A Unified Theory of Syntactic Categories

Studies in Generative Grammar

The goal of this series is to publish those texts that are representative of recent advances in the theory of formal grammar.
Too many studies do not reach the public they deserve because of the depth and detail that make them unsuitable for publication in article form.
We hope that the present series will make these studies available to a wider audience than has hitherto been possible.

Other books in this series:

1. Wim Zonneveld
   A Formal Theory of Exceptions in Generative Phonology
2. Pieter Muysken
   Syntactic Developments in the Verb Phrase of Ecuadorian Quechua
3. Geert Booij
   Dutch Morphology
4. Henk van Riemsdijk
   A Case Study in Syntactic Markedness
5. Jan Koster
   Locality Principles in Syntax
6. Pieter Muysken (ed.)
   Generative Studies on Creole Languages
7. Anneke Neijt
   Gapping
8. Christer Platzack
   The Semantic Interpretation of Aspect and Aktionsarten
9. Noam Chomsky
   Lectures on Government and Binding
10. Robert May and Jan Koster (eds.)
    Levels of Syntactic Representation
11. Luigi Rizzi
    Issues in Italian Syntax
12. Osvaldo Jaeggli
    Topics in Romance Syntax
13. Hagit Borer
    Parametric Syntax
14. Denis Bouchard
    On the Content of Empty Categories
15. Hilda Koopman
    The Syntax of Verbs
16. Richard S. Kayne
    Connectedness and Binary Branching
17. Jerzy Rubach
    Cyclic and Lexical Phonology
18. Sergio Scalise
    Generative Morphology

Joseph E. Emonds

A Unified Theory of Syntactic Categories

1985
FORIS PUBLICATIONS
Dordrecht - Holland/Cinnaminson - U.S.A.

Published by:
Foris Publications Holland
P.O. Box 509
3300 AM Dordrecht, The Netherlands

Sole distributor for the U.S.A. and Canada:
Foris Publications U.S.A.
P.O. Box C-50
Cinnaminson N.J. 08077
U.S.A.

CIP-DATA

Emonds, Joseph E.

A Unified Theory of Syntactic Categories / Joseph E. Emonds. - Dordrecht [etc.] : Foris. - (Studies in Generative Grammar ; 19)
With ref.
ISBN 90-6765-091-9 bound
ISBN 90-6765-092-7 paper
SISO 837.4 UDC 801.56
Subject heading: syntax ; generative grammar.

ISBN 90 6765 091 9 (Bound)
ISBN 90 6765 092 7 (Paper)

Printed in the Netherlands by ICG Printing, Dordrecht.

This book is dedicated to my friends
Mitsou Ronat, Judith McA'Nulty, and Alfredo Hurtado

It would be painful and perhaps pointless to bring out here the many ways in which I feel I benefited both personally and professionally from knowing these linguists. It is a great sorrow to me to think that these three are gone. I wish each of them could read this book.

Acknowledgments

My primary intellectual debt in writing this book is to all those recent authors in generative grammar who collectively continue to create linguistic science. Foremost among them, of course, is Noam Chomsky.

Among these, I would also like to single out some colleagues who have, in the four years I have been writing this book, repeatedly encouraged me in concrete ways to continue to develop my research. Besides those more directly involved who are mentioned below, these include Ann Banfield, Hagit Borer, Margot Griffin, Randall Hendrick, Kazuko Inoue, Masaru Kajita, Jean-Claude Milner, Anne Rochette, Thomas Roeper, and Wendy Wilkins. Indebtedness on a personal scale runs, of course, on a continuum, and there are many friends in linguistics whom I am indebted to besides those mentioned here.

In particular, Alfredo Hurtado invited me to work on the topic of control in Spanish gerunds (Ch. 2) and provided a forum where it could be fruitfully discussed (the Workshop on Spanish Syntax at Simon Fraser University). Shalom Lappin likewise arranged for a series of lectures at the University of Ottawa, which proved to be an opportunity for productive interchange with those that attended. Mark Baltin, Noam Chomsky, Judith McA'Nulty, Lisa Travis, and Hanna Walinska de Hackbeil have each been kind enough to go over a chapter carefully and provide me with very useful material and commentary.

This book has been written while I have been in the Linguistics Department at the University of Washington. It is certainly the members of this department who are most responsible for maintaining a tension-free, intellectually stimulating, personable, and progressive atmosphere in which teaching and writing can be a pleasure as well as labor. They do this in spite of the obstacles of woefully insufficient funding, particularly for graduate students.

All of my colleagues on the faculty, Michael Brame, Heles Contreras, Georgette Ioup, Ellen Kaisse, Frederick Newmeyer, and Sol Saporta have encouraged and assisted me in dozens of ways, directly and indirectly, both with respect to the content of this book and with respect to daily aspects of my job, each in ways characteristic of them alone. I thank each of them for their intellectual, personal, and political integrity. It is a rare privilege to be an academic who has for each of his departmental colleagues a deep personal and intellectual respect.

The students in the University of Washington department have made many important contributions to this project; it is they who have suffered through certainly erroneous preliminary and partly thought-out versions

of the ideas developed here, both in courses and in hours of discussions. Their patience, interest, criticism, and suggestions have been an indispensible means for improving my material. I am especially appreciative of those who have discussed many of the topics in this book while doctoral candidates, even though their treatments and mine diverge in countless ways. These include Abdulaziz Alshalan, Hanna Walinska de Hackbeil, Nobuko Hasegawa, Hajime Hoji, Mi-Jeung Jo, Anne Lobeck, Rosemarie Whitney, Young-Jae Yim, and Karen Zagona. Similarly, I profited from many discussions with Julia Horvath, Sharon Klein, and Vida Samiian, especially while on their U.C.L.A. doctoral committees.

Henk LaPorte and the staff employed by Foris have been conscientious throughout the production process; I especially have appreciated his enthusiastic and cooperative attitude about this project. Anita Hoffman and Rita Liang have been very perceptive and efficient in the difficult preparation of the bibliography and subject index, respectively, and I am grateful to each of them.

My son, Peter Emonds-Banfield, has (usually) been understanding and tolerant when my plea of "having to work," both on this book and at the university, has left him spending many hours on his own during which he might have rightfully expected my company. For his patience I am grateful.

I acknowledge a special debt to the departmental staff at the University of Washington, Anita Tabares-Laws, Peter Skaer, and Corinne Moore, for their cooperation and help throughout the arduous task of preparing the manuscript. Like the faculty, they have facilitated the administrative part of my job in every way, so as to make this research possible. Moreover, they have patiently typed and retyped the greatest part of the manuscript and seen to its being reproduced properly. Anita Tabares-Laws in particular has kept a whole range of departmental affairs running smoothly to the benefit of faculty research. Without her, I could not have chaired the department, written a book, and remained arguably sane.

My gratitude goes out especially to Frederick Newmeyer, Henk van Riemsdijk, and Rosemarie Whitney, who have each edited almost the entire manuscript, reading it both critically and sympathetically. They have given generously of their time, bringing to bear on this task intellectually specialized appreciations of the issues that I have greatly benefited from. Many of their suggestions have been incorporated or otherwise taken into account without specific acknowledgment. Colleagues like them have provided me during the last three years with the confidence I needed to follow through on my own ideas, whether with or against the current.

Table of contents

Introduction

The bar notation and the adjacency hypothesis

For centuries, the determination of the categories of syntax and the principles governing their combination have formed the study of grammar. Within the framework of generative grammar, the central morpheme categories "X" have been determined to be the noun, verb, adjective, and preposition (X = N, V, A, P). All phrasal categories used inside sentences are hypothesized to be "projections" of the lexical categories X^j (j = a small integer), where each X^j has one and only one X as its "lexical head." The centrality of the X and the restriction of phrasal categories to X^j is called the "bar notation" (Chomsky, 1970; Emonds, 1976, Ch. 1; Jackendoff, 1977).

Questions immediately arise: Do all languages realize the same inventory of categories? Do these categories combine in the same way in all languages? Are their principles of combination relatively simple? Are the combinatorial properties of these categories to some extent autonomous, rather than being completely derivable from other principles, such as the principles of the lexicon or of semantics? Finally, and I think centrally, what are the essential, defining properties of N, V, A, and P?

I will comment on the first four questions and then return to the final question concerning the nature of each bar notation category. Basing my conclusions both on other work and on the material presented in this book, I intend to justify strong affirmative claims for the first four questions above, with the following qualifications:

(i) Categorial Uniformity. The categories defined in terms of the bar notation, X^j and SPECIFIER (X) (cf. sections 1.2, 4.1, and 4.2 for specifiers), do not differ from language to language, but their subcategories which are realized in each language's syntax may vary.

(ii) Hierarchical Universality. The range of permitted hierarchical combinations of syntactic categories does not vary from language to language at the level of deep structure.[1] However, different restrictions on the linear order of constituents may be stated for this level.

1. The syntactic level of deep structure, at which Hierarchical Uniformity across languages holds, is understood here essentially as presented in Chomsky (1976). Such deep structures are related to partially "observable" syntactic surface structures by a highly restricted set of transformational operations, many of whose properties will be discussed in this book, particularly in Ch. 3.

Surface structures are "observable" in that the left-right sequence of morphemes in a well-

(iii) Syntactic Asymmetry. The principles determining the hierarchical combinations of syntactic categories at deep structure are simple, in the sense that no device as complex as a set of "phrase structure rules" is needed. However, these principles are not simply a "natural logic" of predicate-argument structure, as envisaged in the tradition of generative semantics; rather, the lexical categories X = N, V, A, P appear in them in limited non-symmetric ways, with a level of complexity somewhat akin to that envisaged by Gruber (1965). The asymmetries are determined by the defining characteristics of each category, to be discussed below.

(iv) Autonomy from the Lexicon. The theory of permitted deep structure categorial combinations is independent of the lexicon, in the sense that these combinations do not follow solely from the organization of the lexicon. The permitted deep structures are partly dependent, however, on the semantic component, i.e., on the principles which determine the semantic interpretation of predicate-argument structures, as I will argue in Ch. 1.

The answers that I will develop for the above questions – that a small inventory of grammatical categories organized according to universal and simple principles at an abstract deep structure level in a syntactically autonomous way suffices for elegant and empirically adequate descriptions of natural languages – clearly identify this work as Chomskyan (for example, according to the criteria in the Introduction of Newmeyer, 1980). If the academic field of linguistics were a science, the above statement, except for the reference to the subject matter (e.g., "grammatical," "syntactically"), would be a truism. However, since linguistics contains subfields in which a scientific approach has not been successful, as well as an overflow of practitioners within syntax and phonology who deny the scientific status of even these areas either explicitly or implicitly, it is appropriate to state at the outset that, as a generativist, I am attempting to construct a scientific theory of syntax, and by this attempt, I affirm that it is presently possible to do so.

Each of the above claims (i)–(iv) has controversial implications, several of which will be developed in detail in this book. Thus, Categorial Uniformity (i) implies that all languages, including verb-initial and verb-

Footnote—*Continued*

formed surface structure always corresponds to an acceptable string of pronounced morphemes in a language. However, surface structures also contain hierarchical structure, empty categories, and possibly indices which are not pronounced. Moreover, some surface structures are ill-formed by virtue of filters or restrictions that apply to the logical and phonological forms derived from them. Finally, some acceptable strings of morphemes may either be not directly generated by the grammar at all (so-called "derivatively generated" strings) or be generated only by virtue of an optional stylistic reordering of a surface structure string.

Since surface structures are not to be equated with "occurring strings of morphemes," for the above reasons, I will typically refer to them by the more abstract term of "s-structures", in conformity with much current usage.

final languages, should have a verb phrase (VP) distinct from the sentence (S) if and only if one language does (Ch. 3). Hierarchical Universality (ii) implies, for example, that if a certain syntactic category exists in a language, it must appear in the same deep structure position as it does in other languages. For example, verbal INFLECTION, being the category of the English modals, can be argued to be a deep structure sister to VP (Emonds, 1976, Ch. 6). By Hierarchical Universality INFLECTION can be a deep structure sister to VP in English if and only if it appears in the same position in French deep structures, since both languages clearly exhibit this syntactic category in their surface structure tense endings on verbs.

While (i) and (ii) are important and far from trivial claims about natural language, what distinguishes this book from other recent generative treatments of the base, or deep structures, are the claims of Syntactic Asymmetry (iii) and Autonomy from the Lexicon (iv).

For concreteness, I will first compare the approach of this book to that of two other generativists who have recently written extensively on the base component, Jackendoff (1977) and Stowell (1981), and then I will discuss my approach on its own terms. Both of these authors have relatively well-worked-out theories and share many of my assumptions. We agree not only on the existence of a transformational component that maps deep structures onto surface structures ("s-structures"), but also more or less on the range of constructions that are to be considered "transformationally derived" rather than "base-generated." We further agree on many aspects of the bar notation theory; i.e., that there are four lexical categories, that phrasal categories are projections of lexical categories, and that other rules of grammar are to be stated in terms of the bar notation categories. Thus, it is of interest to highlight, in a preliminary and cursory way, some basic differences between this book and the work of Jackendoff and Stowell.

Jackendoff's view of (iii) is that, of the asymmetries among N, V, A, and P, the fundamental ones are: (a) only nouns and verbs may take subjects, (b) only verbs and prepositions take prepositionless objects.[2] In my view, these statements are not sufficiently general. With respect to (a), I argue in Chs. 1, 2, and 3 that the fundamental asymmetry between the verb and other lexical categories is that the verb takes an extra projection (or "bar level") not allowed with N, A, and P. From this, several properties peculiar to V, and its projections VP and S, will be shown to follow,

2. At the outset, Jackendoff (1977, 32) states that the feature names that differentiate the lexical categories have only a "heuristic, nontheoretical significance." If this is meant to be true of his features (±SUBJ, ±OBJ), then Jackendoff's phrase structure rules are at best a catalog of descriptive generalizations which, on their own terms, invite further study to separate out the fundamental from the derivative distinguishing characteristics of the lexical categories. However, Jackendoff subsequently claims, for example, that the (surely theoretical) definition of "subject" utilizes the feature SUBJ (p. 41). Thus, my statement accurately reflects Jackendoff's practice, in spite of his disclaimer.

including some differences between subjects in S's and subjects in NP's that do not fall out naturally in Jackendoff's system.

Similarly, I will argue that Jackendoff's (b) should be generalized to "only verbs and prepositions take prepositionless *complements*," and that this statement, when properly formalized, significantly broadens the scope and import of Jackendoff's proposal.[3]

In contrast to the asymmetries just discussed, Jackendoff imposes a parallelism across the lexical categories which commits him to a claim that parallel subcategorizations and interpretations largely determine the deep structure categorial combinations; i.e., his view on (iv) is that the base component is less autonomous than the one I develop here. Specifically, Jackendoff's claim (1977, 36) is that parallel grammatical relations in deep structure are expressed by parallel hierarchical configurations; in Ch. 1 and 2, I strongly dispute this. While I retain structural definitions of "subject" and "object" as NP arguments to a lexical head X which state that they are external and internal, respectively, to X^1, I do not require, as does Jackendoff, that the subject be a sister to a fixed projection of X and that the object be a sister to X. Rather, my definitions of grammatical relations, while simply stated, allow a range of differently situated N and NP to serve as subjects and objects. The resulting interplay between the determination of the grammatical relations (by the semantic component) and the possible deep structures (by the categorial component) permits, as I argue in Ch. 1 and 2, explanations of otherwise unmotivated morphological distinctions and morpheme-insertion rules.

The asymmetries among the lexical categories in Stowell's work are attributed to theoretical statements involving case-theory and government-theory, and are not directly expressed in deep structure categorial combinations. This might seem to be a metatheoretical advantage over my claim (iii), that there are asymmetries in the behavior of various X at deep structure, provided Stowell's statements actually had wider empirical coverage than do the ones I will propose in their stead. But, as I will argue, the opposite is true.

For example, for Stowell, case theory determines that V and P can have prepositionless NP objects, while N and A cannot (the discussion here is of English). As mentioned above, I will establish in Ch. 1 the more general

3. The status of NP complements to A in Chinese (Huang, 1982, Ch. 2), Persian (Samiian, 1983, Ch. 3). and Korean (Jo, in preparation) invites clarification. In languages with morphological case such as German and Latin, these NP's invariably exhibit some oblique case rather than the accusative case (van Riemsdijk, 1983). There is syntactic evidence that oblique cases are PP's at deep structure (cf. Schein, 1981) with a phonologically empty P. Thus, it may be that the NP complements to A established in the above works are of the form $[_{PP}[_{P}\emptyset]NP]$ at deep structure. Morphological case gives additional support for the dative being associated with a deep PP structure: the German dative is the usual case after a lexical preposition, and in Latin, the unmarked prepositional case (the "ablative") is always identical to the dative in the plural. The hypothesis that a dative belies a deep structure PP is discussed in more detail at the end of Ch. 1 and in section 5.7.

proposition that only V and P can have prepositionless *complements*. In order to accommodate this, the theory of NP case must be extended in some way. My proposal (for simplifying the categorial component) is to replace Stowell's enrichment of the stipulations of case theory with an asymmetric principle of θ-role assignment. A simplified general principle of base-dependent case assignment then operates freely on deep structures constrained rather by principles of θ-role assignment. I claim for this system both empirical and theoretical advantages over Stowell's system.

Stowell derives the special properties of the subject of a V indirectly from the role that the special grammatical category INFLECTION plays in his theory of government. Translating terminology somewhat, a second asymmetry across lexical categories in Stowell's system is that VP is the only maximal projection which is always the *sister* to the grammatical formative category INFLECTION that appears with it (by virtue of a special categorial rule expanding S as a projection of INFLECTION that supplements the theory of government).

In contrast, the grammatical formative category that characteristically appears with N, namely DET (determiner), is a *daughter*, not a sister to NP, and similarly for the grammatical category of DEG (degree words) that appears with A. Within the bar notation, this asymmetry follows from my proposal that V, but not N, A, or P, has a third projection in the bar notation. For in the bar notation, each lexical category X is paired with a corresponding grammatical formative category SP(X), called a specifier, which is a daughter to the maximal projection of X. If $S = V^3$, then we can take INFLECTION to be the specifier of V, and it follows that it is the sister to VP($=V^2$), while SP(N) and SP(A) are daughters to NP and AP. Thus, my general claim that only V has a third projection explains why a special grammatical formative category associated with V (i.e. INFL) can appear *outside* V^2, while the same is not true for N and A. But since the claim that $S = V^3$ has other implications as well (set out in some detail in Chs. 1, 2, and 3), this statement is more general than a separate rule, used by Stowell, which stipulates that a special category INFL is a sister to VP.[4]

With respect to Autonomy from the Lexicon (iv), Stowell's approach differs from mine in that he assumes that the theory of the lexicon, yet to be specified, determines the upper limits of complexity for subcategorization frames. While I agree that this is true for the subcategorization of V's, I will argue that a number of systematic discrepancies in the subcategorization frames associated with V's and the corresponding N's and A's can be predicted from the interplay of verb subcategorizations

4. This introductory discussion of the role of INFLECTION in both Stowell's and my theories deliberately glosses over the rather complicated theoretical apparatus we utilize to explain various characteristics of non-finite clauses, and is meant only to give the reader a very general idea of the direction to be pursued in this study. A more complete exposition of my own ideas on non-finite clauses appears in Ch. 2 and 7 of this book.

with asymmetric θ-role assignment principles (Chs. 1 and 2), so that the burden of explanation provisionally placed on the as-yet-undeveloped theory of possible lexical entries is greatly reduced.

I hope that this brief comparison of the approach of this book with those of its "closest relatives," the works by Jackendoff and Stowell, has given some indication of how much room there is for argument, even among those who agree on a relatively wide range of methodological and theoretical points. In order to adequately justify what I feel are considerable improvements over some of their formulations and some points in Chomsky (1981), a book-length study has seemed necessary. Needless to say, if I have succeeded in making such improvements, I owe these authors a great debt for their thorough and insightful contributions to the elucidation of the same basic problems.

I would not want the reader, however, to conceive of this book first and foremost as a comparison of my views with those of other authors. The book has its internal logic. In it, I attempt to treat rather exhaustively all the principles that I feel bear on the deep structures of language.

A universal syntax of deep structure must include statements of combinations allowed in the bar notation, as well as definitions of the basic grammatical relations (subject, object, indirect object) and a theory of how heads assign semantic roles (="θ-roles") to complements.

As these statements are developed in Ch. 1, in particular for the "open" lexical categories N, V, and A, what I take to be fundamental laws governing and setting apart the category N and the category V quickly emerge. These laws are closely linked to the notions of grammatical subject and complement.

Only a noun (phrase) can be a subject.

Only a verb can take complement types freely.

The implication of these extreme restrictions on combining open lexical categories is that a fourth ("closed") head category P emerges. The property of P is that, like V, it can take any complement, and at the same time "transmit" a semantic role from an open lexical category head (N, V, A) to its object. I argue at length in Ch. 1 that P's provide sufficient but also necessary structure for free combinations of open lexical categories. The emphasis placed on P as the *sine qua non* of many grammatical combinations is the basis for the subject matter of the last chapters in the book.

In particular, Ch. 5 argues for the necessity of P in indirect objects and in other structures exhibiting oblique morphological case. Ch. 6 shows that a (non-case-marking) P is structurally required with a wide range of predicate attributes. Ch. 7 claims that any "subordinator," including the much-discussed S-introductory COMP, is also a structurally-induced P, allowing an X to assign a θ-role to an S. Thus, these chapters all testify to the centrality of the structural links provided only by P, as discovered in Ch. 1.

Another fundamental difference between V and all other heads which

emerges in Ch. 1 is that only V has distinct second (VP) and third (S) projections in the bar notation. This distinction leads to separate studies, in Chs. 2 and 3 respectively, of the conditions under which V^2 (VP) can occur alone, and of the properties of V^3 (S).

Since the third projection of V tolerates more transformational deformations of deep structures than do the X^2, Ch. 3 is the natural place to introduce the structure-preserving principle, which sets limits on the divergence between deep structures and s-structures. In particular, the notion of local, language-particular transformation is introduced and exemplified. By seeing the effects of such rules on regular underlying structures, it can be seen how certain recalcitrant language types in fact conform to Categorial Uniformity and Hierarchical Universality at deep structure.

Chs. 4 and 5 further investigate the role of local, language-particular rules in obscuring similarities among underlying categorial combinations across languages. Further principles which apply only to closed categories, the Designation Convention of Ch. 4 and the Invisible Category Principle of Ch. 5, are shown to combine with local rules so as to yield apparently quite diverse surface structures, particularly in the area of inflection. But at the same time, these principles, as well as the local rules themselves, can sanction only a limited range of operations, so that the strong claims about the sparsity and uniformity of deep categorial combinations developed in Chs. 1–3 can stand.

While I have made some claims about what local, language-particular rules can do in Chs. 3–5, I have deliberately excluded much material which will be published separately that elaborates on a restricted theory of language-particular transformations.

In many sections of this book, I develop ideas which I consider crucial for universal grammar, even though they are not presently central points of contention with many authors working in a generative framework. This is especially true in the last two chapters, where I concentrate on the role of the categories P and PP as the main means of subordination and free recursion in syntax. In the last chapter, where I assimilate the category COMPLEMENTIZER to P and the category $\bar{S}$ to $\bar{P}$, I thereby deny that any phrasal category can escape the bar notation. These results, which I think have been fruitful and are certainly potentially controversial, should not be understood as opposed to presently elaborated alternative theories.[5] Rather, my proposals about COMP and $\bar{S}$ are natural simplifications of a presently utilized system of categories, some of whose members happen to stand in a quite complex and previously unnoticed relation of complementary distribution (e.g., $\bar{P}$ and $\bar{S}$ are in this relation).

5. The work of Freiden and Babby (1983), which argues that INFLECTION is the head of $\bar{S}$, has come to my attention after work on Ch. 7 was completed. One of their arguments is that there is an agreement rule between COMP and INFL, analogous to an agreement between SP(N) and N, which suggests to them that COMP=SP(INFL). My

Another hypothesis about the form of universal grammar which is not widely contested but which is centrally important in much present-day theory construction is what I will call the "Adjacency Hypothesis." At several points in this book and throughout my separate work on language-particular transformations, I try to refine and strengthen this proposal, even though it is not the principal subject matter here. Borer's (1984) parametric model, for example, incorporates such a constraint.

> Adjacency Hypothesis: No language-particular rule of any type makes use of a string variable.

For many years, and still in the minds of many linguists, syntax is the only component of linguistic description where it has been thought necessary to use symbols which refer to strings of arbitrary length (="string variables"). If we can demonstrate that such variables are never necessary in language-particular statements, an extremely strong claim about natural language systems embodied in the Adjacency Hypothesis emerges: Given a complete and accurate definition of "adjacent", no child can ever learn a dependency particular to some but not all natural languages which is stated in terms of elements related at a distance.[6]

A quick survey of the components of a typical formal linguistic model reveals that the burden of demonstrating the generality of the Adjacency Hypothesis falls mainly on syntacticians. Thus, it is a commonplace for those working on morphology, whether or not they postulate separate morphological components, to claim that rules of morphology involve only segments that are syntactically adjacent (cf. Roeper and Siegel, 1978, and Ch. 5 here). Similarly, most of the recent developments in formal phonology, both in the "metrical" and the "autosegmental" veins, have been motivated by the twin observations that (a) the majority of phonological processes involve obviously adjacent segments, and (b) the language-particular aspects of those which do not (e.g., prosodic and harmony phenomena) can be recast as local by the proper elaboration of universal phonological theory (personal communication, Morris Halle, Jean–Roger Vergnaud). A recent summary of the formal differences and similarities of autosegmental and metrical phonology (stressing their motivated similarities and the lack of motivation for their differences) makes a point of just this sort:

Footnote 5 — *Continued*
proposals in Ch. 2 and 7 attribute the infinitival *to* to a different cause; I argue that an infinitive form of an S arises if and only if the surface subject NP contains no terminal element. The analysis of *for*-phrase subjects which supports my position is given in Ch. 7, and that of English "raising to object" constructions is given in Emonds (1980a).

6. Culicover and Wilkins (1984) is organized around a locality principle of exactly this type. M. Halle (pers. comm.) attributes the following remark to the physicist L. Tisza: "All physics is an attempt to define adjacency." If so, the Adjacency Hypothesis would suggest that "physics" be replaced by "science

> "A final parallel between the two systems rests on an observation made by John Goldsmith ... In autosegmental phonology, there are two types of spreading. One is maximal and proceeds by a general convention, the well-formedness condition, and the other is confined to a limited domain, e.g. only one or two syllables, and this is achieved by a language-specific rule ... In metrical phonology, analogously, there are essentially two types of tree: unbounded and binary, whether we follow McCarthy's 1979 system or Hayes' 1980 system." (Leben, 1982, 6; I am indebted to Ellen Kaisse for pointing out the passage to me.)

Within syntax proper, there have never been string variables in language-specific base rules. It seems therefore more than plausible that theoretically constrained formal replacements for such mechanisms (such as those developed here in subsequent chapters) can maintain the claim that a deep structure categorial component requires no string variables in its language-specific aspects. Lexical entries, even in their syntactic specifications, are generally assumed to be structured so that no string variables are needed; this is Chomsky's (1965, Ch. 2) claim that lexical subcategorization is local; in this book, local subcategorization will be discussed in Chs. 1 and 2.

The rules of formal semantic interpretation, while probably requiring string variables, supposedly do not vary greatly across languages. To the extent that they do, the required rules can hopefully be stated without recourse to a string variable. Thus, suppose that English and Chinese differ by some semantic rule which is equivalent to "Quantifier Q_i may have wide scope." (For discussion of the relevant data, see Huang, 1982, Ch. 4). If the Adjacency Hypothesis holds true for semantics, then the notion of "wide scope" would have to be automatically determined by universal grammar, even though the actual value of Q_i could vary across languages, and even depend on other factors of the grammar of the language in question. It seems needlessly pessimistic to conclude that universal semantics could not provide definitions of notions like "wide scope", and for this reason I have not included any provisos whatsoever with the Adjacency Hypothesis, since I am confident it can be made to hold in syntax.

Within syntax, the only component that even plausibly falsifies the hypothesis by containing language-specific string variables is the transformational component. Not many years ago, say in 1970, it would have been non-controversial within transformational grammar to claim that the Adjacency Hypothesis is easily falsified even by the most obvious syntactic differences among languages. Thus, one would have said: Japanese and English differ in that English and not Japanese has a transformational rule preposing a phrase marked with the feature WH to sentence-initial position over a string variable Y. That is, the following rule would have been assumed to be part of the grammar of English, but

not of Japanese (X^{max} = a maximal phrase, such as NP, AP, or PP):

$$Y - \begin{bmatrix} X^{max} \\ WH \end{bmatrix} - Z \Longrightarrow 2+1-\emptyset-3$$

Similarly, French, but not English or Japanese, would have been assumed to contain a transformation with a string variable Y by means of which pronominal objects are moved to a pre-verbal clitic position:

$$NP - V + Y - PRONOUN - Z \Longrightarrow 1-3+2-\emptyset-4$$

(Kayne, 1975, Ch. 2).

Today, following a research program initiated in Chomsky (1976), it is assumed practically throughout generative grammar, however different the analyses proposed for such phenomena as "WH-fronting" and "clitic-placement", that these discrepancies in the particular grammars of Japanese, English, and French are not due to language-particular statements containing string variables. The differences are rather attributed to differences in these languages' structures at the "landing site" positions of these movement operations; e.g., Japanese might have no sentence-initial COMP node, and English and Japanese might have no preverbal clitic position defined in their base components; alternatively, within a general transformational "move α" schema, a Japanese-particular statement might specify "$\alpha \neq$ WH" or English and Japanese statements might require "$\alpha =$ PHRASE". Whether it is a question of specifying base positions or of setting restrictions of the form "A $\neq$ B" elsewhere in the syntax, the language-particular statements in question do not contain a string variable.

In Ch. 3 and 5, I will go into more details about how and why we can consider movements across string variables in syntax to be always due to an interplay of universal statements involving such variables and language-particular statements involving category memberships or category adjacency conditions. More extensive syntactic justification for the Adjacency Hypothesis is also the subject matter for another volume on language-particular rules, alluded to above. But even at places where I do not emphasize the Adjacency Hypothesis in this book, the relevance of it to various proposals will be remarked, and in a number of places the hypothesis will influence the choice of formal treatment.

The Adjacency Hypothesis, my hypotheses about $\bar{S}$ and COMP in Ch. 7, and a number of others throughout the book (e.g., those on inflectional morphology in Ch. 5 and the Designation Convention in Ch. 4) are independently justified in a wide variety of ways, and may be easily accepted by readers who might not wish to accept some of the more controversial views of the first three chapters. I would have liked to begin with the less controversial material, but it has seemed in the course of

writing that the basic syntactic and semantic relations between heads and complements, which include subcategorization, control, and what counts as "unmarked movement," must logically precede in my exposition the last four chapters, which deal with grammatical formative categories. So the reader is unfortunately to be exposed to the most controversial (but to my mind, equally well-supported) material first.

The formal plan of the book is then as follows:
Chapter 1: the projections of phrases built around lexical heads X^j; general principles of θ-role assignment, the definitions of grammatical relations, and a theory of abstract case.
Chapter 2: the lexical representation of head-complement relations; subcategorization, the θ-Criterion, and obligatory control; establishing that VP does not imply S.
Chapter 3: the simple unmarked syntactic movements; establishing that S does imply VP.
Chapter 4: grammatical formative categories which are independent words, especially those of category X=N, A, V; the Designation Convention.
Chapter 5: grammatical formative categories which are bound morphemes.
Chapter 6: the non-lexical head-of-phrase P and the range of its projections.
Chapter 7: reducing COMP to P and $\bar{S}$ to $\bar{P}$; the role of the non-recursive, initial symbol E.

In general, the intention of this book is to elucidate as much as possible the formal and empirically justified relations among the bar notation categories, by studying in a relatively complete way those syntactic phenomena in English and French (with some reference to work on other languages when appropriate) which highlight both the differences and similarities among these categories. The goal of such a study, like that of any study of formal syntax, is to set limits on the type of theoretical devices and categories that need to be employed in insightful descriptions of both universal and language-particular grammatical processes. By contemplating these devices and the relations among these categories, we then see a likeness of the power and beauty of the speaking mind, and why it must be respected and developed in each creature that has it.

Chapter 1

The source of categorial asymmetries; indirect θ-roles and generalized case-marking

1.1. Primitive Categories and Heads of Phrases

Traditional and generative grammar agree that central among the categories of syntax are the "major lexical categories": nouns (N), verbs (V), and adjectives (A). The characteristic of major lexical categories is that in English and in most languages they contain usually upwards of a thousand members listed in a lexicon. Also, in typical daily use of language, neologisms ("coinings" of new words) are restricted to these categories. Besides these three major lexical categories, languages have only what can be called "grammatical categories", that is, categories which have at most about twenty or so members, and which are not modified by neologism. Throughout this study, the symbol L varies over exactly the three values N, V, A.[1]

So we begin with the claim that in the realm of syntax, all morphemes are either in a lexical category or in a grammatical category. It is by now familiar in formal accounts of syntax that these lexical and grammatical categories can combine only in certain sequences, and that many of these combinatorial regularities are to be expressed by a set of phrase structure rules or principles of deep structure that generate labelled bracketings of morpheme sequences, called deep structures.

The well-formed morpheme sequences of deep structure we will label here as E (for "expression", following Banfield, 1973). Among the various E, the type that for centuries has rightfully been a principal focus of investigation by students of grammar and semantics is the one which may express a "judgment", in the sense of Frege (see the discussion in Kuroda, 1975). This type or "expansion" of E is called a sentence or, when attention is on its structure rather than on its sense, a clause, and it is notated S. In the familiar Indo-European languages, morpheme sequences which have the structure of an S generally must contain a grammatical category

1. The productive class of adverbs that end in *ly* in English is considered to consist of adjectives with an ending. Cf. the discussions in Jackendoff (1977, section 2.3) and Hendrick (1978).

The category preposition has more than twenty members in many languages. Its special status as a grammatical category which is also a "head of a phrase" is taken up in detail in the latter part of this chapter.

expressing "tense" and must not be contained in a larger S in order to express a judgment. However, even if these two criteria are not met, grammarians usually do not hesitate to assign the label S to the sequences in question, if their deviance from the structure of well-formed judgments is minimal. Thus, in the grammar developed here, all the italicized sequences in (1) are called "S", even though only the first sequence expresses a judgment.

(1) Yes indeed, *somebody will start dishing the children out their lunch.*
Somebody start dishing the children out their lunch.
I don't know if *somebody will start dishing the children out their lunch.*
For somebody to start dishing the children out their lunch wouldn't be appropriate.

Once the categories E and S, the major lexical categories L (=N, V, A), and the grammatical categories are admitted into the theory of syntax, the question that arises is whether any intermediate subsequences of categories should be assigned category labels. Again, it is by now widely accepted that any other such category which can occur in deep structure and which can contain a lexical category is structured around an obligatory "head" category X, and is notated X^i, where i is a small integer. The possible values for X are L and also P, where P (which usually corresponds to the traditional term "preposition") is a head which is a grammatical rather than a lexical category. The categories L^i and P^i are called "phrases", and when i is maximal for a given X (notated "X^{max}"), we say that X^i is an "X-phrase", or a "maximal projection of X". Thus, N^{max} is a "noun phrase", and X^{max} is alternatively notated "XP". This notation for phrases is called the "bar notation", in that X^i can be alternatively written as X with i bars over X (e.g. $X^2=\bar{X}$).

In this regard, I will sometimes use the term "particle" for P, and hence PP = P^{max} = particle phrase. In this study, no formal difference distinguishes "prepositional" and/or "post-positional phrase" from "particle phrase"; for justification, see section 6.2.

The particular variant of the bar notation I will begin with here assumes that each deep structure phrase $X^i, i>0$, consists of a unique and obligatory head X^j, possibly accompanied by certain non-head grammatical categories, by maximal projections, and by S.[2] The sequences of categories which make up X^i or S are called the immediate constituents or daughters of X^i or S.

Some interesting questions can be raised concerning the value of "max" appropriate for each value of X. Jackendoff (1977) advances the "uniform

2. Variants in which certain non-maximal projections can be sisters to X are proposed in Ronat (1973) and Zagona (1982); a restricted use of this idea is made in Ch. 2 of this study.
In coordinated constituents, we probably want to say that there are multiple heads. Cf. Dougherty (1970).

three-level hypothesis", where max = 3 in all cases. In Emonds (1979), I argue that for N, the value of max is 2. In this study, I further claim that V contrasts with N, A, and P in that the value of max for V is 3; I return to this matter below.

It is generally recognized that phrase structure rules are inadequate for expressing linguistic generalizations about well-formed bar notation deep structures. One thorough critique of phrase structure rules can be found in Stowell (1981, Ch. 2). Even Jackendoff, in his attempt to provide a more or less complete set of English phrase structure rules, expresses reservations in the end on whether such rules can adequately express descriptive generalizations (1977, 81–85).

Many of these criticisms can be summed up under the following two very general and I think very telling points: First, phrase structure rules make it impossible to clearly distinguish the contributions of universal grammar from aspects of particular grammars; among other things, certain left-to-right orderings expressed by classical phrase structure rules are language-particular, while most of the hierarchical structure they assign is either arguably universal, or in any case can hardly be asserted to be language-particular. Second, phrase structure notation wrongly implies that too large a number of different sets of base rules for deep structures are possible.

As an example of these two points, consider the fact that grammatical formatives which are not affixes or clitics precede their head in the majority of languages and across all values of X^j. (X varies across N, A, V, P and j varies from 0 to 2.) Even if all languages do not conform to this type, it is surely the case that many languages (e.g., English and French) conform "on the whole"; individual phrase structure rules for expanding the various X^j would thus fail to reflect this property in a revealing way. This generalization can be expressed as (2).

(2) Head Placement for Non-phrasal Modifiers: In deep structure, all immediate constituents of X^i which are not clauses or phrases precede the head of X^i. (Y^j, where $j > 0$, is called a phrase.)[3]

Some ways in which English and French conform to (2) are as follows. The English verb is preceded not only by the auxiliary, but also by negation and by certain ummodifiable adverbs A (of the *scarcely* type; cf. Emonds, 1976, Ch. 5.). French negation words (*pas*, *point*, *guère*, *jamais*, etc.) and certain other unmodifiable adverbs follow the finite verb in surface structure, but it is shown on independent grounds in Emonds (1978) that the finite – but not infinitival – verb moves transformationally to the left over these grammatical formatives, so that in fact French verbal negation confirms (2) in an interesting way.

3. For a treatment of grammatical categories in which a similar principle plays a role in a Categorial Grammar framework, see Flynn (1983).

Similarly, the English and French head noun is preceded in deep structure not only by determiners, but also by negation, by unmodifiable adjectives, by numerals, etc. While such grammatical formatives are not necessarily members of a single archi-category such as SPECIFIER (X), they are not phrases Y^i, $i>0$, and hence, by (2), they precede the head:

(3) *Not one* person did I see. (only order possible)

The book came on time, but { *not* the pencil. / *the pencil *not*. }

The *three* houses on the block are old.
*The houses *three* on the block are old.
The (*very) *other* reason for this is...
*The reason *other* for this is...
John's (*most) *principal* objections to that are...
*John's objections *principal* to that are...
A *mere* mention of that would...
*Too *mere* a mention of that would...
*A mention *mere* of that would...
John has (*very) *barely* finished.
*John has finished *barely*.

A possible objection to (2) might be made on the basis of the English post-verbal "particle" node PRT, which often appears in discussions of the English VP. However, it has been argued in Emonds (1972) and never seriously refuted that such particles are instances of PP in deep structure.[4]

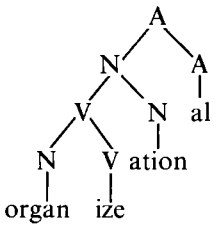
A striking confirmation of the validity of Head Placement (2) comes from a consideration of morphology. In Chapter 5, it will be argued that all English and French *inflectional* morphology (which follows the head in surface structure) is derived either from pre-head positions in deep structure or from a transformational adjunction to X^i in such a way that (2) is not violated.[5] Thus, in deep structure, the categories from which the English tense suffixes, the English plural marker, the English adjectival comparative markers, etc. are derived all *precede* the X^i to which they are attached in surface structure.

4. Jackendoff (1977) considers such particles to be a bar notation phrase, and as such, they could not be a counterexample to (2). However, Jackendoff attributes phrasal status to all categories, so a generalization akin to (2) in his system would have to be stated in terms of his features ±COMP and ±DET. Without some revision in the distribution of these features, the generalization cannot be straightforwardly expressed in such terms.

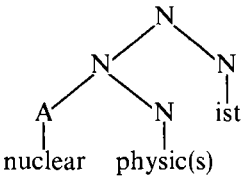
5. My argumentation that inflectional morphology is transformationally derived would not have been controversial ten years ago. A different position is put forward in Lieber (1980). In Chapter 5, I will counter some of her argument, but I accept what seems to be her most interesting claim – that rules of the same type are needed to express certain generalizations about derivational and inflectional morphology (cf. her Ch. 2). I think what is crucial is that when syntactic phrases $Y^j (j>0)$ are being composed, it is forbidden to use subcategorizations whose domain is Y^0. Beyond this, it may well be that (derivational) rules which apply inside Y^0 and (syntactic) rules which apply outside Y^0 are of the same type.

The scope of (2) extends even to derivational morphology. If we follow the interesting argumentation of Williams (1981) and Lieber (1980) to the effect that the *head* of a word composed by derivational morphology is its category-determining derivational suffix (speaking for example of English, Latin, Polish, French, etc.), as exemplified in (4) and (5), then (2) holds for the value $i=0$, as well as for the syntactic cases, where $i>0$.

(4)



(5)



Taking up Lieber's suggestion, the difference between a derivational affix which is, say, an N and a lexical N is simply the presence or absence of a subcategorization feature. Thus, *(a)tion* is + N, + V____, while *organ* is simply + N.

Further, it is rather obvious, as has been observed to be by H. Hoji, that subcategorizations of particular items do not vary significantly across languages with differences in word orders: so the subcategorization mechanism should not refer to linear order. Lieber points out some exceptions to (2) among English affixes: thus, the verb-forming *be-* as in *befriend*, *besiege* is a prefix. She acknowledges their atypical status. We can assume the verb-forming prefix *be*, parallel to *ize*, is listed as taking a +N complement with the added exceptional stipulation that *be* is a prefix:

(6) [V *ize*], + N____ . [V *be*], + N____, "is a prefix".

Now let us assume that linguistic theory requires at a given level of structure that less general statements, being marked, supersede more general statements. Here, the algorithm that determines (6) to be less general than (2) is the obvious one that (6) contains a constant (*be*) while (2) does not.[6]

6. Actually, this algorithm is what Sommerstein has proposed for phonology, where he also incorporates the claim that less general statements precede more general ones:

> Proper Inclusion Precedence: If every (logically possible) form meeting the structural description of rule A (here rule 6, JE) also meets the structural description of rule B (here rule 2, JE), and the converse is not the case, then rule A has precedence over rule B (Sommerstein, 1977, 186).

Added support for Lieber's contention that affixes like *be-* and *-able* have the categories of heads comes from the fact that a synchronic grammar can reflect the naturalness of certain grammatical formatives representing both lexical categories and derivational affixes by the use of parentheses in subcategorization: *be*, +V, + ____(N); *able*, +A, +(V)____.

Even though the Head Placement Principle (2) holds for a variety of both verb-second and verb-final languages, it is not clear at this point whether a language-particular statement is involved, or whether Head Placement is in fact a universal, subject only to language-particular exceptions such as (6). If Head Placement should turn out to be language-particular, then its eventual formal statement must, according to the Adjacency Hypothesis discussed in the Introduction, be stated without an internal variable. This might be done in a number of ways, depending on how languages which are marked with respect to head placement fail to conform to (5); presumably, the language-particular statements would describe departures from (5).

It is more likely that Head Placement, given its applicability to widely differing languages, is a consequence of universal grammar. But it could nonetheless follow from very different theories of grammar: (i) In one theory, the head follows everything in deep structures in the unmarked case, and verb-second and verb-initial languages share a language-particular stipulation that places the head before (only) maximal phrases and clausal complements. (ii) In another theory, Head Placement (5) holds universally, but it is not formally related to the relative order of the head and its maximal phrase complements. (iii) In a third theory, Head Placement is the consequent in an implicational universal that applies to languages with a fixed word order base.

Given these divergent possibilities, it seems sufficient to leave Head Placement is the consequent in an implicational universal that applies to of factoring its content out of the statements that determine the hierarchical constituency relations in deep structures.

1.2. A Preliminary Account of Specifiers

Many, and probably most, of the grammatical categories which are not prepositions (that is, not of the form X^{i}) are assigned in deep structure to categories called the "specifiers" of X, notated here SP(X). Languages like English and French accord with Head Placement (2), in that the specifiers of various heads precede the head. Previously (e.g., in Chomsky, 1970) this was expressed in the following phrase structure rule:

(7) $X^{max} \rightarrow SP(X) - X^{max-1}$

The most clear-cut representatives of SP(X) in English are as follows, ignoring for the moment the many contextual restrictions on the various specifier morphemes.

(8) SP(N) = DETERMINER = *this, that, these, those, the, a(n), each, every, all, both, half, some, any, no, which, what.*
SP(A) = INTENSIFIER = *very, so, as, more, most, less, least, too, enough, how, somewhat, rather, quite, real, this, that.*

SP(V) = AUXILIARY = *will, would, can, could, may, might, shall, should, must, ought, need, dare.*
SP(P) = *right, clear, straight.*

I return in section 1.4 to why SP(P) has so few members. The grammatical behavior of the members of SP(X) is dealt with in some detail in Ch. 4 and 5 below. Here, it suffices to make two general points about the category SP(X).

The first point about the category symbol SP(X), as expressed, for example, in (7), is that a fundamental and uniform syntactic relation between SP(X) and X across values of X(=N, A, V, P) is implied. But moderate reflection on the nature of SP(V) (=tense and modal categories), SP(N) (=demonstratives, quantifiers, numerals), and SP(A) (=expressions of degree and intensity) strongly suggests that with respect to their semantic interpretation, any parallels among these categories are at best secondary to the fundamentally unique roles they play in logical semantics – so much so that Jackendoff (1977, 37) denies that the various SP(X) are formally related at all.

However, to deny syntactic status to SP(X) would miss the striking fact that each of the central lexical categories L is tightly associated with a particular closed grammatical category SP(L) both in the base and in the operation of the local transformations (cf. Ch. 5 below). Moreover, minor parallels have been discovered among the SP(X), especially between SP(N) and SP(A) (e.g., the *this–that* contrast and the possibility of a WH member: *which, what, how*). Thus, there are reasons for treating all SP(X) as parallel, and for having a single symbol available to refer to them.

In contrast, the uniqueness among the SP(X) is in *interpretation*. The rules for logical interpretations of tense morphemes based on surface configurations given in Emonds (1975) have no counterparts in the other SP(X) systems; similarly, rules of interpretation for other SP(X) proposed by other authors (Jackendoff, 1977, Ch. 5–6; Milner, 1978, Ch. 7–8) usually do not generalize across values of X. The category SP(X) therefore conforms to the following principle, in large part suggested by Stowell (1981), and which is a working hypothesis throughout this chapter.

(9) Category-neutral Syntax: Syntactic principles of the base component generalize across values of X; rules of semantic interpretation are often based on category-specific values of X.

I take this principle to be the source of the autonomy of syntax and semantics. In my view, this pervasive discrepancy between parallel syntax and asymmetric semantics has been either misread or at best insufficiently articulated by previous authors. Jackendoff explicitly claims that fundamental semantic relations and deep syntactic structures are parallel (1977, 37). Stowell (1981, Ch. 2) raises the possibility that apparent asymmetries in the base are due to rules of logical form, but, as will become clear, I do not believe he accurately locates these asymmetries.

At this point, I will not develop further the membership system, the syntactic behavior, or the semantic interpretation of SP(X), but I do consider that the general properties of these categories give credence to the above claim that rules of semantic interpretation are often category-specific, while syntactic principles of the base, including those which involve SP(X), tend to generalize across values of X.

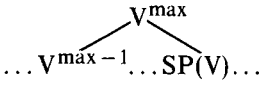
A second point about the category SP(X) is that we must specify its existence in universal grammar by factoring out the linear order expressed in (7) now captured by Head Placement (2). The residue of (7), which is almost certainly a universal, is expressed in (10a):

(10) (a) SP(X) can only be a daughter of X^{max} and a sister of X^{max-1}.

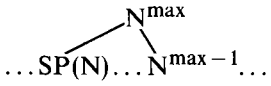
This formulation allows us to accomodate the possibility that some languages may be "flatter" than others; that is, the value of "max" might be less for some languages than for others, and/or some languages might require that the head of X^j always be X^{j-1}. I will not be concerned with these possibilities here.

Two types of trees generable by (10a) are exemplified in (10b–c):

(10) (b)



(c)



Of course, (10b) is ruled out as a deep structure in a language which conforms completely to Head Placement of Non-phrasal Modifiers (2). (10c) is exemplified by English.

Formally, I will write the operation in (10) as (11):

(11) $X^{max} \rightarrow SP(X),\ X^{max-1}$

$C \rightarrow A, B$ should be read as "C may dominate the immediate constituents A and B." It is not implied that A and B are the only daughters of C in a well-formed tree conforming to such a rule. Moreover, C can be left unexpanded. I call such rules "base composition rules." In my conception, this kind of statement is typically part of universal grammar. Statements of this sort express the Categorial Uniformity and the Hierarchical Universality discussed in the Introduction.

I do not totally exclude the possibility that a base composition rule may be language-particular. However, I tentatively propose that *language-particular base composition rules are limited to expansions of non-phrasal nodes* – for example, differing possibilities for compound L or for expansions of COMP in different languages. An English-particular base composition rule is also provided by (12);

(12) $SP(X) \rightarrow NP,\ X \neq V$; English-specific

The English possessive NP is one type of phrase which is generated by (12). Such phrases alternate with a range of determiners; in particular, a possessive NP has a distribution almost identical to that of *that/those* and *this/these*. Chomsky (1970) provides evidence that some possessive NP's are base-generated.

Rule (12) should not be considered to play a role only in the possessive construction. Pre-head measure phrases, which almost certainly fit into the specifier system of both A and P, also can be generated by (12). Like possessives, they alternate with individual specifier morphemes.

(13) (a) The house is very high. The house is ten feet high.
*The house is very ten feet high. *The house is ten feet very high.
John is standing right behind the house.
John is standing a short distance behind the house.
*John is standing right a short distance behind the house.
*John is standing a short distance right behind the house.

Measure phrases modifying nouns, as in *John's two mile driveway*, are not generated by (12). Like adjectives, these phrases follow the SP(N), and they also fail to exhibit the plural morpheme.

French, in which the distribution of phrases in deep structure is otherwise almost identical to that in English, exhibits neither lexical possessive phrases nor lexical pre-head measure phrases inside AP and PP. Thus, it seems to be a correct generalization that the expansion of SP(X) as an NP is English-specific.

(13) (b) *La maison est dix mètres haute. "The house is ten meters high."
*Jean se trouve une petite distance derrière la maison.
"John is standing a short distance behind the house."

Finally, specifiers, like other non-head constituents, are typically optional, even though classes of heads, such as count nouns, may require a specifier. Exceptionally, the SP(V) appears to be an obligatory constituent of S; I return to this point in Ch. 3.

In summary, while base composition rules for expanding a morpheme category may be language-specific, I claim that the possible expansions for phrases are determined by universal rules whose proper form factors out linear ordering conditions and is typified by rule (11). These syntactic rules are universal and category-neutral.

1.3. Subject Phrases

As mentioned above, my formalization of base composition rules does not imply that the category on the left of such a rule may dominate only categories specified on the right of that same rule. The same category may

appear on the left in different base composition rules; each base composition rule is a maximally simple expression of a single generalization. Some generalizations about the possible dominance relations in the base overlap in interesting ways, even though their interplay cannot and should not be expressed in a single formula.

For example, the rule for specifiers (11), which applies across all values of X, co-exists, in my view, with a rule for certain generating subject phrases Y^{max} which holds at least for the value X = V.

(14) $X^{max} \rightarrow Y^{max}, X^{max-1}$

For X = V, (11) and (14) together yield the structure $[_{V^{max}} Y^{max} - SP(V) - V^{max-1}]$. For reasons to be given below, this reduces to the familiar expansion of S, $[_S NP - AUX - VP]$.

The base composition rules being proposed thus contrast with phrase structure rules or a bar notation schema in that a single composition rule does not necessarily specify all the daughters of a single parent node. However, no base composition rule can be used more than once to expand a single symbol; for example, multiple subjects and specifiers for a single X^j are not allowed.

Regarding (14), there are two questions: What are the possible values of X (that is, which types of phrases may contain subject phrases)? And, what are the possible values of Y (that is, which types of phrases may serve as subjects)? Stowell (1981, Ch. 4) takes the position that both X and Y may vary over all head-of-phrase categories. While I agree with Stowell's research heuristic of a category-neutral base component, I think he is mistaken on both counts with regard to the issue of "subjects across categories."

First, let us consider the question of whether all phrasal types may serve as subjects. I contend that there is a universal consonance between subjects and the category NP, which Stowell has failed to undermine the arguments for. His account of this correlation is that categories which receive abstract case (e.g., NP) appear in case-marked subject positions, while categories which do not are excluded. But AP's typically exhibit morphological case (in Indo-European case-marking languages) and presumably take abstract case, yet they cannot regularly be subjects. Conversely, PP's and S's, which do not receive abstract case, should freely appear as subjects of infinitives, yet they do not. That is, whatever devices might be called upon to explain these facts, there are no patterns which suggest that the theory of case alone can explain why non-NP's consistently fail to occur as subjects.

Moreover, it is precisely when the correlation between NP's and subject position appears weakened (by, for example, verbs which seemingly have $\bar{S}$ subjects) that more thorough investigation has revealed that NP (and N) are inexorably present as deep structure subjects. While Stowell accepts the descriptive generalization of Emonds (1976, Ch. 4) to the effect that $\bar{S}$

and PP are not in fact in subject position in surface structure (they must be topicalized or extraposed), he does not address the explanation I gave for *why* any $\bar{S}$ and PP generated as deep structure subjects (and NP objects) move. In brief, my explanation was and is that these are NP positions, and that the empty head N required by the bar notation must be either removed or co-indexed during the transformational derivation in order to yield a well-formed surface structure. Hence, it follows from the bar notation and the trace theory of movement rules that the $\bar{S}$ generated as the sole lexical phrases in base NP positions (necessarily with an empty N sister) must move and leave a co-indexed element in these empty surface NP positions.[7]

Therefore, the explanation for why $\bar{S}$'s do not appear in surface NP positions depends directly on a restriction like (15):

(15) The Subject Principle: Phrasal arguments of X external to $\bar{X}$ (i.e., subject phrases) must be NP's.

The Subject Principle shound not be coalesced with (14) because it has wider applicability. For example, (15) ensures that a possessive subject phrase generated as a daughter of SP(N) by (12) will be an NP; that is, in light of (15), the symbol NP in (12), intended to encompass both subject phrases and measure phrases, can be replaced by the category which includes both NP's and AP's, since AP's also serve as measure phrases (*three dollars cheaper, very much cheaper; three dozen books, too few books*, etc.). In succeeding sections of this chapter and in Ch. 2, I will also establish that subjects can occur outside the maximal projection of their predicate, and in these positions also they are consistently NP's, even though they are not generated by (14).

7. I return in sections 1.6, 7.7.1, and 7.7.2 to details of how this co-indexing is achieved and to how the empty N's are removed when $\bar{S}$ is topicalized or extraposed. At the time I originally made my proposal, trace theory was embryonic, and my own formulations incomplete and in certain ways *ad hoc*.

In more detail, Stowell's proposal is that $\bar{S}$'s and PP's are generated directly as subjects (with no empty N), and that they are forced to vacate this position because they are incompatible with nominative case. But there are two interpretations of the "case-resistance" of categories like $\bar{S}$ and PP; one is that they don't receive case and that the proximity of a case-assigning category has no effect on them, and the other is that the proximity of a case-assigning category actually forces them to move. The only other instance of such "forced movement" adduced by Stowell concerns his claim that $\bar{S}$ (but not PP) inside $\bar{X}$ can be interpreted only by virtue of being extraposed. But in just those instances where my original proposal contrasts with his (where I claimed there was no $\bar{S}$ extraposition involved: the $\bar{S}$ complements to verb classes like *seem, murmur,* and *persuade*, and to nouns and adjectives), Stowell fails, in my view, to show that $\bar{S}$ moves. In some cases, he utilizes *ad hoc* devices; in others, the verb classes are not dealt with; and in others, incomplete and unrepresentative data suggests patterns that are just not there (here I refer to his treatment of $\bar{S}$ complements to nouns). Thus, I don't believe that Stowell has established that $\bar{S}$ or PP exhibit any "forced movement" other than not being compatible with subject position or subcategorized object NP position.

For these reasons, the Y in (14) should remain category-neutral, in accord with Stowell's (9). The Subject Principle (15) always restricts this Y to the value N. This restriction can be thought of as part of category-specific semantics, in the sense that (15) concerns the way predicates are related to arguments; the external argument of X must be an NP.

Actually, the Subject Principle as stated above may well be a "syntactified" version of an even more obviously logical or semantic requirement, and might easily be correlated with general observations brought forward by Keenan and other writers that subjects often are required to be "referential" in ways that other NP's are not (cf. Keenan, 1976). We might rephrase (15), for example, as "Specifiers of external arguments must be able to express quantification or co-reference." This directly suggests its semantic nature.

Let us turn to whether all maximal projections can contain subjects; should X in (14) be restricted? While I agree that members of all head categories X may impose selection restrictions on and be in a grammatical relation with a subject phrase external to $\bar{X}$, I will argue in the next chapter that the structural definition of subject phrase ("=argument external to $\bar{X}$") is satisfied by a variety of configurations, and not only by daughters of a maximal projection as required by Stowell, and also Jackendoff (1977, section 3.4). In this I agree with Travis and Williams (1983) and Williams (1983). The contribution of (14) to generating subject phrases is that of satisfying the requirement that clauses, in contrast to other X^{max}, must contain expressed or understood syntactic subjects of X. In English, X takes on only the value V in (14); that is, it is not the case that the subject of X is always inside X^{max} (cf. Ch. 2).

In Chs. 2 and 3, I will argue that X is limited to V in (14) not by direct stipulation, but from the fact that universally, at least in the unmarked situation, only V has three rather than two bar notation projections. This is expressed by reformulating (14) as (16), and imposing (17), a category-specific statement which describes the unmarked case.

(16) $X^3 \rightarrow Y^{max}, X^2$

(17) Only V can have 3 rather than 2 projections.

Taking into account the Subject Principle and the fact that subjects are typically clause-initial, (16) reduces in phrase structure rule format to "$X^3 \rightarrow NP - X^2$." This is the putative universal rule proposed in Williams (1984) for languages with verbless deep structure clauses, if we equate X^3 here with Williams' S. (Alternatively, X^3 could be replaced by V^3 in (14), and verbless sentences would be "exocentric"; the category of the projection and of the lexical head would not agree.) If Williams has the right analysis for such sentences, (17) is an unmarked category-specific option for the category-neutral (16). If there are no verbless deep structure S, so

that (17) is universal, it can still be thought of as category-specific semantics; only verbs can have arguments external to X^2.[8]

Category-specific generalizations such as (15) for nouns and (17) for verbs delineate the fundamental distinguishing properties of the lexical categories, and can account for much of their asymmetric behavior in individual paradigms. Such generalizations come closer to explaining these pervasive asymmetries, which we will be examining in detail in the rest of this chapter, than do formal cross-classifications of N, V, A, and P in terms of two binary features. The generalizations of (15) and (17) are essentially correlations between the type of meanings expressed by SP(X) (reference for N and modality for V) and the deep structures or logical forms the corresponding X^{max} can appear in. The differences among N, V, A, and P seen in this way cannot be referred to freely either in the transformational component or in other statements of deep structure and logical form. This restrictiveness is a welcome development, even though it might appear at present to be too strong a claim to say that (15) and (17) are the *only* statements in grammar that distinguish among N, V, and A.

The base composition rules (11) and (16), the Subject Principle (15), and the Head Placement Principle (2) allow for three types of deep structure clauses: NP – AUX – V^2, AUX – NP – V^2, and AUX – V^2 – NP, where AUX = SP(V). In order to complete the specification of the English deep structure S as the first of these, we must insure that *the deep structure subject is initial in S*. Since this is most likely the unmarked case in universal grammar, there is no need for a further stipulation in the grammar of English. How exceptions to this word order constraint are stated for other languages is important, particularly because of the Adjacency Hypothesis of the Introduction. However, serious consideration of language-particular rules outside of English and typologically similar languages is beyond the empirical scope of this study.

Since subjects of clauses are arguably obligatory (Chomsky, 1981, Ch. 2), there arises the question of when a category that appears on the right of a base composition rule is obligatory. Beyond the fact that each X^j must have an obligatory head X^k in deep structure, we know that the principles of the base should allow complement phrases to be optional, with obligatory occurrence being stipulated by subcategorizations of individual lexical items. How far should we extend the notion that non-heads are syntactically optional? In Ch. 3, I return to the question of what renders subjects obligatory.

8. In Emonds (1980b), three arguments are given, apparently contrary to the idea that the subject of a V is inside V^{max}, that the subject NP is external to VP. But these three arguments are easily made compatible with the present treatment by replacing X^{max} in the principles I formulate there with X^k, $k \leqslant 2$. That is, the arguments in Emonds (1980b) can be construed either as arguments for Hornstein's (1977) position that $S \neq V^{max}$ *or* for the position that max for V is greater than max for the other head categories. This latter position will be developed in Chs. 2 and 3. In either case, these arguments are evidence against Jackendoff's "uniform 3-level hypothesis" in which $S = V^{max}$.

With preliminary but I think plausible principles and base composition rules for heads, specifiers, and subjects in mind, we can now turn to the central problem addressed in this chapter, the asymmetries across syntactic categories found in the complement structures to N, A, V, and P.

1.4. Two Base Composition Rules

In this section, I will argue that the expansions of X^k, $k = 1$ or 2, provide solid evidence for the principle of Category-neutral Syntax (9). The Subject Principle (15) will play an important role throughout. Careful consideration of the broadest generalizations holding of these expanisons will provide more support as well for factoring out universal statements of dominance relations from language-particular ordering statements.[9]

Concerning left-right order, the fixing of a general word-order parameter has the following effect on head-initial languages such as English:

(18) Head Placement for Phrasal Complements: A phrase cannot be a left-sister to the head of a *deep structure* X^k, $k \leqslant 2$.[10]

The verb-second and the verb-first languages of Greenberg (1963) are those which are subject to (18). The status of verb-first languages is treated in some detail in Emonds (1980b), and will be returned to here in Ch. 3. My claim that subject phrases in sentences are outside V^2, as expressed in (16) and (17) above, allows subjects of sentences to escape the effect of (18), which does not apply to a third projection of X. Sentence-initial adverbial also escape (18), and the free recursion they exhibit (*Unfortunately for the average person who has a moderate income*, ...) confirms that V should have a third bar notation projection.

In all other cases where phrases apparently precede heads in English, an argument can be made that either (i) a deep structure configuration is not involved, (ii) the apparent phrase actually is an X^0 in deep structure, or (iii) the phrase is a daughter of SP(X), and not itself a left-sister to the head.

An English pre-head phrase which exemplifies sometimes (i) and sometimes (iii) is the possessive NP. As argued in section 1.2, base-generated possessive NP are daughters of SP(N). Such pre-head NP's are not found in general across head-initial languages, and they do not violate (18), if they are generated by a rule like (12).[11]

9. Some of the empirical considerations leading to my conclusions are also taken into account in Stowell (1981). However, my conclusions are different in many ways, as will be noted at appropriate points.

10. Head Placement rules out a left-branching verbal complex in French, of the type proposed in Emonds (1978), unless my V′ there is re-interpreted as V^0.

The principle of Head Placement allows the "complements" of traditional grammar to be defined as phrases internal to $\bar{X}$, and "modifiers" to be defined as either "external to $\bar{X}$ or a non-phrase." These characterizations are independent of left-right order, and thus seem superior to the definitions in Jackendoff (1977).

11. Even if my claim that English possessive and measure NP's are daughters of SP(X)