Chapter 1
A General Theory of Word Structure

Within the context of generative grammar, a variety of approaches to morphology have been pursued. In the first work on the topic, Lees (1960) proposed that complex words—compound words as well as those involving derivational or inflectional affixation—be derived through the operation of syntactic transformations from deep structures including only noncomplex words. Chomsky (1970) presented important arguments against this approach to derivational morphology, concluding that derivationally complex words must be present in deep structure. The same sorts of arguments lead one to conclude that compounds are present in deep structure (cf. Allen (1978)). In this monograph, as in other generative works on morphology (Dell (1970, 1979), Halle (1973), Jackendoff (1975), Siegel (1974), Aronoff (1976), Allen (1978), Lapointe (1980a), Strauss (1979a,b), Williams (1981a), Lieber (1980)), the conclusion that words with derivational morphology and compound words are not formed by syntactic transformation is taken as a point of departure. Along with this view, I adopt the somewhat less universally held assumption that inflectional affixation is not accomplished by syntactic transformation, but that, with derivational affixation and compounding, it instead forms part of a morphological component of grammar.[1]

My purpose is to examine what I will call the *syntax of words*, by which I mean the *structure of words* and the *system of rules* for generating that structure. While much has been said in the recent linguistic tradition about the syntactic structures of which words form the basic units, considerably less attention has been paid to the structure of the words themselves. Perhaps this has come about because word structure seemed perfectly obvious, apparently a mere extension of syntactic structure. This is the view of word structure implicit in Chomsky and

Halle (1968; hereafter SPE), for example, and one that is adopted in most subsequent works. However, it is an error to view word structure as merely the "lower" portion of a syntactic representation that is entirely homogeneous in character. It can be argued that, aside from the category Word itself, the categories involved in word structure are distinct from those of syntactic structure and, moreover, that the two types of structure combine these categories in significantly different ways. It in fact seems possible to construe word structure as an autonomous system. In my view, the category Word lies at the interface in syntactic representation of two varieties of structure, which must be defined by two discrete sets of principles in the grammar.[2] Yet I will argue that word structure has the same general formal properties as syntactic structure and, moreover, that it is generated by the same sort of rule system.[3]

In order to underline this fundamental similarity, I will often employ the terms *W-syntax* and *W-syntactic* rather than the terms *morphology* and *morphological* in speaking of the structure of words. The terms *syntax* and *syntactic* will be reserved for their normal usage, though it may be convenient from time to time, for reasons of contrast, to employ the terms *S-syntax* and *S-syntactic* instead.[4]

First, I will argue that, just as it is the appropriate formal device for generating syntactic structures, a context-free grammar is appropriate for characterizing the notion "possible word structure of L." Second, I will show that certain fundamental notions of the so-called $\bar{X}$ theory of phrase structure (i.e., S-structure) can be profitably extended to the theory of W-structure.[5] As for the question of whether transformations have a role to play in the grammar of words (mapping one W-structure into another), I consider it open. The phenomena I will be examining give no support to the notion that transformations form part of the W-syntactic component of grammar.

A caveat may be in order here. It should be understood that my arguments concerning the formal properties of word structure and word structure rules are founded in large part on the W-syntax of English and related languages. It is enough to consider the Semitic system of derivational morphology, for example, to understand that a context-free grammar such as that of English is but one of perhaps a variety of types made available by universal grammar. McCarthy (1979, 1981) has shown that the characterization of the derived verb forms in Semitic requires a simultaneous "three-dimensional" representation of morphemes. The segments of a morpheme on one "tier" (e.g., the con-

sonants of the verb root) are intercalated with the segments of a morpheme on another (e.g., the "vocalization") by means of their common association with a template, which is a morpheme on a third tier. Semitic derived verb forms thus have no immediate constituent structure. Semitic words also include affixation of the more conventional sort, which can be described in the same way as affixation in other languages. The W-syntax of the Semitic languages, then, includes two components, only one of which is of the sort I am attempting to characterize here. It therefore seems that the context-free grammar theory of word structure I will be defending here forms part of a typology of word structure systems. The substantive claim that I am making is that English word structure can be properly characterized solely in terms of a context-free grammar.

1.1. A Context-Free Grammar for Words

As a context-free constituent structure grammar, word structure rules assign a labeled tree (a structural description) to every word of the language. This sort of grammar thus captures the intuition of native speakers of English (and other languages) that words have an internal constituent structure, the constituents of which may be assigned to different categories. A context-free rewriting system also allows for the recursiveness or self-embedding evidenced by morphological structure; it embodies the claim, a correct one, that there is no principled upper bound on the length of words.

Part of the interest of the claim that English word structures (involving both compounding and affixation) are generated by, and only by, a context-free rewriting system is its controversiality in the current context. Roeper and Siegel (1978), for example, have proposed that a set of what they call "lexical transformations" are operative in the generation of verbal compounds. This transformational analysis will be rejected in favor of a rewriting rule account in section 2.3. In addition, treatments of affixation such as those proposed by Allen (1978) or Lieber (1980) have been based on word formation rules which could possibly be viewed as rules of a categorial grammar. A categorial grammar is at best a notational variant of a context-free rewriting grammar. In what follows, I have chosen to couch the formal theory of word formation in the most familiar terms of the latter sort of model, and leave it to others to argue that it should be done otherwise.

A context-free rewriting system by itself is capable of generating all of the words of a language, but only at a certain cost. Members of a certain class of morphemes, the affixes, display idiosyncratic distributional properties. The suffix *-ity,* for example, attaches only to an adjective and with it forms a noun: *obesity* = $_{N}[_{A}[\text{obese}]_{A}\text{ -ity}]_{N}$. The suffix *-ify,* on the other hand, always constitutes a verb along with a sister adjective or noun: *codify* = $_{V}[_{N}[\text{code}]_{N}\text{ -ify}]_{V}$, *purify* = $_{V}[_{A}[\text{pure}]_{A}\text{ -ify}]_{V}$. The most obvious way of capturing these distributional peculiarities within a context-free rewriting system is to introduce each affix directly by a rule specific to it, as in (1.1), which means, in effect, positing a separate rule for every affix of the language:

(1.1) N → A *ity*
V → N *ify*
V → A *ify*

(Such a rule system would also involve rules rewriting the preterminal categories N, A, V with elements of the terminal vocabulary; e.g., N → *code, object, boy,* etc., A → *pure, nice,* etc.) This treatment of affixes is inadequate, as we will see in section 3.4, because it fails to assign affixes a categorial status and to capture generalizations about possible word structures in a direct way.[6]

The alternative to this purely context-free generation of word structure that I will defend here is not subject to these same objections. It involves a "mixed" theory of morphology analogous to the *Aspects* model of the syntactic base component. The morphological component is seen as consisting of a set of context-free rewriting rules (the *word structure rules*), which (like the phrase structure rules of the *Aspects* model) do not introduce elements of the terminal string; a list of lexical items, including affixes and other bound forms (the *extended dictionary*); and a *(morpho)lexical insertion transformation.* The word structure rules are roughly of the form (1.2) or (1.3). (The term *Af* is a temporary expedient.)

(1.2) a. P → φ Q Af Ψ
b. P → φ Af Q Ψ
c. P → φ Q R Ψ

(1.3) a. P → Δ
b. Q → Δ
c. R → Δ

where P, Q, R stand for individual category symbols, φ and Ψ are variables over category symbols (including Af), and Δ is the "dummy" symbol (cf. Chomsky (1965))

Such a rule system generates labeled trees with terminal strings consisting of dummy elements.

The extended dictionary lists all of the lexical items of the language, including the affixes. The (morpho)lexical insertion transformation completes the structures generated by the rewriting rules by inserting items from the dictionary, subject to the (lexically specified) conditions that the particular items may impose. Among these restricting conditions are the category of the item itself (which must be nondistinct from that of the preterminal category dominating the Δ for which the item substitutes) and its subcategorization frame (the sister category of the item in word structure must be nondistinct from that specified in an item's subcategorization frame).

On this theory, the idiosyncratic properties of an affix are listed as part of its lexical entry. These properties of the affix include:

(1.4) a. Its category (involving a specification of its type (the level *Af*) and of its categorial features, syntactic and diacritic);

b. Its subcategorization frame (involving a specification of the category to which the affix may be sister in morphological structure);

c. Its meaning (usually a characterization of what sort of function it is; cf. section 2.3);

d. Its phonological representation (minimally, a distinctive feature matrix), as well as a list of its other idiosyncratic phonological properties.

The first two specifications govern the distribution of the affix in morphological structure: an affix morpheme and its dominating category α substitute for a particular affix category β of a tree generated by the word structure rules on condition (i) that α be nondistinct from β and (ii) that the sister of β in that tree satisfy the lexical subcategorization form of the affix morpheme in question. The treatment of affixes here is entirely analogous to the *Aspects* treatment of verbs and their distribution in S-syntactic structure.

This, then, is the model that I propose for the rule system generating word structure. It embodies the claim that morphological structures are labeled trees with possible self-embedding. It also embodies the claim that affixes belong to a morphological category. However, this model implies nothing more specific about the nature of morphological structure. Any further specification of the general properties of morphologi-

cal structure involves, first, a theory of the categories of morphological structure and, second, a theory of the possible relations between categories in morphological structure.

1.2. $\bar{X}$ Theory in Word Syntax

Beyond the assertion that the rule system appropriate to generating word structure is a context-free grammar, there is another claim: that certain notions of $\bar{X}$ theory, a theory of S-structure (cf. Chomsky (1970), Bresnan (1976), Jackendoff (1977)), are required for an insightful characterization of W-structure. We must isolate two basic ideas of $\bar{X}$ theory. The first is that, formally speaking, a syntactic category is a pair $(n, \{F_i, F_j, \ldots\})$ consisting of a category *type* or *level* specification n (the number of "bars" of the category) and a feature specification $\{F_i, F_j, \ldots\}$, where F_i is a syntactic or morphological feature. I will call the feature specification the category *name*. Thus, in $\bar{X}$ theory, the symbol X (or Y, Z, etc.) is a variable standing for the set of category names Adjective (A), Noun (N), Verb (V), Preposition (P), etc., and the superscript integer (or, equivalently, the number of "bars") defines the level or type of category. The syntactic word is a category of type zero (it is the "lowest" category of syntactic structure). Words of the category Noun, Adjective, Verb, etc., will thus have the category symbol N^0, A^0, V^0, respectively (or, more simply, N, A, V). The class of words itself is designated by the symbol X or X^0 (or some other upper-case letter). Categories of level X^1 and higher are phrases. X^1, for example, is the category level which dominates the head X and its complements, such as a verb and its direct object (e.g., $_{V^1}[_V$[devoured]$_V$ [the sandwich]]$_{V^1}$) or a noun and its prepositional complement (e.g., $_{N^1}[_N$[facts]$_N$ [[about] [the case]]]$_{N^1}$).

The second basic idea of $\bar{X}$ theory, intimately related to the first, is that the phrase structures of language conform to certain restrictive patterns, the characterization of which requires the $\bar{X}$ theory of categories. Specifically, the hypothesis, first put forward by Chomsky (1970), is that phrase structure rules conform in general to a schema such as (1.5):

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

That is, every syntactic category dominates a category bearing the same name, but one level down in the $\bar{X}$ hierarchy. This amounts to the

claim that all S-structures have a head.[7] In what follows, I will extend these two basic ideas to the area of W-syntax.

I will defend the claim that W-syntactic (i.e., morphological) categories are entities that are formally identical in character to syntactic categories, which is to say that each morphological category is, formally speaking, a pair $(n, \{F_i, F_j, \ldots\})$. Note that it is already necessary to view *some* morphological categories as identical to syntactic categories, for the rules of the W-syntax in fact share a set of categories with rules of the syntax—the word-level categories Noun, Verb, Adjective, etc. My hypothesis is simply that all W-syntactic categories, be they of the type Word or "lower" than Word, are in the $\bar{X}$ hierarchy. (It also makes available the possibility that word structure rules may be formulated in terms of these different types.) In principle, it could turn out that Word and only Word is the (recursive) category type at play in language. As we will see, however, a theory permitting only this type is not sufficiently differentiated to allow for the expression of a fair array of linguistic generalizations in various languages, while a limited extension of the theory of morphological category types does provide a means of expressing them. A case can be made, for example, for the existence of a type X^{stem} (where *Stem* is simply a convenient term for the type X^{-1} that is one down in the $\bar{X}$ hierarchy from Word $(= X^0)$) and for seeing Stem as a recursive category type. A case can also be made for a yet lower (recursive) category level X^{root} (or X^{-2}) contained within Stem.[8] The category X^{af} (for *Affix*) is also required, though its position in the system is somewhat special, in that it is not ordered within this hierarchy (it cannot be assigned an integer) and is to all appearances preterminal. (See section 3.4.)

The features which play a role in word syntax (i.e., form part of W-syntactic categories) can be assigned to two classes: (i) the *syntactic category features* [±Noun], [±Verb], etc., which represent the distinctions among Noun, Verb, Adjective, Preposition, Adverb, etc. (cf. Chomsky (1970), Jackendoff (1977)[9]), and (ii) all of the others, which will be termed *diacritic features*. The diacritic features include those relevant to the particulars of inflectional and derivational morphology. The inflectional features might include, for example, conjugation or declension class markers, features for tense (e.g., [±past]), gender (e.g., [±feminine]), person, number, and so on. The derivational features may include ones such as [±latinate] (cf. Aronoff (1976), Williams (1981a)) and [±learnèd] (cf. Dell and Selkirk (1978)). The terms *inflectional* and *derivational* are meant only to provide a loose classification,

for the systems of derivational and inflectional morphology are not strictly disjoint. An inflectionally marked element, for example, may serve as the base for derivational processes (cf. Lieber (1980)). Below we will see evidence that the categories of word structure, both preterminal and "higher" in the tree, must be specified in terms of features, both diacritic and syntactic. Of particular importance is the claim that such features are associated with affixes, for this amounts to the claim that affixes have a categorial status.

There seem to be severe limitations on the possible relations between the type of dominating category and the types of categories it dominates. First, major constituents of the syntax do not appear within morphological structures generated by the word structure rules. Nouns such as *ne'er-do-well, speak-easy, will-o'-the-wisp,* with apparently syntactic structure, are exceptions and are not representative of general processes of word formation. As for the constituents of compound words which display an apparently syntactic phrasal structure, such as the left-hand constituent of $_{N}[_{N}[_{A}[\text{American}]_{A}\ _{N}[\text{history}]_{N}]_{N}\ _{N}[\text{teacher}]_{N}]_{N}$, they can be viewed as mere compounds. (There is independent motivation for such a compound structure, as shown below.) Second, a morphological category of a higher level does not seem to appear in structures where it is dominated by a category of a lower level. For example, it does not seem possible to introduce a Word below the level of Root. These observations suggest that word structure rules, unlike phrase structure rules, are required by universal grammar to be of the form (1.6):

(1.6) $X^n \rightarrow \varphi\ Y^m\ \Psi$

where $0 \geqslant n \geqslant m$

In other words, a category may not be rewritten in terms of another category (or categories) higher than itself in the $\bar{X}$ hierarchy. In what follows, this interestingly restrictive hypothesis concerning the relation of category types in word structure will be assumed.[10]

It is in particular this (putative) characteristic of word syntax—that it is organized into levels (in the sense that a category lower in the hierarchy cannot dominate a higher one)—which distinguishes it from the syntax of sentences. Indeed, in $\bar{X}$ syntax, it is assumed that any nonhead category introduced on the right side of the arrow in a phrase structure rule is necessarily the *maximal projection* of the category (that is, the one having the highest possible level or type specification),

so that rewriting rules are of the following form (see, for example, Emonds (1976, chapter 1)):

(1.7) $X^n \rightarrow \ldots Z^m \ldots X^{n-1} \ldots$ *or* $X^n \rightarrow \ldots X^{n-1} \ldots Y^m \ldots$
where m is maximal

Should the claim about "level-ordering" in word syntax in fact hold up, then, it would provide important evidence that the systems of word syntax and phrase syntax are truly distinct.

The possible relations between the feature specifications of categories in word structure would also seem to be governed by general principles. For example, like syntactic structures, word structures tend to be "headed" (Williams (1981a)). That is, a W-syntactic constituent X^n with a particular complex of category features will contain a constituent X^m, its head, which also bears those features. Word structure rules thus apparently conform (in general) to the format (1.8), which is to be read as stipulating that each constituent contain a head.

(1.8) $X^n \rightarrow \varphi\ X^m\ \Psi$
where X is a variable standing for a complex of categorial features, both syntactic and diacritic

Taken together, (1.6) and (1.8) amount to the claim that the context-free rewriting rules of any grammar employing such a system conform to the format (1.9) or, put another way, that such a system is capable of generating the word structures of such languages:

(1.9) $X^n \rightarrow \varphi\ X^m\ Y^p\ \Psi$
$X^n \rightarrow \varphi\ Y^p\ X^m\ \Psi$
where $0 \geqslant n \geqslant m,p$

Of course, this format underdetermines the range of possible morphological structures in a language. One purpose of the following chapters is to show the need for particular statements in grammars concerning which specific categories have what composition in terms of other specific categories. We will see that any given language has a (particular) grammar of word structure (just as it has a particular grammar of phrase structure), one which nonetheless conforms to certain quite general principles governing possible word structures in language. We will see that these particular word structure rules mention all aspects of W-syntactic categories: syntactic features, diacritic features, and category types.[11]

1.3. The Place of Word Structure Rules in the Grammar

In principle, the question of the nature of the rule system that generates word structures is independent of the question of the rules' "location" in the grammar. In an earlier unpublished version of this work, I took the position that word structure rules were part of the system of base rules of the syntactic component. However, this position does not follow in any logical sense from the fact that word structure rules and phrase structure rules have the same general formal properties. It is merely consistent with that fact. Equally consistent would be a model of grammar according to which the word structure rules were "in the lexicon," that is, in an entirely distinct component of the grammar. It is not completely clear to me at this point what this distinction amounts to.

To keep the focus on the issues which I consider to be central here—the issues of the nature of word structure and the rule system for generating it—I will assume, along with the various other generative theories of morphology, that the rules of word structure form part of what one may call the *lexical component* or simply the *lexicon* (understood in a broad sense). As it is viewed here and in most earlier theories, the lexical component contains a variety of subcomponents. First, it contains a list of freely occurring lexical items (which I will assume to be words, in English). We may call this the *dictionary* (or *lexicon,* in the restricted sense). Second, it contains a list of the bound morphemes of the language. This, together with the dictionary proper, I will call the *extended dictionary*. Third, the lexical component includes the set of rules characterizing the possible morphological structures of a language, the *word structure rules* of the present theory. The word structure rules, along with the structures they define, are the central concern of this monograph. Together with the extended dictionary, they form the core of the *word structure component* of the lexicon or, shall we say, the *morphological base*. Other divisions within the morphological subcomponent have been proposed, including allomorphy rules (Aronoff (1976)) and morpholexical rules (Lieber (1980)), but these will not be important in our discussion.

Where this monograph parts company with previous studies on morphology in the generative framework is in its concern for questions of word syntax, and in the explicitness of the proposal concerning the mechanism for generating word structure. With few exceptions,[12] researchers have given little attention to these issues. For the most part,

"morphological rules" have been stated relatively haphazardly, with no particular emphasis on the nature of the rules themselves. Presumably the issues were not considered particularly important, though I hope that the reader will judge, with me, that they are in fact significant and well worth pursuing.

One characteristic that distinguishes morphology from syntax, to be sure, is the fact that many of the entities defined as well formed by the rules of morphology are fixed expressions. Most words we speak and understand we have heard before, while sentences are for the most part novel to us. More precisely, what distinguishes words from sentences is that most words are in the dictionary.

There are a number of reasons for saying that the list of items called the dictionary forms part of the speaker's knowledge (or grammar) of a language. First, speakers have intuitions about what is or is not an actual word of the language (as well as intuitions about what constitutes a possible word of the language). Second, and more important, the individual characteristics of words are not always predictable. The meaning of a simple word is totally unpredictable, and even in the case of complex words the meaning often cannot be predicted on the basis of its component parts. The conclusion is that a word and its (idiosyncratic) meaning must be paired in a list; that list is the dictionary. (See Aronoff (1976), who develops this point at length.) Phrases whose meaning is not compositional—that is, those phrases that are usually called *idioms*—will also have to be listed in the lexicon.

Consider now the fact that the multimorphemic words of this list must be said to have an internal structure. Speakers have intuitions about the structure of existing words of their language. These intuitions are presumably based on their knowledge of the word structure rules of the language, and indeed it seems that in general the existing lexical items of a language (more exactly, the words of the lexicon) have structures generable by the morphological component of the language. But the word structure rules cannot be viewed as generating these words anew each time they are used, for this contradicts the notion that they are listed; no distinction would then be drawn between existing and possible (or newly generated) words, and no means would be available for representing their idiosyncratic, noncompositional features. In the case of existing lexical items, then, it would seem appropriate to view the word structure rules as redundancy rules or well-formedness conditions on lexical items. More generally, it seems possible to impose the following condition:

(1.10) For every word of the language, there must exist a *derivation* via the word structure rules of the language.

This condition allows us to treat existing words and possible words in uniform fashion. If a word (existing or possible) is to be well formed, its structure must be among those generable by the word structure rules of the language.[13]