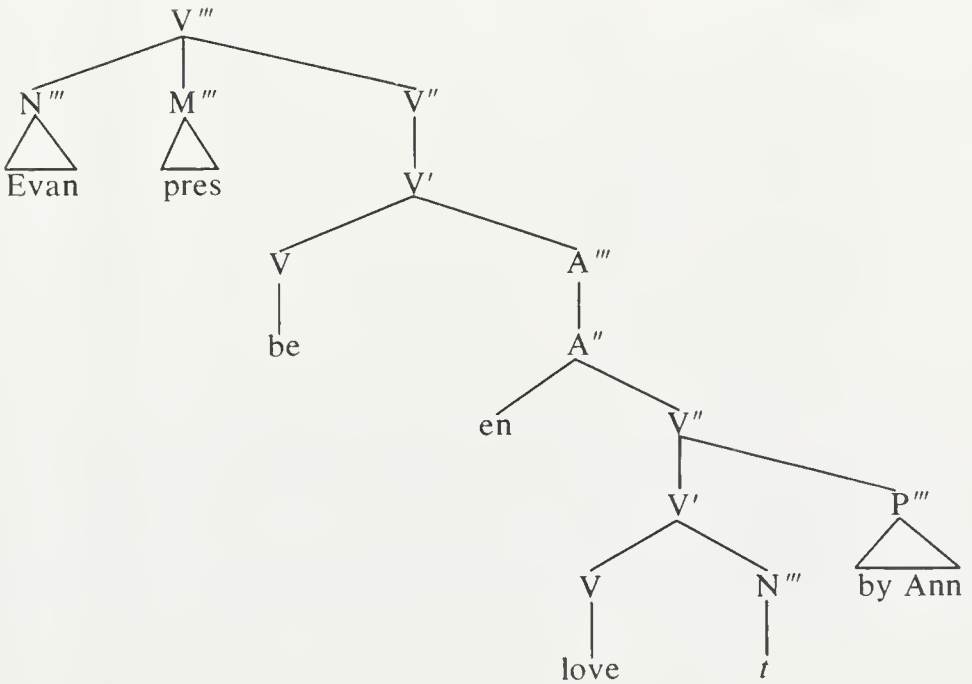
(9.28) a.



b.


V‴
N‴
M‴
V″
pres
V′
V
A‴
be
A″
en
V″
V′
P‴
V
N‴
by Ann
love
Evan

c.


V‴
N‴
M‴
V″
Evan
pres
V′
V
A‴
be
A″
en
V″
V′
P‴
V
N‴
by Ann
love
t

The *t* in direct object position in (9.28c) is the "trace" left behind by movement rules in Fiengo's theory, a pronoun which is anaphorically bound to the NP originally occupying that position. We will not go into Fiengo's motivation for this theory, as it goes far beyond the scope of the present study. However, we observe that the analysis requires a deverbalizing phrase structure rule, here represented as (9.29), in that there is an AP without an adjectival head.

(9.29) *Passive VP*
$A'' \rightarrow$ en - V''

One piece of evidence for identifying the passive verb phrase as AP is that passive verb phrases occur, with the correct meaning, after certain verbs other than *be*, for example (9.30).

(9.30) a. Fred got arrested by the cops.
b. Mary got Kathy arrested by the cops.
c. George had Jerry studied by a team of shrinks.

Since these environments also permit APs (*Fred got sick, Mary got Kathy sick, George had Jerry angrier than a swarm of bees*), an A'' structure like the one in (9.28) would make the generation of these otherwise troublesome examples rather simple. Such a theory would, of course, have to develop a semantic explanation of why passive VPs do not occur in all adjectival environments, a problem which I must leave for future research.

9.6. Final Remarks

We have shown that there is evidence for at least five instances of rule schema (9.1) in English; we repeat these here.

(9.5) *Gerundive Nominals*
$N'' \rightarrow$ ing - V''

(9.7) *Tensed Complements*
$$V''' \rightarrow \begin{Bmatrix} \text{that} \\ \text{as} \\ \text{than} \\ \text{wh} \end{Bmatrix} - V'''$$

(9.12) *Free Relatives*
$N''' \rightarrow$ wh - V'''

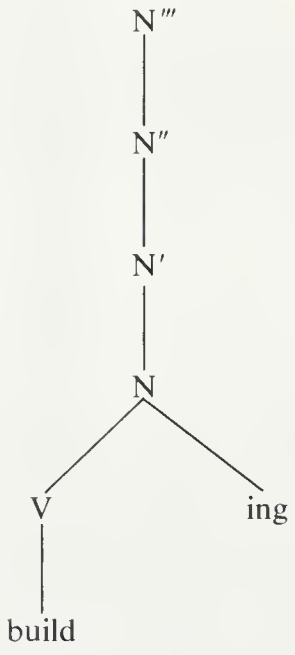
(9.15) *Gerundive PPs*
$P'' \rightarrow$ ing - V''

(9.29) *Passive VPs*
$A'' \rightarrow$ en - V''

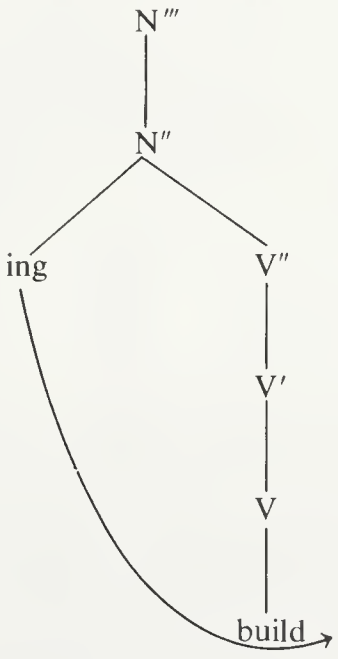
Though not all of these rules have been equally well motivated, their similarity justifies the claim that they constitute a principled class of exceptions to the Uniform Three-Level Hypothesis.

It is interesting to notice what happens if we let $i = 0$ in schema (9.1), a possibility we have not considered so far. The resulting rules are of the form $X \rightarrow af - V$ and are congruent with a subset of the word formation rules in the lexicon, for example those relating derived nominals to verbs. And in fact, the historical roots of (9.5), (9.15), and (9.29) are undoubtedly in word formation rules which still exist in the lexicon of English. Corresponding to gerundive nominals are derived nominals of Germanic verbs, such as *building* and *writing*. It is not difficult to imagine a historical restructuring in which the affixing is taken to be at a higher level, thus enabling the verb transparently contained in the noun to take some of its verbal complements. The reanalysis is from (9.31a) to (9.31b).

(9.31) a. *Deverbal Noun*

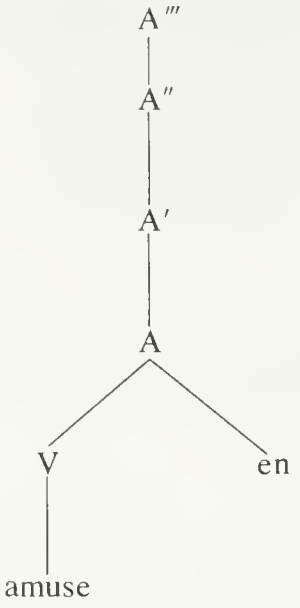


b. *Gerundive Nominal*

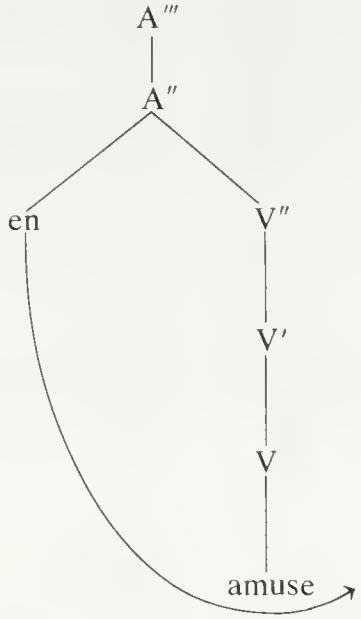


Similarly, the past participles used in passive sentences are paralleled by past participles used as passive adjectives such as *amused, disturbed, shut,* and so forth. Again, one can imagine a reanalysis of the affix at a higher level, so that alongside the lexical process (9.32a) arose the syntactic process (9.32b). Wasow (1977) has an interesting discussion of syntactic tests that distinguish the two constructions from one another.

(9.32) a. *Passive Adjective*

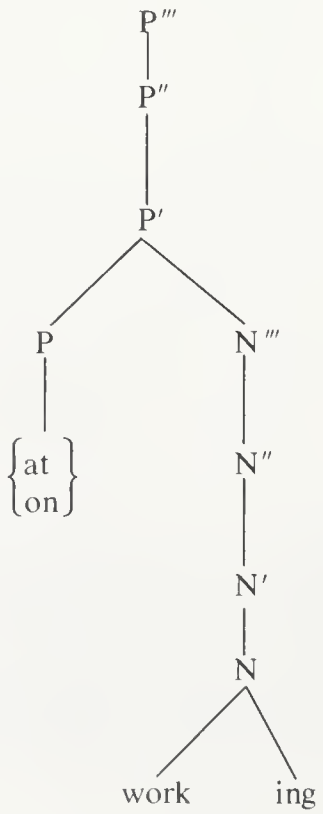


b. *Passive VP*

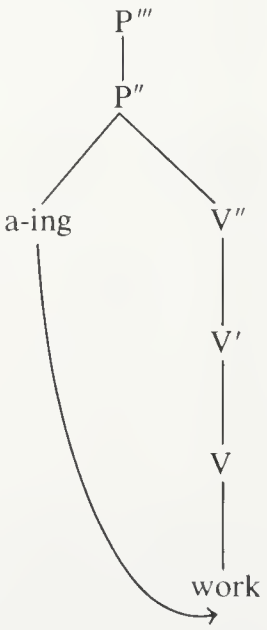


Gerundive PPs are historically derived from the deverbal noun preceded by the preposition *on* or *at*, a possible relic of which is the proform *at it* used to refer to these constructions. The preposition degenerated to the prefix *a-*, which in turn dropped off altogether. The transition is represented as (9.33); the structure of the intermediate step is conjectural.

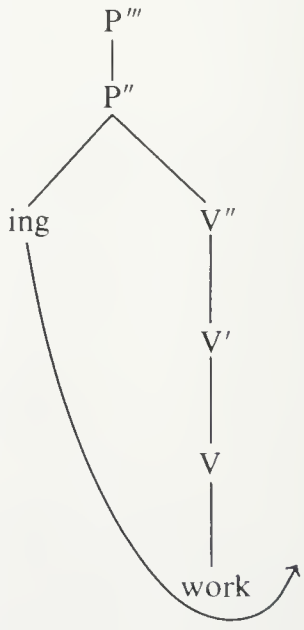
(9.33) a.



b.

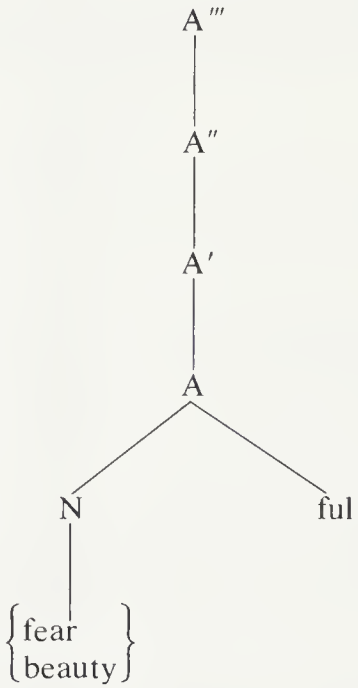


c.

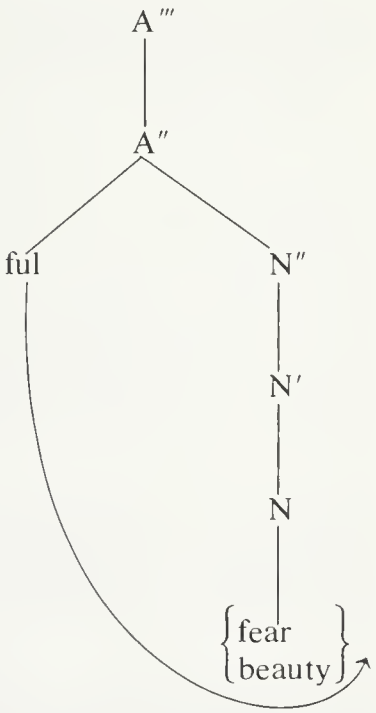


Now consider the implications of these changes for universal grammar. There are word formation rules for deriving almost every lexical category from almost every other lexical category. Is it a coincidence that the only types that have developed into phrase structure rules are those which involve deverbal constructions? Rule schema (9.1) claims that it is not, since it restricts the class of category-switching expansions in syntax to those in which V is on the righthand side. For example, (9.1) predicts that we would never find a historical change like the one from (9.34a) to (9.34b).

(9.34) a.



b.



If there were a syntactic structure (9.34b), we would expect it to have N″ modifiers such as adjectives and quantifiers, e.g. **great fearful* or **much beautiful*. No such category-switching constructions have come to my attention in which the head is other than V. By contrast, changes in the opposite direction, changing syntactic structure into lexical structure, do occur; for example, the adjective *full* developed into an affix that converts nouns into adjectives. Thus rule schema (9.1) explains why certain historical changes could have taken place and others could not have.

This completes our brief survey of deverbalizing rules. I consider them an extremely interesting class of rules, putting on a principled basis a sizable number of what previously had to be treated as altogether arbitrary phrase structure expansions. In particular, they seem to me to provide a framework for a richer syntax of complementation than has previously been considered, one that can stimulate a fresh approach to the crucial problem of constraints on extraction. Moreover, as we have just mentioned, they provide a previously unsuspected locus for historical change in syntax.

In turn, it is only because we adopted such a highly constrained theory of phrase structure as the Uniform Three-Level Hypothesis that the class of deverbalizing rules has emerged as a unified whole. The fact that the exceptions to the main phrase structure rule schema can be so clearly and systematically delimited is evidence that we chose the correct set of constraints in the first place. Thus the discussion in this chapter serves indirectly as an argument for the main thesis of this book, the formalization of the $\bar{X}$ Convention in terms of syntactic features and primes.[6]

[6] After this book went to press, it came to my attention that there are also free relatives dominated by PP, for example the subordinate clauses in *She put it where I wanted it* and *I'll go whenever you do*. This implies a further deverbalizing rule:

(i) $P''' \rightarrow$ wh - V'''

10: More General Implications

10.1. Explanatory Adequacy

Chapter 3 formulated the $\bar{X}$ Convention as a theory of grammatical categories. This theory claims that the possible lexical categories are defined by a set of syntactic distinctive features, and that the syntactic categories are defined in turn as the supercategories of the lexical categories by way of the bar (or prime) notation. The theory was made more explicit by proposing the syntactic features *Subj, Obj, Comp,* and *Det* to describe the lexical categories of English; finally, the theory of syntactic categories was elaborated into a theory of phrase structure rules, the Uniform Three-Level Hypothesis.

The Uniform Three-Level Hypothesis claims that every lexical category X has exactly three supercategories X′, X″, and the major phrasal category X‴, and that all modifiers are either major phrasal categories or grammatical formatives. This claim was formalized as the rule schema (3.9),

(3.9) *Uniform Three-Level Hypothesis*
$X^n \rightarrow (C_1) \ldots (C_j) - X^{n-1} - (C_{j+1}) \ldots (C_k)$, where $1 \leq n \leq 3$, and for all C_i, either $C_i = Y'''$ for some lexical category Y, or C_i is a specified grammatical formative.

We further claimed that there are only two kinds of exceptions to (3.9), coordinate conjunction and deverbalized constructions, each constrained by a characteristic phrase structure rule schema.

(3.32) *Coordinate Conjunction Schema*
$X^i \rightarrow X^i - (\text{Conj} - X^i)^*$

(9.1) *Deverbalizing Rule Schema*
$X^i \rightarrow \text{af} - V^i$

We claim, then, that all possible phrase structure rules of natural language are described by these three schemata.

Starting in section 3.4 we embarked on a tour of English, concerned with showing that its phrase structure is described by this theory. We took particular care to show that many previous analyses of phrase structure which are not instances of the three

schemata are either empirically inadequate or replaceable by descriptively equivalent rules which do conform to the present theory. It is now worth stepping back to try to assess the value of this mass of particular analyses as justification of the theory.

There are three separate issues that must be dealt with in evaluating a theory. First, does it generate enough grammars that the correct grammars are among them? Second, does it constrain the class of grammars sufficiently to exclude various kinds of conceivable but nonexistent rules? Third, does it characterize the notion "linguistically significant generalization" in such a way that the language learner can arrive at the correct grammar on the basis of fragmentary evidence?

The first issue, that of the descriptive adequacy of the theory, has been dealt with at length in this study, through detailed investigation of the grammar of English. At those points where there was some question as to the need to expand the class of grammars described by the theory, we have tried to suggest the most feasible ways of doing this. In order to evaluate the descriptive adequacy of the theory more fully, one would have to look more seriously at other languages. A notable lack in the theory, for example, is a mechanism to deal with free word order languages; the feature "transportable" we introduced for the free order of English adverbs in V″ and V‴ is doubtless only a makeshift.

As for the issue of constraining the class of grammars, this theory is to my knowledge the most highly constrained theory of phrase structure rules to date (other than the trivialized generative semantics base, which contains only the two rules S → V – NP – (NP) and NP → N – (S)). Inasmuch as it excludes many well-known analyses in the literature, forcing us to develop and defend alternatives, the constraints it imposes are hardly self-evident or gratuitous. Of course, it is entirely possible that ways can be found to constrain the theory more. Having been up to now preoccupied with showing that it is not too constrained, I have not given much consideration to this problem.

The third issue, that of characterizing the notion "linguistically significant generalization", is a knottier one. In formal terms, what is desired is an evaluation measure within the theory, which determines which of two given grammars is more highly valued, i.e. more "general". The type of linguistic generalization we have been concerned with in this study is the cross-category generalization: a term of a rule is simpler, i.e. more general, if it is satisfied by a set of syntactic categories expressible by fewer syntactic features. For example, a term satisfied by any of the categories N″, Q″, A″, Adv″ would be more general than a term satisfied by any of the categories V‴, V″, N″, and P‴, because the former could be expressed with two features, whereas the latter would take four. It would also be simpler than a term satisfied by N″ alone, which would take three features. This sort of feature-counting is precisely parallel to that in phonology.

Though the notion of embodying the evaluation measure in a feature-counting mechanism is perfectly straightforward when dealing with individual terms of rules, as demonstrated by our success in stating generalized grammatical relations, it has proven a great deal more problematic in dealing with entire rules. As shown in the discussion

of section 4.5, the generalizations among phrase structure rules are characteristically only partial, giving rise to several sorts of situations in which the notational conventions traditionally incorporated in the evaluation measure fail to appropriately reduce the number of features in the rule. Since the evaluation measure as a simple feature-counting device depends on the existence of appropriate abbreviatory conventions, there are only two solutions to the problem: either develop new abbreviatory conventions, or develop an evaluation measure that does something more than just count features. We have shown that one of these alternatives is necessary in the theory of the phrase structure component.

It may not be out of place here to point out that Jackendoff (1975a), discussing the nature of lexical generalizations, shows that they too resist analysis in terms of a feature-counting evaluation measure. A different sort of evaluation measure is proposed there, better adapted for the partial generalization endemic to the lexicon: instead of attempting to collapse related lexical entries into a single more general entry, it directly measures the redundancy between fully specified entries. A parallel approach in the phrase structure component would use the feature and prime system not to collapse rules but to enable the evaluation measure to judge, given two or more rules with all categories fully specified, how much alike they are, i.e. how easy it is to learn one given the other. Whether such an approach would be feasible I leave for future research: it has been shown, in any event, that the usual approach is inadequate.

Of course, the evaluation measure (at least as presently conceived of) would value most highly a grammar in which all categories of the same level had exactly the same phrase structure expansion, aside from the head. Such a grammar seems hardly likely, and there are probably many substantive universals which prevent this ideal from being achieved. For example, we have good reason to expect M''' only under V''' and Art''' only under N''', and these very specific restrictions are undoubtedly part of universal grammar. For a more controversial case, observe that the grammar of English would be simplified if nouns and adjectives came to be followed directly by NP, as verbs and prepositions are. Yet this change seems highly unlikely. One possible explanation is that universal grammar places such general restrictions on phrase structure expansions as prohibiting surface objects in N' and A', prohibiting subjects in A''' and P''', and generally excluding all X' modifiers from the minor lexical categories. (This would amount to the claim that my choice of feature names has theoretical significance.) Whether or not these specific restrictions are correct, they are indicative of the way in which substantive universals could in principle prevent the phrase structure component from achieving maximal generality.

Having discussed the implications of this study for the syntactic component, we will conclude by turning to its impact on semantics.

10.2. The Challenge for Semantics

Throughout this study I have made strong assumptions about the capabilities of the semantic component of the grammar. Though many readers will no doubt see these

assumptions as "sweeping difficulties under the rug", my sense of the situation is that, given an apparently arbitrary restriction on syntax, the account of it that puts the least burden on the language learner is that it follows from independently motivated rules and restrictions on semantic interpretation; it can then be claimed that knowing the meaning of the construction in question automatically results in the application of the observed constraint. If instead the constraint has to be added to the syntax, either through universal grammar, or worse yet, in the language-particular syntax, no real explanation has been provided. It seems to me that the success of this methodology has been amply proven in Jackendoff (1972; 1974b; 1975c; 1976) and need not be further defended here.

In the present study, however, I have been somewhat lax about actually proposing semantic solutions to syntactic constraints. In a few cases, such as the binding of restrictive relative clauses to their articles and the interpretation of relative pronouns and the PRO in degree clauses, I have given explicit solutions. But the cases left open, such as the Partitive Constraint, pseudopartitive determiner constraints, the binding of degree clauses to their degree words, and the nonrecursion of the complementizer node, have been far more numerous. Thus the present study leaves unsolved a great number of particular problems, some of which are demonstrably semantic, some of which have been called semantic admittedly on the basis of little more than hope.

More generally, we have explored a wide range of syntactic constructions whose semantics have never been seriously studied. And though for the most part we have mentioned the semantics of these many constructions only peripherally, it is clear that they provide a rich field of inquiry for natural logic. For example, consider the range of modifiers within the NP. At the N′ level there are functional arguments, which are well understood by logic. At the N″ level there are restrictive modifiers (generally treated in logic as conjoined predications or restrictions on quantifiers), the quantifiers *many* and *few* (not among the quantifiers traditionally treated by logic), and numerals, measure phrases, and pseudopartitive group nouns (outside the province of traditional logic). In N‴ there are subjects (again functional arguments), genitive NPs of intrinsic connection (not treated by logic), demonstratives and N‴ quantifiers (only partially within traditional logic), appositives (again treated as conjoined predication), and gerundive PPs (semantics unknown). The PP, which has in the past been treated with equal disdain by syntax and semantics, has proved to have an extremely rich syntax, including a characteristic system of quantification and degree modification; this syntax implies the existence of a corresponding semantics, up to now totally unexplored. In order for the semantic component of language to achieve even observational adequacy, it is clear that there is a vast range of phenomena yet to account for.

But of course the syntactic investigation here was not directed toward merely the complexity of phrase structure but toward uncovering the richness of cross-categorial generalization in phrase structure and in grammatical relations. Since grammatical relations, as the environments over which selectional restrictions are defined, are in fact the structural descriptions of rules of semantic interpretation, cross-categorial

syntactic generalization automatically implies cross-categorial semantic generalization as well. Each of the major lexical categories gives rise to functional arguments, to measure and degree modification, to quantification, and to restrictive and nonrestrictive modification. Many of the minor lexical categories take some of these modifiers too. In one particularly extreme example pointed out in chapter 6, it was observed that the degree expression *almost* can serve as a modifier to seven different categories. Such cross-categorial semantic generalizations have never to my knowledge played a role in any semantic theory.

As I see it, such generalizations have profound consequences for semantic theory. Under the usual assumption that the aspects of meaning represented in deep structure are compositionally determined from the structure, each X''' in a sentence must be provided with an interpretation into which all of the modifiers of that X''' have been integrated. The fact that the same grammatical relations appear in different categories suggests that the interpretations of the categories have internal structure that is parallel in important ways. On the other hand, the differences in possibilities among the categories argue that the similarities are not to be captured by reducing all categories semantically to predicates, as the traditional logical program would have it.

The argument I have just given parallels the syntactic arguments of chapter 2, aimed at showing that various categories are syntactically distinct and not derived one from the other. And I think the conclusions are parallel as well: alongside a lexicalist syntax there must be a lexicalist semantics, which takes seriously the categorial pluralism of language. Such a semantic theory would take as a basic assumption that language has such a variety of syntactic devices at its disposal in part because of the variety of semantic entities and relationships it has to express. And where syntactic relationships are parallel or nonparallel across categories, that may often be a reflection of parallelisms or nonparallelisms in semantic structure.

To be sure, we should not expect point-by-point parallelism between syntax and semantics: we have seen cases both where a single syntactic position served two semantic purposes and where the same thing was expressed with two different syntactic constructions. But to me it is inconceivable that something as syntactically rich as natural language serves the purpose of expressing something as ontologically impoverished as any present-day form of logic.

Given the relationships between syntax and semantics that have been discussed in this study, a plausible initial hypothesis emerges as to the proper strength of their parallelism in universal grammar. A primary claim is that each major phrasal category (X''') corresponds to a semantic (or ontological) category. In the simplest cases, NPs correspond to a semantic category which essentially picks out things; Ss to a category which picks out events and states of affairs; APs to properties (*not* predicates); and PPs to places (i.e. spatial relationships to things). Each of these major semantic categories has an internal structure whose possibilities are provided by universal grammar; the constituents of these internal structures are the interpretations of the grammatical relations possible within the major phrasal categories.

What one might guess not to be universal about the correspondence is exactly how the phrases expressing internal constituents of the major semantic categories are distributed syntactically within the major phrasal categories. For instance, we might expect to find modifiers of the same semantic type occurring within X″ in one case and within X‴ in another. We have found at least two such examples in English alone: the occurrence of semantic quantifiers at both the N‴ and N″ levels, and the alternation of nonrestrictive relative clauses in N‴ with nonrestrictive adjective phrases in N″. It would not be surprising to find much greater variability across languages in this respect.

On the other hand, one might expect the correspondences within each X‴ not to be totally arbitrary. For example, nonrestrictive modifiers probably cannot be more deeply embedded in X‴ than restrictive modifiers, and functional arguments other than the subject quite possibly must occur in X′. Among the tasks of cross-linguistic investigation would be to discover exactly how constrained these relationships are, and to specify how extensive cross-categorial generalizations are both in universal grammar and, where universal grammar leaves the question open, within particular languages.

Such a program of lexicalist semantics, then, would conceive of the hierarchical aspect of rules of phrase structure as being to a significant extent dictated by semantic considerations. However great or small the relationship actually turns out to be, the result is of interest. For to whatever extent the form of the syntax is not determined by semantics, this is evidence for Chomsky's (1975) contention that the language faculty is an autonomous mental organ whose structure is as different from other mental organs as the heart is different from the liver. On the other hand, the closer that syntactic form is related to semantic form in particular or universal grammar, the more it is possible to use purely syntactic arguments as a basis for hypotheses about semantic structure.

As may be gathered from the preceding remarks, I believe that the relationship between syntax and semantics is a strong one, and that it can and should be exploited in order to counteract the irresponsibility toward syntactic form that is characteristic of most present-day work in logic. Nor should this relationship be used, as many linguists have used it, to bring evidence from a mathematically conceived formal semantics to bear on syntax. Rather, the syntax itself provides precise and highly differentiated data; a lexicalist semantics would use this evidence to shed light on the much murkier question of how natural logic is structured. I consider Jackendoff (1975c; 1976) as examples of such an approach, using syntactic paradigms to motivate novel semantic analyses of some generality.

Since semantic representation is presumably closely related to cognitive structure, such a program of research has a further consequence. To the extent that a lexicalist semantics proves successful, the study of language provides an unparalleled tool for investigating not just a single organ of the human mind, but perhaps some of the most fundamental principles of mental organization. The potential implications for psychology are vast. It is hoped that the present study provides some impetus toward their discovery.

References

Akmajian, A. (1973) "The Role of Focus in the Interpretation of Anaphoric Expressions," in S. Anderson and P. Kiparsky, eds. (1973), 215–225.

Akmajian, A. (1975) "More Evidence for the NP Cycle," *Linguistic Inquiry* 6.1, 115–129.

Akmajian, A. (1977) "The Complement Structure of Perception Verbs in an Autonomous Syntactic Framework," in A. Akmajian, P. Culicover, and T. Wasow, eds. (1977).

Akmajian, A., P. Culicover, and T. Wasow, eds. (1977) *Studies in Formal Syntax*, Academic Press, New York.

Akmajian, A. and A. Lehrer (1976) "NP-like Quantifiers and the Problem of Determining the Head of an NP," *Linguistic Analysis* 2.4, 395–413.

Akmajian, A. and T. Wasow (1975) "The Constituent Structure of VP and AUX and the Position of the Verb BE," *Linguistic Analysis* 1.3, 205–245.

Anderson, S. and P. Kiparsky, eds. (1973) *A Festschrift for Morris Halle*, Holt, Rinehart, and Winston, New York.

Aronoff, M. (1974) *Word-Structure*, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Aronoff, M. (1976) *Word Formation in Generative Grammar*, Linguistic Inquiry Monograph No. 1, MIT Press, Cambridge, Massachusetts.

Bach, E. and R. T. Harms, eds. (1968) *Universals of Linguistic Theory*, Holt, Rinehart, and Winston, New York.

Banfield, A. (1973) "Narrative Style and the Grammar of Direct and Indirect Speech," *Foundations of Language* 10, 1–39.

Berman, A. (1974) *Adjectives and Adjective Complement Constructions*, Doctoral dissertation, Harvard University, Cambridge, Massachusetts; published as Formal Linguistics Report NSF-29, Susumu Kuno, principal investigator, Harvard University Department of Linguistics.

Bierwisch, M. and K. E. Heidolph, eds. (1970) *Progress in Linguistics*, Mouton, The Hague.

Bowers, J. S. (1968a) "Some Adjectival Nominalizations in English," unpublished mimeograph, MIT, Cambridge, Massachusetts; *Lingua* 37, 341–361 (1975).

Bowers, J. S. (1968b) "Adjectives and Adverbs in English," Indiana University Linguistics Club, Bloomington; *Foundations of Language* 13, 529–562 (1975).

Brame, M. (1968) "A New Analysis of the Relative Clause: Evidence for an Interpretive Theory," unpublished paper, MIT, Cambridge, Massachusetts.

Brame, M. (1975) "On the Abstractness of Syntactic Structure: The VP Controversy," *Linguistic Analysis* 1.2, 191–203.

Bresnan, J. W. (1970) "On Complementizers: Towards a Syntactic Theory of Complement Types," *Foundations of Language* 6, 297–321.

Bresnan, J. W. (1971) "Sentence Stress and Syntactic Transformations," *Language* 47, 257–281.

Bresnan, J. W. (1972) *The Theory of Complementation in English Syntax*, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Bresnan, J. W. (1973) "Syntax of the Comparative Clause Construction in English," *Linguistic Inquiry* 4.3, 275–344.

Bresnan, J. W. (1974) "The Position of Certain Clause-Particles in Phrase Structure," *Linguistic Inquiry* 5.4, 614–619

Bresnan, J. W. (1976a) "On the Form and Interpretation of Syntactic Transformations," *Linguistic Inquiry* 7.1, 3–40.

Bresnan, J. W. (1976b) "Toward a Realistic Theory of Transformational Grammar," paper presented at AT&T conference on language at MIT, March 1976.

Chomsky, N. (1955) *The Logical Structure of Linguistic Theory*, Plenum, New York (1975).

Chomsky, N. (1957) *Syntactic Structures*, Mouton, The Hague.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1970a) "Remarks on Nominalization," in R. Jacobs and P. Rosenbaum, eds. (1970), 184–221; in N. Chomsky (1972), 11–61.

Chomsky, N. (1970b) "Deep Structure, Surface Structure, and Semantic Interpretation," in D. Steinberg and L. Jakobovits, eds. (1971), 183–216; in N. Chomsky (1972), 62–119.

Chomsky, N. (1972) *Studies on Semantics in Generative Grammar*, Mouton, The Hague.

Chomsky, N. (1973) "Conditions on Transformations," in S. Anderson and P. Kiparsky, eds. (1973), 232–286.

Chomsky, N. (1975) *Reflections on Language*, Pantheon, New York.

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, Harper & Row, New York.

Dean, J. (1966) "Determiners and Relative Clauses," unpublished mimeograph, MIT, Cambridge, Massachusetts.

Dineen, F., ed. (1966) 19th Monograph on Languages and Linguistics, Georgetown University Press, Washington, D.C.

Dixon, R. M. W., ed. (1975) *Grammatical Categories*, Australian Institute of Aboriginal Studies.

Dougherty, R. C. (1968) *A Transformational Grammar of Coordinate Conjoined Structures*, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Dougherty, R. C. (1970–1971) "A Grammar of Coordinate Conjoined Structures," *Language* 46.4, 850–898, and 47.2, 298–339.

Emonds, J. E. (1970) *Root and Structure-Preserving Transformations*, Doctoral dissertation, MIT, Cambridge, Massachusetts; published by Indiana University Linguistics Club, Bloomington, Indiana

Emonds, J. E. (1972a) "Evidence that Indirect Object Movement Is a Structure-Preserving Rule," *Foundations of Language* 8, 546–561.

Emonds, J. E. (1972b) "A Reformulation of Certain Syntactic Transformations," in P. S. Peters, ed. (1972).

Emonds, J. E. (1976) *A Transformational Approach to English Syntax*, Academic Press, New York.

Faraci, R. (1974) *Aspects of the Grammar of Infinitives and* For-*Phrases*, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Fiengo, R. (1974) *Semantic Conditions on Surface Structure*, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Fillmore, C. (1968) "A Case for Case," in E. Bach and R. T. Harms, eds. (1968), 1–88.

Fillmore, C. and D. T. Langendoen, eds. (1971) *Studies in Linguistic Semantics* Holt, Rinehart, and Winston, New York.

Fodor, J. A. and J. J. Katz, eds. (1964) *The Structure of Language*, Prentice-Hall, Englewood Cliffs, New Jersey.

Freidin, R. (1975) "The Analysis of Passives," *Language* 51.2, 384–405.

Greenberg, J. (1963) "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements," in J. Greenberg, ed. (1963).

Greenberg, J., ed. (1963) *Universals of Language*, MIT Press, Cambridge, Massachusetts.

Gruber, J. S. (1965) *Studies in Lexical Relations*, Doctoral dissertation, MIT, Cambridge, Massachusetts; published by Indiana University Linguistics Club, Bloomington, Indiana.

Hale, K. (1975) "The Adjoined Relative Clause in Australia," in R. M. W. Dixon, ed. (1975).

Halle, M. (1973) "Prolegomena to a Theory of Word Formation," *Linguistic Inquiry* 4.1, 3–16.

Harris, Z. (1946) "From Morpheme to Utterance," *Language* 22, 161–183.

Harris, Z. (1951) *Methods in Structural Linguistics*, University of Chicago Press, Chicago, Illinois.

Higgins, F. R. (1973) "On J. Emonds's Analysis of Extraposition," in J. Kimball, ed. (1973).

Horn, G. (1975) "On the Nonsentential Nature of the Poss-Ing Construction," *Linguistic Analysis* 1.4, 333–388.

Hornstein, N. (1975) "S and the $\bar{X}$ Convention," *Montreal Working Papers in Linguistics*, Vol. 4, McGill University Department of Linguistics, Montreal.

Horvath, J. (1976) "Focus in Hungarian and the $\bar{X}$ Notation," *Linguistic Analysis* 2.2, 175–197.

Jackendoff, R. (1968) "Quantifiers in English," *Foundations of Language* 4.4, 422–442.

Jackendoff, R. (1969a) "An Interpretive Theory of Negation," *Foundations of Language* 5.2, 218–241.

Jackendoff, R. (1969b) "Les constructions possessives en anglais," *Langages* 14, 7–27.

Jackendoff, R. (1971) "Gapping and Related Rules," *Linguistic Inquiry* 2.1, 21–36.

Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, Massachusetts.

Jackendoff, R. (1973) "The Base Rules for Prepositional Phrases," in S. Anderson and P. Kiparsky, eds. (1973), 345–356.

Jackendoff, R. (1974a) "Introduction to the $\bar{X}$ Convention," published by Indiana University Linguistics Club, Bloomington, Indiana.

Jackendoff, R. (1974b) "A Deep Structure Projection Rule," *Linguistic Inquiry* 5.4, 481–506.

Jackendoff, R. (1975a) "Morphological and Semantic Regularities in the Lexicon," *Language* 51.3, 639–671.

Jackendoff, R. (1975b) "*Tough* and the Trace Theory of Movement Rules," *Linguistic Inquiry* 6.3, 437–446.

Jackendoff, R. (1975c) "On Belief-Contexts," *Linguistic Inquiry* 6.1, 53–93.

Jackendoff, R. (1976) "Toward an Explanatory Semantic Representation," *Linguistic Inquiry* 7.1, 89–150.

Jackendoff, R. and P. Culicover (1971) "A Reanalysis of Dative Movements," *Foundations of Language* 7.3, 397–412.

Jacobs, R. and P. Rosenbaum, eds. (1970) *Readings in English Transformational Grammar*, Ginn, Waltham, Massachusetts.

Jespersen, O. (1909–1949) *A Modern English Grammar*, George Allen & Unwin, London.

Katz, J. J. (1966) *The Philosophy of Language*, Harper & Row, New York.

Katz, J. J. (1972) *Semantic Theory*, Harper & Row, New York.

Keyser, S. J. (1968) Review of S. Jacobson, *Adverbial Positions in English*, *Language* 44.2, 357–373.

Kimball, J., ed. (1973) *Syntax and Semantics*, Vol. 2, Academic Press, New York.

Kimball, J., ed. (1975) *Syntax and Semantics*, Vol. 4, Academic Press, New York.

Klima, E. S. (1964) "Negation in English," in J. A. Fodor and J. J. Katz, eds. (1964), 246–323.

Klima, E. S. (1965) *Studies in Diachronic Syntax*, unpublished Doctoral dissertation, Harvard University, Cambridge, Massachusetts.

Kuroda, S.-Y. (1969) "English Relativization and Certain Related Problems," in D. Reibel and S. Schane, eds. (1969), 264–287.
Lakoff, G. (1970a) "Global Rules," *Language* 46.3, 627–639.
Lakoff, G. (1970b) "Repartee, or a Reply to 'Negation, Conjunction, and Quantifiers'," *Foundations of Language* 6.3, 389–422.
Lakoff, G. (1971a) "On Generative Semantics," in D. Steinberg and L. Jakobovits, eds. (1971), 232–296.
Lakoff, G. (1971b) *On Syntactic Irregularity,* Holt, Rinehart and Winston, New York.
Lakoff, G. and J. R. Ross (1966) "A Criterion for Verb Phrase Constituency," in *Mathematical Linguistics and Automatic Translation,* Report NSF-17, Harvard Computation Laboratory, Harvard University, Cambridge, Massachusetts.
Lasnik, H. and R. Fiengo (1974) "Complement Object Deletion," *Linguistic Inquiry* 5.4, 535–572.
Lees, R. B. (1961) "Grammatical Analysis of the English Comparative Construction," *Word* 17, 171–185; in D. Reibel and S. Schane, eds. (1969).
Lightfoot, D. (1974) "The Diachronic Analysis of English Modals," *Montreal Working Papers in Linguistics* Vol. 3, McGill University Department of Linguistics, Montreal, 115–145.
McCawley, J. D. (1971) "Where do Noun Phrases Come From?" in D. Steinberg and L. Jakobovits (1971), 217–231.
McCawley, J. D. (1975) "The Category Status of English Modals," *Foundations of Language* 12, 597–601.
Milner, J.-C. (1973) "Comparatives et relatives," in J.-C. Milner, ed. (1973).
Milner, J.-C., ed. (1973) *Arguments Linguistiques,* Maison Mâme, Paris.
O'Neil, W. (1976) "Clause Adjunction in Old English," ms., to appear in *General Linguistics.*
Perlmutter, D. (1970) "On the Article in English," in M. Bierwisch and K. E. Heidolph, eds. (1970).
Peters, P. S., ed. (1972) *Goals of Linguistic Theory,* Prentice-Hall, Englewood Cliffs, New Jersey.
Postal, P. M. (1966) "On So-Called 'Pronouns' in English," in F. Dinneen, ed. (1966); in D. Reibel and S. Schane, eds. (1969), 201–224.
Postal, P. M. (1971) *Crossover Phenomena,* Holt, Rinehart and Winston, New York.
Postal, P. M. (1974) "On Certain Ambiguities," *Linguistic Inquiry* 5.3, 367–424.
Rardin, R. (1975) *Studies in Derivational Morphology,* unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.
Reed, A. (1974) *The Structure of English Relative Clauses,* unpublished dissertation, Brandeis University, Waltham, Massachusetts.
Reibel, D. and S. Schane, eds. (1969) *Modern Studies in English,* Prentice-Hall, Englewood Cliffs, New Jersey.
Rosenbaum, P. (1967) *The Grammar of English Predicate Complement Constructions,* MIT Press, Cambridge, Massachusetts.
Ross, J. R. (1967) *Constraints on Variables in Syntax,* Doctoral dissertation, MIT, Cambridge, Massachusetts; published by Indiana University Linguistics Club, Bloomington, Indiana.
Ross, J. R. (1969) "Adjectives as Noun Phrases," in D. Reibel and S. Schane, eds. (1969), 352–360.
Schachter, P. (1976) "A Nontransformational Account of Gerundive Nominals in English," *Linguistic Inquiry* 7.2, 205–241.
Selkirk, E. O. (1970) "On the Determiner Systems of Noun Phrase and Adjective Phrase," unpublished paper, MIT, Cambridge, Massachusetts.

Selkirk, E. O. (1972) *The Phrase Phonology of English and French,* unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Selkirk, E. O. (1974) "French Liaison and the $\bar{X}$ Convention," *Linguistic Inquiry* 5.4, 573–590.

Selkirk, E. O. (1977) "Some Remarks on Noun Phrase Structure," in A. Akmajian, P. Culicover, and T. Wasow, eds. (1977).

Siegel, D. (1974) *Topics in English Morphology,* unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Smith, C. (1964) "Determiners and Relative Clauses in a Generative Grammar of English," *Language* 40.1, 37–52; in D. Reibel and S. Schane, eds. (1969), 247–263.

Sommerstein, A. (1972) "On the So-Called Definite Article in English," *Linguistic Inquiry* 3.2, 197–210.

Steinberg, D. and L. Jakobovits, eds. (1971) *Semantics: An Interdisciplinary Reader,* Cambridge University Press, New York.

Stockwell, R., P. Schachter, and B. Partee (1973) *The Major Syntactic Structures of English,* Holt, Rinehart, and Winston, New York.

Thompson, S. A. (1971) "The Deep Structure of Relative Clauses," in C. Fillmore and D. T. Langendoen, eds. (1971), 76–96.

Van Riemsdijk, H. (1973) "The Dutch P-Soup," unpublished mimeograph, University of Amsterdam.

Van Riemsdijk, H. (1976) "Extraction from Prepositional Phrases and the Head Constraint," unpublished mimeograph, University of Amsterdam.

Vergnaud, J.-R. (1974) *French Relative Clauses,* unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Wald, J. D. (1976) *Stuff and Words: A Semantic and Linguistic Analysis of Non-Singular Reference,* unpublished Doctoral dissertation, Brandeis University, Waltham, Massachusetts.

Wasow, T. (1971) *Anaphoric Relations in English,* unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Wasow, T. (1975) "A Note on Empty Nodes," *Foundations of Language* 12, 603–607.

Wasow, T. (1976) *Anaphora in Generative Grammar,* Story-Scientia, Ghent.

Wasow, T. (1977) "Transformations and the Lexicon," in A. Akmajian, P. Culicover, and T. Wasow, eds. (1977).

Wasow, T. and T. Roeper "On the Subject of Gerunds," *Foundations of Language* 8.1, 44–61.

Williams, E. (1975) "Small Clauses in English," in J. Kimball, ed. (1975), 249–273.

Other recent titles in the series Linguistic Inquiry Monographs

Recent Transformational Studies in European Languages
edited by Samuel Jay Keyser

This is the third publication in the Linguistic Inquiry Monograph Series, which presents new and original research beyond the scope of the article format.

Contents: "*On a Morphologically Governed Vowel Alternation in French*" by François C. Dell, CNRS, Paris, and Elisabeth O. Selkirk, University of Massachusetts at Amherst; "*Why Subject Sentences Don't Exist*" by Jan Koster, University of Amsterdam; "*Trace Theory and French Syntax*" by Jean-Yves Pollock, Université de Paris XII et VII; "*A Restructuring Rule in Italian Syntax*" by Luigi Rizzi, Scuola Normale Superiore, Pisa; "*Result Clauses and Conditions on Rules*" by Alain Rouveret, Université de Paris VIII; "*On the Diagnosis of Wh Movement*" by Henk van Riemsdijk, University of Amsterdam.

Studies in Abstract Phonology
by Edmund Gussmann

The issue of abstractness is currently one of the central questions in the ongoing research into the theoretical foundations of modern phonology. Linguists are divided on the degree of abstractness admissible in phonological representation. While some would impose far-reaching restrictions on admissible abstractness, others have felt that such restrictions are unrealistic. This monograph offers extensive arguments for admitting a high degree of abstractness in phonology.

The empirical material on which the monograph is based is the phonology of modern Polish. A substantial portion of the phonological rules of the language are studied and their theoretical implications are investigated in detail. Although there exists a rich literature on Polish phonology, very little of it has been written with clear theoretical issues in mind, and practically none of it is in English. The results obtained in the monograph, especially with regard to the vowel-zero alternations, hold not only for Polish but with some straightforward modifications can be shown to hold for all other Slavic languages as well.

An Encyclopedia of AUX:
A Study in Cross-Linguistic Equivalence
by Susan Steele,
with Adrian Akmajian, Richard Demers, Eloise Jelinek, Chisato Kitagawa, Richard Oehrle, and Thomas Wasow

This significant contribution to the generative theory of syntax confronts a fundamental question inherent in the study of universals of grammar: How can claims of universality be empirically verified? This book puts the question of the similarity of categories across the grammars of particular languages on a substantive empirical basis by providing a principled method of cross-linguistic comparison, a method based on equivalence classes of language-particular properties relative to a universally applicable definition. While the critical issue is general to a consideration of any syntactic category, the framework proposed is applied specifically to a category called AUX. A considerable body of evidence is introduced for both the framework and the category through the detailed examination of a number of languages—Lummi, Egyptian Arabic, Japanese, Luiseño, German, and English.

The MIT Press
Massachusetts Institute of Technology
Cambridge, Massachusetts 02142