

Syntactic words and morphological words, simple and composite*

Arnold M. Zwicky

1. SYNTAGMATIC UNITS AND PARADIGMATIC UNITS

What is the relationship between the simple or elementary objects of grammar, word-like things, and its composite or complex objects, phrase-like things? I focus here on a small piece of this gargantuan topic, having to do specifically with the elementary objects of morphosyntax, rather than with the grammar as a whole: 'syntactic words', the syntagmatic units I will call *Ws*; and 'morphological words', the paradigmatic units I will call *moremes*.

The distinction at issue is a familiar one – it is made clearly, though not with this terminology, in careful discussions of the notion of *word*, such as those in Lyons (1968: sec. 5.4) and Matthews (1974) – but for some reason generative grammarians have for the most part failed to take the distinction seriously, preferring instead to use references to 'X⁰' units as if the small objects of syntax and the large objects of morphology have the same status, in fact, as if they coincided with one another. But they are objects of quite different character – the former are expression tokens, the latter are expression types – and the question of whether they are in some sense coincident with one another is an empirical question, to be decided by considering a wide range of problematic data, not via a terminological or notational stipulation. Indeed, familiar data suggest quite strongly that coincidence is merely the default relationship between the objects of syntax and morphology.¹

The complications that I am addressing in this paper are two: that both *Ws* (section 3) and *moremes* (section 4) can themselves be simple or composite, a fact that might at first seem problematic for the distinction between word-like and phrase-like units; and (section 5) that the interface between morphology and syntax can match, or instantiate, a single *moreme* by a sequence of two or more *Ws*, or a single *W* by a composite of two or more *moremes*, again a fact that might at first seem problematic for the distinction between word-like and phrase-like units.

An important side issue is the question of how we could detect the various sorts of word-like units. A standard tool for picking out 'words' is the *intervention constraint*, a condition prohibiting the interruption of two adjacent expressions (within a 'word') by other material. I will be claiming that there are at least three

Author's address: 722 Ramona St., Palo Alto, CA 94301, U.S.A.

different types of morphosyntactic intervention constraints, each with its own characteristic profile of possible and impossible interruptions.

Two of the types arise from conditions on syntactic rules – from a condition that some constituent covered by a rule must be a **W** (section 3.1), or from a condition that two constituents covered by a rule must be immediately adjacent to one another (section 3.4). There is no general prohibition, however, against interrupting a **W**. On the other hand, there is a general prohibition against interrupting a moreme, and this is the source of the third, and most stringent, type of intervention constraint (section 4.4).²

As prelude to the discussion in sections 3 through 5, I provide in section 2 a pretheoretical sketch of the morphosyntactic portion of a grammar.

2. SYNTAX AND MORPHOLOGY

The morphosyntactic portion of a grammar concerns itself with *expressions* of a language, where an expression is a pairing of semantic content with phonological shape. More specifically, it concerns itself with expression types that have the potential to be instantiated³ as free-standing expression tokens. I will refer to such an expression type as a *tacteme*.

2.1. Sentemes and moremes

Tactemes come in two major varieties, which we can think of crudely as big ones and little ones.

Big tactemes are the initial objects of description in syntax, essentially Bloomfield's (1933: ch. 11) 'sentence types': the expression types realizable as free-standing expressions with illocutionary semantics. I will call them *sentemes*.⁴ Each expression in (1) instantiates a different senteme of English.

- (1) a. In the garden with a hammer.
- b. Why not give up?
- c. Me play the saxophone?!
- d. Be quiet, my love.
- e. I am asking you.

Little tactemes are the objects of description in morphology. I will call them *moremes*.⁵

Some complication is introduced into these matters by the fact that moremes often come in a number of (inflectional) *forms*, and that particular forms often come in a number of phonological *shapes*, so that moremes are two degrees more abstract than the little chunks of stuff that in some sense occur within sentemes. For instance, the English verb moreme JUMP has several forms with the shape *jump*, several with the

shape *jumped*, and several with the alternative shapes *jumping* and *jumpin'* (not to mention the fact that each of these has alternative shapes with various tones: falling as in *I am jumping*, level as in *I have been jumping since Tuesday*, rising as in *Am I jumping?*).

2.2. Rules and descriptions

The units of description in both morphology and syntax are *rules*.⁶

In morphology, there are (at least) rules embodying generalizations about how one set of moremes is related to another set (*derivational rules*) and rules embodying generalizations about the properties of a moreme: rules embodying generalizations over the set of stems for a moreme (*stem rules*), rules embodying generalizations over the set of forms for a moreme (*inflectional rules*), rules embodying generalizations over the set of shapes for a form (*shape rules*), and rules embodying generalizations about how various phonological, morphological, syntactic, or semantic properties of a moreme are related to others (different types of *lexical redundancy rules*).

In syntax, there are (at least) rules embodying generalizations about the makeup of sentemes (an English rule stipulating that a declarative senteme uses a clause with finite head, a nominative subject, and cross-referencing of subject properties on the head), rules embodying generalizations about the makeup of constituents with particular properties (the English Subject + VP rule, stipulating that a clause can comprise two constituents, the second being a VP compatible with the first constituent as its subject), and rules embodying generalizations about the compatibility of heads with a list of dependents, each with specified properties (an English rule permitting a head to occur with exactly two argument constituents, a subject and a direct object).

A complete syntactic *description* of some expression-token E in a language is then an assemblage of all syntactically relevant information about E and its parts. Such a description specifies (perhaps with considerable redundancy) all the properties of E that are potentially relevant to E's ability to occur as a part of other expressions of the language, it lists all the subexpressions of E that contribute in a regular fashion to E's form and meaning (that is, all the *constituents* within E), and it specifies all the properties of these subexpressions that are relevant to their occurrence within E.

3. SYNTACTIC WORDS

It is clear that among the syntactically relevant properties of subexpressions are some having to do with the relative 'size' of constituents. We are not surprised to find generalizations that are sensitive to the difference between 'included' and 'including'

constituents – generalizations that, for instance, treat a VP constituent of a VP expression differently from the VP expression including it.

But some such distinctions are not just quantitative, but constitute qualitative distinctions to which virtually every generalization in syntax is sensitive. These are the distinctions of ‘rank’: *clause* C, the largest rank; *word* W, the smallest; and *phrase* P, the distinctive intermediate rank.

Generative syntacticians have come to speak of rank distinctions as matters of ‘bar level’, and a considerable theoretical tradition has developed around bar level. My concern is neither to advance nor to reject this tradition. My claims are pretheoretical: that

- (2) Any adequate theory
- A. must incorporate some version of at least the distinction between W, P, and C as properties of constituents;
 - B. must characterize the prototypical non-W constituent as composed of a head constituent and stipulated dependents (arguments or modifiers); and
 - C. must allow for rank-shifts in particular languages, according to which certain constituents of one rank are stipulated to be available for certain syntactic functions prototypically associated with constituents of another rank.

In particular, an adequate theory must allow for certain types of C functioning as arguments of a V (as complement clauses, serving in a prototypical NP function), modifiers of an N-type constituent (as relative clauses, serving in a prototypical AdjP function), or modifiers of some other type of constituent (as adverbial subordinate clauses, serving in a prototypical AdvP function).

3.1. Intervention constraints of type 1

Claim (2A) embodies the observation that syntactic rules routinely require constituents in particular slots of a construction to be of a stipulated rank – W or P or C, depending on the rule and the slot.

For instance, the initial slot in the English WH Question construction must be filled by something of rank P (and category other than V). And constituents of rank W are called for in a number of English rules: the first slot in the Subject Auxiliary Inversion construction must be filled by an expression of rank W (*will* or *won't*, but not *will not* or *will soon* or *soon will*); the first slot in the WH Cleft construction must (for many speakers) be filled by an expression of rank W (*what* or *where*, but not *which one* or *where in Paris* or *whose book* or *near where* or *from where*, though all of these are available for the corresponding slot in the WH Question construction); the slot to which the postmodifier *else* is attached must be filled by an expression of rank W (*anyone* or *someone* or *everywhere* or *what*, but not *any person*

or *some day* or *everywhere nice*); and the second slot in a V+P transitive construction must be filled by an expression of rank W (*up* or *off* or *on*, but not *right up*: *give right up*, *give the fight right up*, but **give right up the fight*).

Whenever a syntactic rule calls for a constituent of rank W, a type of 'intervention constraint' is induced: dependents (whether modifiers or arguments) of this constituent will not be able to intervene between this constituent and any adjacent constituent (because then this constituent would be of rank P rather than W), though dependents of the adjacent constituent will normally be able to intervene (since they will be part of a constituent that is not obliged to be of rank W), as will loose-construction modifiers. I will refer to this configuration of possibilities as an intervention constraint of *type I*. English WH Clefts will illustrate these observations:

- (3) a. Typical example:
Where + tourists meet in Paris is Notre-Dame.
- b. Intervening dependent within crucial constituent:
**Where in Paris + tourists meet is Notre-Dame.*
- c. Intervening dependent within adjacent constituent:
Where + happy tourists meet in Paris is Notre-Dame.
- d. Intervening loose-construction modifier:
Where, as you know, happy tourists meet in Paris is Notre-Dame.

3.2. Word-phrases

Claim (2B) is compatible with the existence of constituents of rank P that happen to comprise only a head constituent, where that head is itself of rank W: NPs like *kangaroos*, VPs like *vanished*, PPs like *in*, AdjPs like *happy*, AdvPs like *carefully*. What we have here are expressions that happen to have the property of being of rank P and the property of being of rank W. There is no inconsistency in this. Grammatical generalizations that call for rank W constituents will apply to such expressions, and so will grammatical generalizations that call for rank P constituents.

This is the first way in which 'words' (in one sense or another) can be phrases; these are *word-phrases*.

3.3. Phrase-words

Claim (2C) is compatible with the existence of rank-shifts in which expressions of rank P (or even C) function syntactically like expressions of rank W. The result would be a second way for Ws to be phrases – what I will call (following Bloomfield 1933: 180) *phrase-words*.

This is the analysis I would suggest for the first elements in 'phrasal compounds' like *slept all day look* and *God is dead theology*, cited by Booij (1990) (with more discussion and references to be found in Dressler 1988 and Hoeksema 1988). These first elements have the internal syntax of constituents of rank P or C but the external

syntax of a W, like the first elements of such garden-variety compounds as *liberation theology* or *Dracula look*.

I suggest a phrase-word analysis for two further constructions in English. In both, phrasal expressions with V heads function as heads taking direct object arguments, that is, they have a prototypical VW function. The two constructions I have in mind involve head V combined with Prt, that is, with an adverbial P (*send away* in *send away the money*); and head V combined with indirect object NP (*give Pat* in *give Pat money*).⁷

The implicit claims are that there are (at least) three separate constructions involving V and Prt (intransitive combinations, as in *Ronnie ran away*; separated transitive combinations, as in *We sent the money away*; and contiguous transitive combinations, as in *We sent away the money*), only the last of these of concern here, and that there are (at least) two separate constructions involving V combined with indirect object (those with a separated PP, as in *We gave the money to Pat*, and those with a contiguous NP, as in *We gave Pat money*), only the second of these of concern here. These claims are supported in part by well-known facts – that somewhat different subcategories are involved in each of these combinations (for instance, that the V and Prt possibilities are not quite the same for the separated and the contiguous transitive combinations), and that the different combinations are subject to different conditions (for instance, that the contiguous transitive combination requires a Prt of rank W, and so calls up an intervention constraint of type I, while the other two combinations permit a Prt of rank P, that is, a modified Prt: **We sent right away Pat* versus *We sent Pat right away* and *Ronnie ran right away*).⁸

A similar treatment seems to be appropriate for French ‘causative clause union’ examples like *fait partir* in *Je fait partir Jean* ‘I made Jean leave’ and might be advanced as well for the German ‘verb-clusters’ discussed by Bierwisch (1990).

In all of these examples, what I am proposing is that some combination of material acts as a W syntactically. There is no necessary claim that such a combination instantiates a moreme; being word-like syntactically does not entail being word-like morphologically, though the standard situation is, of course, for a W to instantiate a moreme (as for the ordinary compound nouns I discuss in section 4.3 below).

Note that acting as a W syntactically does not in itself call up an intervention constraint. There is no general principle of ‘W integrity’, parallel to the principle of moreme integrity (to be illustrated in section 4.4 below) that usually goes under the name of ‘lexical integrity’. Insofar as some part of a phrase-word satisfies the conditions for a syntactic construction, it can participate in that construction. For instance, NP indirect objects can be ‘passivized’ (*Pat was given the money*) and (for some speakers) ‘extracted’ (*Pat I gave the money, Who did you give the money?*).

3.4. Intervention constraints of type II

Given a constituency or senteme rule with two or more slots, there are several possible conditions the rule might place on the temporal ordering of the expressions that fill its slots.

The rule might place no such condition on a pair of slots N_1 and N_2 , in which case expressions filling them can occur in either order with respect to one another. Or the rule might stipulate that an expression filling N_1 precedes an expression filling N_2 ; if there are still other slots, then N_1 could be separated from N_2 by expressions filling these other slots. Or the rule might stipulate that an expression filling N_1 immediately precedes N_2 (Zwicky & Nevis 1986, Ojeda 1988), in which case no such intervening material is possible; we then have an intervention constraint of type II. Note that such a constraint does not necessarily involve word-like units at all (neither Ws, as in intervention constraints of type I, nor moremes, as in intervention constraints of type III).

I suggest that a number of English constructions involve immediate-precedence conditions and therefore give rise to intervention constraints of type II: the first slot in Subject Auxiliary Inversion immediately precedes the second (*Is Dana obviously the culprit?* and *What did Alexis apparently see?*, but **Is obviously Dana the culprit?* and **What did apparently Alexis see?*); and the V slot in transitive VPs immediately precedes, at least in the default case, the direct object slot (*put the box on the counter* and *word the letter carefully*, but **put on the counter the box* and **word carefully the letter*).

In an intervention constraint of type II, dependents of *either* constituent can intervene between them so long as such dependents are parts of these constituents, but not otherwise; and loose-construction modifiers can intervene between them, though usually only with some degree of awkwardness. English transitive VPs will illustrate these observations:

- (4)
- a. Typical example:
put + boxes on the counter quickly
 - b. Intervening dependent within first constituent:
put down + boxes on the counter quickly
 - c. Intervening dependent within second constituent:
put + heavy boxes on the counter quickly
 - d. Intervening sister of the two constituents:
*put + on the counter + boxes quickly
*put + quickly + boxes on the counter
 - e. Intervening loose-construction modifier:
?put, as I had asked, boxes on the counter quickly

4. MOREMES

So far, I have distinguished certain classes of expression types (tactemes, of both the senteme and the moreme variety) from certain classes of expression tokens (constituents, of clause, phrase, or word rank). Sentemes are composed of moremes in roughly the same way that substance types like NaCl are composed of elemental substance types like Na and Cl. Clauses are composed of phrases, and these of words, in roughly the same way that a chunk of salt is composed of salt crystals, and these of salt molecules, and these of sodium and chlorine atoms.

What relates the two sorts of linguistic analysis is *instantiation* of types by tokens. Sentemes, for instance, are instantiated as (sequences of) constituents. Sentemes are not necessarily instantiated as single constituents; (1a), *In the garden with a hammer*, is not a single constituent. Sentemes that are instantiated as single constituents are not necessarily clauses; *In the garden* is a PP, casual-style *Saw Ronnie yesterday* is a finite VP. Clauses do not necessarily instantiate sentemes; *they be happy* in *I insist they be happy* is such a clause.

4.1. *The morphology-syntax interface*

Similar complexities of instantiation arise for moremes. Just as sentemes are instantiated as (sequences of) constituents, moremes are instantiated as (sequences of) Ws. The ordinary arrangement is for one W to instantiate one moreme, but as we shall see there are circumstances in which a sequence of Ws instantiates a moreme, and such a moreme-instantiating expression is not always a constituent (just as a senteme-instantiating expression is not always a constituent).

There is a further complexity here introduced by the existence of constituents of rank W that include constituents of rank W – phrase-words like *give Pat* in *give Pat money*, as discussed above, and syntactic compounds like *television table*, to be discussed below. It is the ultimate constituents of rank W, the *minimal Ws* – *give* and *Pat* in the W *give Pat*, *television* and *table* in the W *television table* – that serve as the interface between the requirements placed by syntax on expressions and those placed by morphology. What morphology provides is the list of moremes available for instantiation as (sequences of) minimal Ws. To be instantiable as a chunk of stuff, a moreme must have properties compatible with those called for by syntactic rules relevant to that chunk. The properties in question include syntactic category and subcategory (GIVE can be instantiated as a VW combining with an indirect object NP, as in *give Pat*, but DONATE cannot), the availability of particular forms (WALK can be instantiated within the perfect construction in English, as in *have walked to Canada*, but STRIDE cannot, since it lacks the past participle form this construction calls for), and the availability of particular shapes (to be illustrated below).

The conditions placed on expressions by syntactic rules are not moreme-specific at all (except insofar as it might turn out in some circumstances that there is only

one moreme with the appropriate properties for instantiation at some point within an expression). Strictly speaking, in discussions of syntax alone, we should be citing the templates of expression properties that are licensed by the rules of syntax, rather than citing expressions with moremes instantiated within them.⁹ An expression, as opposed to such a template, is licensed:

- (5) A. If it has properties satisfying the conditions in all the relevant syntactic rules (except insofar as its properties are 'listed', as aspects of an idiosyncratic syntactic pattern); and
 B. If every minimal W in it instantiates one, and only one, moreme; and
 C. If the properties of each of these moremes are compatible with the properties of the W(s) instantiating it; and
 D. If each of these moremes has properties satisfying the conditions in all the relevant morphological rules (except insofar as its properties are 'listed', as aspects of an idiosyncratic morphological pattern).

Since condition (5A) says that an expression must satisfy all the relevant syntactic rules and condition (5D) that its parts must satisfy all the relevant morphological rules, I will speak of them as enforcing *co-satisfaction* of the two sets of rules. Condition (5B) is a version of the 'total accountability' criterion of the structuralists (Hockett 1947: sec. 14). And condition (5C) is a static version of the 'lexical insertion' operation of classical transformational grammar (Chomsky 1965:122: "Lexical entries substitute for the lexical categories of a preterminal string").

The usual conceptualization of the roles of syntax and morphology in describing expressions is rather different from this one. The standard story, in generative circles at any rate, is that syntactic and morphological rules describe objects of the same type. Generative morphologists differ as to whether these rules comprise a single component of grammar or, as in the 'autolexical' approach of Sadock (1980, 1985) and the 'coanalysis' proposal of Di Sciullo & Williams (1987), two (or more) distinct components that impose largely compatible (but sometimes incompatible) analyses on the same material. I have observed that the (syntagmatic) objects described by constituency and compatibility rules in syntax have a different character from the (paradigmatic) objects described by derivational rules in morphology – the former are expression tokens, the latter expression types – so that some sort of coanalysis is necessary (though undoubtedly not the sort that either Sadock or Williams has in mind).

4.2. Derivation

The subject matter of morphology is the list of moremes of a language – what I will call the *morphicon*.¹⁰ Morphological rules can be thought of as capturing generalizations about the contents of this list, just as syntactic rules can be thought of as capturing generalizations about the contents of the list of sentemes in a language.

An individual moreme is an assemblage of properties: semantic, pragmatic or stylistic values, syntactic category and subcategory, specifically morphological properties like paradigm class, and a variety of types of information about phonology. The phonological information comprises a stem list, providing phonological material for derivational and inflectional rules, and a form/shape list, a two-dimensional assemblage of information providing phonological material corresponding to each coherent and relevant combination of grammatical categories, and alternative phonological realizations for each of these according to their morphosyntactic contexts. The form/shape list for the moreme BE, for instance, provides the shapes *is* (the default shape) and *'s* (the enclitic shape) for the third-person singular present form, while for the third-person singular past form it provides the default shape *was*, corresponding to *is*, but provides no enclitic shape at all, nothing corresponding to *'s*.

Once again I must emphasize that I am not proposing a theory here, in particular, any sort of theory of representations for moremes and the morphicon. Rather, my intention is only to catalogue the sorts of information and the sorts of generalizations that any adequate theory of morphology must find a place for.

A derivational rule predicts (some of) the properties of one set of moremes on the basis of (some of) the properties of another. The phonological side of this prediction will be of some concern to us here. In particular, an ordinary derivational rule predicts some stem for a 'derivative' moreme on the basis of some stem for a 'source' moreme. The stipulated stem of the source moreme might be used unaltered – this is 'zero derivation' – or as altered by one or more phonological operations on it, for instance by suffixation of some material to it.

At least one more degree of freedom is available to derivational rules, however. Rather than using a stipulated stem, a rule can use a stipulated form (the plural form of a noun, the feminine singular form of an adjective, or whatever) of the source moreme, in its default shape. The stipulated form might be used unaltered (as when past participle forms of English verbs are used unaltered as adjectives), or again it might be used as altered by one or more phonological operations on it.

4.3. Compound moremes and syntactic compounds

Moremes can also be composite, in the sense that they can be derivatives based on two or more source moremes. A rule describing a class of these *compound moremes* predicts some of their properties on the basis of some of the properties of source moremes.

As for the phonological side of such a rule, it posits slots for the source moremes, enforces a temporal ordering on phonological material filling these slots, distinguishes one or more of the slots as *locus* (where grammatical categories belonging to the compound moreme as a whole will be realized phonologically), and specifies phonological properties for each of the slots. Potentially, each slot could have its phonological properties stipulated in the same way that an ordinary derivative

moreme has the phonological properties of its stem stipulated: as an unaltered stem of the source; as a stem of the source, altered by phonological operation(s); as an unaltered form of the source; or as a form of the source, altered by phonological operation(s). It is customary, and entirely reasonable, for locus slots to use a stem (rather than a stipulated form), unaltered.

German compound nouns provide illustrations of three of the four possible customary arrangements. The compounds in question are all locus-final.

The first slot in one type is filled by an unaltered noun stem (*Wortbildung* 'word building, derivational morphology'), in a second type by a noun stem altered by suffixation (*Arbeitszeit* 'work time'), and in a third type by the plural form of the source noun (*Bücherfolge* 'series of books').

What makes these facts germane to the morphology-syntax interface is that some compound moremes are syntactically as well as morphologically composite. They are *syntactic compounds*, things of rank W with two or more constituents of rank W.

The W constituents of such expressions cannot be very 'active' syntactically. Most syntactic rules, in particular 'extraction' rules, call for constituents of rank P in at least one of their slots; such a constituent might happen also to be of rank W, or the rule might even stipulate that it has to be of rank W as well as of rank P (as the English WH Cleft rule stipulates for its first slot), but such a rule cannot 'go into' a syntactic compound, since the compound's constituents are not of rank P. As we shall see shortly, alternative orderings of the constituent Ws and the addition of dependents to either of these Ws, or between them, will be prohibited by the fact that these syntactic compounds instantiate compound moremes and so must satisfy the requirements imposed by derivational rules on such moremes, as well as the requirements imposed by syntactic rules.

What seems to be left is just coordination, as in *a television or microwave table* and *a television table or shelf*; ellipsis, as in *You prefer the Bloomfield approach, but I favor the Sapir*; and external modification, as in the Greenlandic incorporation examples discussed by Sadock (1980).

4.4. Intervention constraints of type III

Whenever a sequence of two or more Ws in an expression instantiates a moreme, a particularly strong type of intervention constraint (one of type III) will be induced.

Since two adjacent Ws within this expression must satisfy all the conditions in relevant derivational rules (as well as all conditions in relevant syntactic rules), and since derivational rules have their own slot requirements, each slot being filled by phonological material standing for a particular moreme, there is, in general, no place for dependents of the Ws within the instantiation of the moreme in question, nor any place for intervening loose-construction modifiers. English compound nouns will illustrate these observations:

- (6)
- a. Typical example:
television + tables
 - b. Intervening dependent within first participant:
*television from Japan + tables
 - c. Intervening dependent within second participant:
*television + big tables
 - d. Intervening loose-construction modifier:
*television, or so I think, tables

Intervention constraints of type III are the ones customarily appealed to in textbook discussions of the criteria for (syntagmatic) wordhood. Since inflectional and derivational affixes are proper parts of an expression that instantiates a simplex moreme, neither can be separated from its stem:

- (7)
- a. Typical examples:
jump + ing kind + ness
 - b. Intervening dependent within base participant:
*jump up + ing *kind enough + ness
 - c. Intervening loose-construction modifier:
*jump, as I thought, ing *kind, as I expected, ness

These are minimal Ws, of course, so that (unlike syntactic compounds) their parts are unavailable for coordination (**jump- and singing*) or ellipsis (**were jumping before sing-*, **had jumped before others were -ing*).

5. SUPERMOREMES

Moremes are not always instantiated by syntactic constituents, much less constituents of rank W. There are large morphological objects – *supermoremes*, as I will call them – that are instantiated by expressions comprising two or more Ws, where these Ws do not (in general) together make a syntactic constituent.

What I have in mind here are, among other things, certain ‘clitic groups’, ‘serials’ (in particular ‘serial verbs’), and ‘portmanteaus’. There is, for instance, an English clitic group instantiated by the expression *Pat’s* in *Pat’s seen it* and comprising the syntactic Ws *Pat* and *’s* (an alternative shape for the third-person singular present form *has* of the moreme HAVE). And there is an English serial instantiated by the expression *go see* in *I will go see who’s at the door* and comprising the syntactic Ws *go* and *see*. Both examples are composites, with two source moremes each.

And there is the French portmanteau instantiated by the expression *du*, as in *du pain* ‘of the bread’, and comprising the syntactic Ws *de* and *le*. This is a simplex supermoreme – simplex from the point of view of morphology, though expressions

instantiating it comprise two Ws and so are composite from the point of view of syntax.

Clitic groups, serials, and portmanteaus are three further ways, in addition to word-phrases, phrase-words, and compounds, in which some sort of word-like thing (here, a supermoreme) can be some sort of phrase-like thing (here, a sequence of Ws).

5.1. Characteristics of supermoremes

Supermoremes and the derivational rules that describe them (*superderivational rules*) can be special in at least six ways, in addition to the syntactically exceptional instantiations that supermoremes have. I am not claiming that each supermoreme class in a language will exhibit all six characteristics, only that each can be expected to exhibit one or more of them.

First, a superderivational rule can, in effect, have a variable across categories in one or more of its slots. Second, such a rule can, in effect, have a variable over forms in one or more of its slots. And third, it can call for particular shapes of forms. For instance, the superderivational rule for English enclitic expressions like *realized's* allows any moreme (of whatever category) in its first slot, allows any form for the moremes that fill both slots, and stipulates a special shape for the forms it uses in its second slot (comprising just the final consonant of the default shape for the form).

Fourth, composite supermoremes can have variable numbers of source moremes, as in the serial expression *Go run see who's at the door*, or in the clitic groups of Tagalog. Fifth, like composite moremes in general, they can constitute domains for the distribution of certain phonological properties. And sixth, composite supermoremes of the clitic-group type are parallel in very many ways to complex inflectional forms, so that such supermoremes can be viewed as hybrids between compounding and inflectional morphology; these parallels are enumerated in section 3.11.2 of Zwicky (to appear).

5.2. Intervention constraints in supermoremes

Supermoremes are moremes, so that they will exhibit intervention constraints of type III, in much the same way that syntactic compounds do. Perlmutter's (1971: 95-97) discussion of English serial verbs cites just such an intervention constraint, as in the following examples:

- (8) a. Typical example:
go + run to the store
- b. Intervening dependent within first participant:
*go out + run to the store

- c. Intervening dependent within second participant:
*go + quickly run to the store
- d. Intervening loose-construction modifier:
*go, as they suggested, run to the store

5.3. Co-satisfaction and interface links

How do the requirements imposed by superderivational rules and those imposed by syntactic rules interact with one another? The natural assumption, and the one I have been making throughout my discussion, is that of co-satisfaction: an expression must satisfy both (5A) all the conditions in relevant syntactic rules and also (5D) all the conditions on the moremes (including supermoremes) that are instantiated in it.

Consider now how syntactic and morphological conditions interact to yield the intervention constraints on syntactic compounds like *television table*. The syntactic rule describing such a compound stipulates that its constituents are both of rank W; consequently, we expect an intervention constraint of type I. But such compounds also instantiate compound moremes, so that we expect an intervention constraint of type III as well. Since a type III constraint incorporates the effect of a much weaker type I constraint, these expectations are satisfied.

But something is missing in this story. Nothing I have said so far would prevent us from viewing *television table* as a syntactic compound of two Ws, each instantiating one ordinary moreme, but not as itself instantiating a compound moreme. After all, there are phrase-words (as in section 3.3), which are composite for syntactic purposes but not for morphological purposes. Expressions like *television table* could be licensed by the syntax alone, and the fact that there are also N+N compound moremes in the language would not necessarily be relevant, since nothing would require that *television table* and similar expressions be instances of such compound moremes.

What is needed is a link between the syntactic rule(s) licensing certain syntactic compounds and the morphological rule(s) describing compound moremes. Such an 'interface link' between morphology and syntax enforces a stronger sort of co-satisfaction than conditions (5A) and (5D) of section 4.1. It is not enough for an N + N sequence to satisfy the conditions required by some syntactic rule and for it to be an instance of a moreme satisfying the conditions required by some morphological rule; there are some language-particular stipulations requiring specific conjunctions of rule satisfaction.

Such links are standardly – perhaps invariably – required for supermoremes. My analysis of the French portmanteaus (Zwicky 1987) depends crucially on such links, and it should be obvious that English serial verbs also involve such a link: these composite moremes must be instantiated via a particular head-V + VP-complement construction, since other such constructions fail to exhibit the shape constraint that characterizes these serial verbs for so many American speakers, allowing the VP *go*

look at the sunset to occur in its base form and present-tense non-third-singular forms, but not any others.

6. CONCLUSION

Let me stress again the pretheoretical character of the enterprise; I am not proposing schemes of representation or formalisms for rules. Rather, I have pointed out that we must distinguish several different senses of *word* and *phrase* and several different patterns according to which a 'word' might happen to be a 'phrase'. And I have observed that the way we make these distinctions has considerable consequences for the architecture of an adequate theory of grammar.

NOTES

* My thanks to those who provided critical commentary on earlier versions of this paper – most notably to Wolfgang Klein, the official discussant of the paper as presented in Nijmegen on 20 December 1989 at the Max Planck Institute for Psycholinguistics workshop, 'When words happen to be phrases', but also to Melissa Bowerman, Greg Carlson, David Dowty, Larry Hyman, Allard Jongman, Aditi Lahiri, Pieter Seuren, and two anonymous reviewers for the *Yearbook of Morphology*. This is a condensation (of 16 June 1990) of the draft distributed in early 1990.

1. I am not denying that there are at least two further types of word-like and phrase-like units in grammar, namely those to which semantic interpretations can be given and those serving as the domains for automatic, or prosodic, phonology. These units are closely aligned with those of morphosyntax, but (as is well known) the alignment is not perfect.
2. Intervention constraints of a fourth type can arise from conditions on prosodic phrase formation.
3. I say 'have the potential to be instantiated', rather than 'are instantiated', so as to leave open the possibility that some expression type will fail to be so instantiated for entirely phonological (or for that matter, semantic or pragmatic) reasons.
4. I would have preferred the term *sentences*, but I realize that it would only engender confusion.
5. In earlier work I used the term *lexemes*. But apparently this is so frequently used for material characterized by idiosyncrasy of semantics and/or phonology that audiences are confused.
6. I am not proposing a theory of rules here, but rather merely attempting to set out the sorts of generalizations that any adequate theory must somehow manage to state.
7. Whether the NP *Pat* here is correctly characterized as a direct object or an indirect object is not the point at issue.
8. The careful summary and critique of the literature on V-Prt combinations (in English, Dutch, and Afrikaans) by Le Roux (1988) discusses analyses that treat some such combinations as (compound) 'words' and interprets these as claims that the combinations are morphological objects, moremes in my terminology. At least some of the problematic aspects of the 'lexicalist' analyses Le Roux considers disappear when such analyses are interpreted instead as claims that the combinations at issue are Ws.
9. I am reluctant to devise a representational scheme specifically for such moreme-free templates, but it would in fact clarify discussions to have such a scheme, rather than using an orthographic representation to stand for both a template and its corresponding expressions, as I have been doing.
10. I would, of course, have preferred the term *lexicon*, but this word seems unavoidably, and inappropriately, to suggest either a list of idiosyncratic and unpredictable expression types, or some sort of psychological construct, or both.

REFERENCES

- Bierwisch, M. 1990. "Verb-cluster formation as a morphological process". In this volume.
- Bloomfield, L. 1933. *Language*. New York: Henry Holt.
- Booij, G. 1990. "The boundary between morphology and syntax: separable complex verbs in Dutch". In this volume.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Di Sciullo, A.M. and E. Williams. 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.
- Dressler, W.U. 1988. "Preferences vs. strict universals in morphology: word-based rules". In M. Hammond and M. Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press, 143-154.
- Hockett, C.F. 1947. "Problems of morphemic analysis". *Language* 23, 321-343.
- Hoeksema, J. 1988. "Head-types in morpho-syntax". *Yearbook of Morphology* 1, 123-37.
- Le Roux, C. 1988. *On the Interface of Morphology and Syntax: Evidence from Verb-particle Combinations in Afrikaans*. *Stellenbosch Papers in Linguistics* 18.
- Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Matthews, P.H. 1974. *Morphology: An Introduction to the Theory of Word-structure*. Cambridge: Cambridge University Press.
- Ojeda, A.E. 1988. "A linear precedence account of cross-serial dependencies". *Linguistics and Philosophy* 11, 457-492.
- Perlmutter, D.M. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.
- Sadock, J.M. 1980. "Noun incorporation in Greenlandic". *Language* 56, 300-319.
- Sadock, J.M. 1985. "Autolexical syntax: a theory of noun incorporation and similar phenomena". *Natural Language and Linguistic Theory* 3, 379-439.
- Zwicky, A.M. 1987. "French prepositions: no peeking". *Phonology Yearbook* 4, 211-217.
- Zwicky, A.M. To appear. "Some choices in the theory of morphology". In S. Davis and R. Levine (eds.), *Formal Grammar: Theory and Implementation*. Vancouver: University of British Columbia Press.
- Zwicky, A.M. and J.A. Nevis. 1986. "Immediate precedence in GPSG". *Ohio State University Working Papers in Linguistics* 32, 133-138.