by **D. Terence Langendoen**

Linguistics Program, CUNY Graduate Center, 33 West 42 Street, New York, N.Y. 10036 and Department of English, Brooklyn College, Brooklyn, N.Y. 11210

Linguistics must be computational too. A generative grammar of a language is a device that computes the sentences of a language. However, the computation is only partial: sentences are computed from words in the syntactic component (Chomsky, 1965), and words are computed from morphemes in the morphological component (Aronoff, 1976; Siegel, 1979), but morphemes themselves are not computed. Rather, they are simply listed in a dictionary.

A morpheme is a rather complicated object, made up of an underlying phonological representation, a semantic representation, and a "morpholexical" representation that expresses its distribution in words. Thus the question arises as to why morphemes, too, are not computed like words and sentences. The answer is that there is only a finite number of morphemes, but from this it only follows that morphemes *can* be listed, not that they *should* be. Suppose that the morphemes of a language are computed by a component that we dub the morphemic component. What would such a component be like?

Ever since Saussure (1915) formulated the doctrine of the arbitrariness of the linguistic sign, it has been assumed that the pairing of phonological and semantic structures in morphemes is arbitrary. This can be expressed by the claim that the rules of the morphemic component are of the form (1), where *A* is a morphemic category, *f* is a phonological structure, and *s* is a semantic structure (we ignore morpholexical structures here and throughout this discussion).

(1) $A \rightarrow [_A\, f, s]$

But as Saussure also pointed out, the arbitrariness of the linguistic sign is not absolute. Even if we put sound symbolism aside, we find many cases in which whole sets of morphemes are semantically and phonologically related. For example, English has many pairs of morphemes, such as *break* (verb) and *break* (noun), that are phonologically identical, and whose semantic difference is predictable by rule. It is often claimed that the verb stem of such pairs is basic and that the noun stem is derived by a phonologically zero affix, but while we may consider the verb stem to be basic, there is no evidence that the noun stem is morphologically complex. Thus, rather than describing the relation between such pairs in the morphological component, we do so in the morphemic component of English as follows. Let *E* (for "empty"), *N* (for "noun stem"), and *V* (for "verb stem") be morphemic categories; let *N* and *V* be axioms of the morphemic component; Let *brāk* be the phonological structure of *break*; and let s_{break} and s_{break}' be the semantic structure of *break* (verb) and *break* (noun) respectively. The rules of the morphemic component include those given in (2).

(2) a. $N \longrightarrow [_N\, V\, E]$
b. $V \longrightarrow [_V\, br\bar{a}k,\ s_{break}]$ / ___ (E)
c. $[_N[_V\, br\bar{a}k,\ s_{break}]\, E] \longrightarrow [_N\, br\bar{a}k,\ s_{break}']$

Now consider the relation between the morphemes *break* (verb) and *broke*, past tense of *break* (verb). The latter is an inflected form of the former and is therefore derived from it; nevertheless, *broke* is not morphemically complex (many linguists have wasted a lot of time trying to figure out ways to analyze forms like *broke* into two morphemes). Accordingly, the relation between these forms is also to be described in the morphemic component. To do so, let $\overline{V}$ (for "inflected verb stem") and *P* (for "past tense") be morphemic categories; let *brōk* be the phonological structure of *broke*; and let s_{broke} be the semantic structure of *broke*, past tense of *break* (verb). The rules of the morphemic component of English also includes those in (3), in which (3b) replaces (2b).

(3) a. $\overline{V} \longrightarrow [_{\overline{V}}\, V\, P]$
b. $V \longrightarrow [_V\, br\bar{a}k,\ s_{break}]$ / ___ $(\{E, P\})$
c. $[_{\overline{V}}[_V\, br\bar{a}k,\ s_{break}]\, P] \longrightarrow [_V\, br\bar{o}k,\ s_{broke}]$

The morphemic analyses just given obviously only begin to provide a glimpse of the kind of work that remains to be done before we can say that a complete account of how morphemes are computed in English has been given. Nevertheless, this glimpse is sufficient to show that even when such an account is provided, a generative grammar is still not completely computational. We still have to describe how phonological and semantic structures are computed.

Only a finite number of phonological and semantic structures enter into the construction of morphemes. By means of the recursive devices in the morphological and syntactic components, however, an infinite number of phonological and semantic structures enter into the construction of words, phrases, and sentences. The set of phonological structures that enter into the construction of the morphemes, words, phrases, and sentences of a language nevertheless does not exhaust the set of phonological structures of a language; very many such structures for example, English *brāg,* still remain unpaired with semantic structures at any level. Such "meaningless" phonological structures of a language may be called " accidental gaps" (Halle, 1962); these contrast with phonological structures called "systematic gaps," which are not part of the language at all (for example, *bnāk* is not part of English). To compute the phonological structures of a language, we propose a phonological component, whose terminal vocabulary consists of the phonemes of that language, together with certain other elements. The rules of this component compute the phonological structures from phonemes much in the way that the syntactic component computes the sentences from words. This component is, in other words, a genuine "phonological grammar" (Householder, 1959) that is quite unlike the phonological component of a standard generative grammar (Chomsky & Halle, 1968). It computes phonological structures completely without regard for their roles in the meaningful constructs of the language. Rather, the phonological structures of those constructs are constrained to be members of the set computed by the phonological component.

Likewise, to compute the semantic structures of a language, we propose a semantic component whose terminal vocabulary consists of the semantic primitives ("sememes") of the language, together with certain other elements. The rules of the semantic component compute the semantic structures of the meaningful constructs of the language from sememes much in the way that the phonological component computes the phonological structures from phonemes. Again, there is nothing comparable to a semantic component of this sort in a standard generative grammar; the "semantic component" of such a grammar (Katz & Postal, 1964; Katz, 1972) must be distributed among the morphemic, morphological, and syntactic components proposed here, since its function is exclusively to show how semantic structures combine in the formation of derived readings for morphemes, words, phrases, and sentences. It says nothing about how such structures are computed in the first place.

The question arises whether there are any accidental semantic gaps in a language, parallel to the accidental phonological gaps discussed above. Since we have no direct intuitions about them, it is certainly possible that they do not exist (Katz, 1979). On the other hand, perhaps they do exist, but we simply do not have direct access to them.

Many linguists have been claiming of late that generative grammars should be "psychologically real" and modifications in the theory of grammar have been proposed in an effort to achieve this (Bresnan, 1978). However we do not have any clear idea of what psychological reality for grammars really is. This is why I think A & C's paper is important. It focuses on the need for psycholinguistic models to be both neurologically based and fully computational. Quite independently, linguistic models also need to be fully computational. It remains to be seen whether the optimal psycholinguistic computational model resembles the optimal linguistic computational model in any way.