IN THE BEGINNING WAS THE WORD

Laurie Bauer
(Victoria University of Wellington)

Introduction

There have been a lot of strange goings-on recently in the
study of word-formation.   After having been almost exclusively
a European preserve for over a century, word-formation has been
"discovered" by the Americans, many of whom appear to be ignor-
ant of the vast amount of European work on the subject.   To
see that this is true, it is only necessary to compare the
bibliographies in, say, Halle (1973) and Jackendoff (1975),
which mention a minimum of European research, with those in
Brekle and Kastovsky (1977), which show just how much recent
work there is on word-formation on both sides of the Atlantic.
Despite this oversight, some interesting work on the subject
is coming from linguists in America, and one of the most
interesting items is Aronoff's (1976) monograph.   One of
Aronoff's main claims in this work is that word-formation pro-
cesses are word-based.   As he puts it (1976: 21):

> "All regular word-formation processes are
> word-based.   A new word is formed by apply-
> ing a regular rule to a single already
> existing word.   Both the new word and the
> existing one are members of major lexical
> categories [defined elsewhere as noun, verb,
> adjective and adverb]."

First let me say something about the status of this
claim.   Aronoff clearly marks it as a "hypothesis"; he also
admits that it does not work all of the time: "new words are
by and large formed from old ones" (1976: 23, my stress),
though the exceptions may be what Aronoff (1976: 20) terms
"oddities" -- acronyms like *NATO* and blends (or 'portmanteau
words') like *smog* and *chunnel*.   It becomes clear, however,
that this claim is a very basic one as far as Aronoff is
concerned, and one which he would be reluctant to give up;
since it provides a motivated limitation on the structure of

W[ord]F[ormation]R[ule]s. It thus seems to me appropriate to consider the word-based morphology hypothesis critically as a basic tenet of an existing theory of word-formation.

It must also be stressed that Aronoff, very sensibly, restricts his theory to the productive formation of words, and does not expect it necessarily to hold for all complex words in the dictionary. For example, there is a word *cavalcade* in the dictionary, but it was not formed "by applying a regular rule to a single already existing word" in the way that the later formations *aerocade*,[1] *Beatlecade*[1] and *motorcade*[1] were: *cavalcade* was a loan. Theories of word-formation which attempt to explain all the words in the dictionary run into innumerable problems of this type; for theories like Aronoff's which only attempt to explain productive formations, words like *cavalcade* are no problem at all.

Aronoff (1976: 28ff) discusses one type of counter-evidence to the word-based morphology claim, the

"case in which there are several words formed from the same stem, but in which the stem never shows up as a word itself."

He gives examples like the following (the asterisks are Aronoff's, see later):

incision	incisive	incisor	*incise
locomotion	locomotive	locomotor	*locomote
malediction	maledictory		*maledict
illusion	illusory		*illude

but is able to argue for these examples (I think convincingly) that it is the noun which is basic and from which the other words are derived. He has three arguments for this position: (i) the history - diachronically speaking the nouns were loans and earlier than the other forms; (ii) the distribution - there are more nouns without corresponding adjectives than vice versa; (iii) the meaning - in those cases where a verb does exist, the meaning of the adjective follows the meaning of the noun rather than the meaning of the verb. It seems to me that there are other types of counter-evidence which can be brought against the claim, notably instances where the base (or "operand" to use Matthews's terminology) is either (a) larger than a word or (b) smaller than a word.

Base larger than the word

The main place where a base larger than the word is used is in
the first elements of compounds, as in, for example, *oh-what-a
wicked-world-this-is-and-how-I-wish-I-could-do-something-to-
make-it-better-and-nobler expression,*[2] *carry-on luggage,*[3] *a
what-do-you-think movement,*[3] *last night's catch-as-catch-can
sleeping arrangements,*[3] *a pain-in-stomach gesture.*[3]   Such
examples are not counter-examples to the word-based morphology
claim as it is formulated by Aronoff, since compounds clearly
fall outside the area that the claim is meant to cover ("a
single already existing word"), involving, as they do, more than
one base.   There are also plenty of examples to be found, how-
ever, where an affix is added to a base of more than one lexeme,
and these, it seems to me, do provide counter-evidence to the
claim.   Some examples taken at random are *cold mooner,*[4] *come-
uppance, cut-down-ness,*[5] *get-at-able, oneupmanship, seeoffable,*[6]
stick-to-it-iveness,[7] *up-to-date-ness,*[4] *with-itness.*[4]   These
examples might look as though they are fairly marginal and
rather light-hearted, but this is not always true.   I did a
survey recently (Bauer, 1979b) based on the word-list of
Barnhart *et al.* (1973) of words ending in the suffix *-er* which
denoted persons.   This showed that *-er* was added to a phrase
made up of more than one lexeme in 15% of cases.   Admittedly,
the sample was a fairly small one, no more than 100 items, but
it seems to me that this proportion is too large simply to be
ignored.   While it is clear that, in the examples listed, a
phrase has been rank-shifted to act as a lexeme in the majority
of cases, it is not obvious that this helps.   There are two
reasons for this.   The first is that rank-shift is not a notion
which applies within the model embraced by Aronoff; this reason
is relatively trivial.   The second reason is that it is not
always a phrase which is used as a base: consider *a pain-in-
stomach gesture* and *get-at-able* where (parts of) two phrases are
involved.   Thus it is not a higher unit being rank-shifted.
Incidentally, it is a problem for any model to decide precisely
what constraints do apply on the formation of such bases, since
it is not the case than any string of lexemes can be used: we
can not have an **any-string-can model* or an **is-not-the-case
claim* for instance.   I have no thoughts to offer on this topic.

Base smaller than the word

The major problem with bases smaller than the word arises in
connection with neo-classical compounds: English words coined
from Greek or Latin elements, such as *telephone* (the original
formulation of this problem appeared as Bauer, 1979a).   In

words   such as *electro-filter*,[4] *geo-hygiene*,[4] *sociolinguistics*,
the elements *electro-*, *geo-* and *socio-* are fairly clearly pre-
fixes: they are of fixed form, fixed meaning and are attached
before bases (operands) which are lexemes.   Similarly, in words
such as *negrophile*, *infanticide* and *Pekinology*,[4] *-phile, -cide*
and *-ology* are suffixes.   But what, then, is the status of
words like *electrophile*,[4] *geocide*[4] and *sociology*?   Are they
made up of a prefix and a suffix?   Prefixes and suffixes are
usually defined in terms of being attached before or after a
base.   Yet if one of the two elements is a base, how can it be
decided in a principled manner which it is?   Is *electrophile*
prefix *electro-* plus base *-phile* or base *electro-* plus suffix
-phile?   Unless some method of deciding between these last two
analyses is found, and unless the point of view that such words
are made up of a prefix and a suffix can be clearly rejected, a
theory of word-formation which accepts the word-based claim
cannot account for words of this type.

A second problem with bases smaller than words is where a
clipped form of the base is used in the word-formation process,
as in, for example, *commitology*[8] or *evacuee* from *committee* and
evacuate respectively.   Here there is obviously no doubt that
the word-formation process has applied to "a single already
existing word"; what is in doubt is whether the rule is in fact
regular.   In the case of *evacuee* from *evacuate* it is, since
-ate is regularly deleted before *-ee* suffixation, cf. *designee*,[8]
nominee, vaccinee.[4]   With *commitology*, however, the case is not
so clear-cut, and there are other examples of this type, though
(perhaps fortunately) they are not particularly common:
centrism,[4] *erotology*,[4] *graviton*,[4] *jockette*,[4] *obituarese*,[4]
scientology.   Words like this provide difficulties not only for
a theory of word-based morphology, but for any generative theory
of word-formation.

Backformation

Aronoff evidently sees no problem for his claim in backformation.
He says (1976:   28) that

> "when we back-form [sic!] ... we essentially ask
> ourselves 'what word *might* this one have been formed
> from?'"

And with examples like *aggress* from *aggression*, which is a loan
word, there is no obstacle to such a process within the theory
outlined by Aronoff.   Consider, however, the recent backforma-
tion *flappable*[4] from *unflappable*.   The problem for the word-

based hypothesis is not the backformation *per se*, but the operand for the backformation process: how was *unflappable* formed? It is a recent coinage (listed by Barnhart *et al*, 1973), but it cannot have been formed by a regular rule from a single already existing word, because that word would have had to be *flappable*, and the whole point of the backformation is that *flappable* is a diachronically later formation than *unflappable*.

Words from foreign bases

There is also a fifth argument against the word-based morphology claim, but I present this one far more tentatively than the others, because I am not really sure of its status. It concerns data like that cited from Aronoff (1976: 29) previously. Aronoff points out that there are sets like *valediction* and *valedictory*, although there is no verb **valedict* and goes on to show that in these cases it is the *-ion* noun which is basic, and the adjective is derived from the noun. His claim that the noun is basic in such paradigms is not invalidated by the fact that many of the verbs he marks with an asterisk are listed in a dictionary like *Chambers's Twentieth Century Dictionary* (verbs like *illude, locomote, maledict)* since such verbs are back-formations from the appropriate nouns. Aronoff does not, however, discuss where the nouns come from. In the instances he considers, this is not vital, since they all appear to be loans from Latin. There are, however, English words coined on Latin bases without being direct loans. As an example, consider the relatively recent coinage *ludic*.[4] This contains a Latin base *lud-* and an English suffix *-ic*. Yet there does not appear to be an English word which might have been a model for *ludic* in the way that *valediction* could have been a model for *valedictory*. The possibilities, as far as I can see, are *Ludo, ludicrous* (which no longer has any obvious semantic association with playing as far as I'm concerned) and the family of *allude, delude* and *elude* (which again have no association with playing). To the extent that this kind of process is productive, it seems to me that it provides a problem for the word-based morphology hypothesis, since the "existing word" on which processes of English word-formation have operated are not English words, but Latin (or sometimes Greek). Obviously the people who coin such words are familiar with the classical languages -- at least to a certain extent -- but I would be very doubtful about the desirability of including the lexica of Greek and Latin (and possibly also French) in the word-formation component of the grammar of English. Other recent words which seem to come into the same class as *ludic* are *ebulism*,[4] *phillumenist*,[4] *vinefy*[4] and *viridian*.[4]

77

Conclusion

I have presented five arguments against the hypothesis of word-based morphology, though, of course, not all of these arguments are of equal weight. First I showed that word-formation could apply to units larger than the word; then I showed that there were at least two cases where it could apply to units smaller than the word; then I showed that the operand in backformation could provide problems; and finally I considered the problem raised by foreign bases for such a hypothesis. Thus, however much this hypothesis may have support in older writings, and however tempting it may be as a hypothesis about word-formation, it cannot be held as a general limitation on rules of word-formation. Unfortunately, word-formation is more complex than this.

Moreover, let me very briefly return to the original formulation of the hypothesis by Aronoff. He says

> "both the new word and the existing word are
> members of major lexical categories."

Again, however nice it would be if this were the case, it is simply not true, as the large numbers of words formed from minor lexical categories show: words like *downer*,[4] *enoughnik*,[9] *iffy*, *ifnik*,[9] *inness*,[10] *muchness*,[10] *phutnik*,[9] *suchness*,[10] *thereness*,[10] *thusly*,[11] *uppity*, *whatnik*,[9] *whyness*.[10] Word-formation is an area where generalizations tend not to last long!

I have spent some time pointing out flaws in a hypothesis which, though obviously not tenable overall, is nevertheless valuable since, it cannot be denied, it does work in the vast majority of cases. It might, I think, fairly be asked what I would wish to put in the place of the word-based hypothesis. At the moment I have no solid answer, but I think there are some lines of research that can be followed. Firstly, we have to stop viewing the word-based hypothesis as an absolute universal, and start thinking of it as a statistical universal (although it is reading a lot into Aronoff's work to take it as a universal at all). It may be the case (I really haven't any idea whether this would hold) that every language that uses derivation has forms which fit with the word-based hypothesis, but that only a few extraordinary languages (like English) have forms which provide problems for such a claim. It may be, again, that some of the types of problem are more widespread than others. It might also be the case that the types of problem that arise recur across languages, and are less "exceptional" than it appears when only English is considered.

Secondly, it seems to me that the major problem I have raised, simply in numerical terms, if in no other way, is the

problem posed by the neo-classical compounds. These surely deserve more study. In particular, in this context, it needs to be decided what properties the elements of neo-classical compounds have in common with lexemes, if any: for example, *-crat*, unlike *-ation*, seems to contain lexical meaning as well as information on the new part of speech, but then so does *-er*; *-crat* is probably a member of an open class, like a lexeme, but so is *-er*, and so on. It can also be asked whether neo-classical compounds can be termed "compounds" in English for reasons other than the merging of the two classical roots: if it is held that the semantic relationships between the elements of compounds can be reduced to a small but exhaustive list (which I personally doubt, but see e.g. Levi, 1978), it may be that the relationships between the elements of neo-classical compounds fit into the same classes. All this may lead us to re-evaluate some cases of prefixation and suffixation, since if *sociology* is a compound, *sociolinguistics* may also be one; it may effectively eliminate the barrier between derivation and compounding, since the dividing line between elements like *socio-* and elements like *un-* may turn out to be one of degree rather than one of type; but until such time as we have a clearer picture of the processes involved in these formations, we have to be very careful with the word-based morphology hypothesis. This hypothesis can only be saved from the problems provided by neo-classical compounds if it can be shown that elements like *geo-*, *socio-* and *-ology* are lexemes, or if they form a class with lexemes which provides the real bases for word-formation processes.

BIBLIOGRAPHY

Adams, V. 1973. *An Introduction to Modern English Word-Formation*. London, Longman.

Aronoff, M. 1976. *Word Formation in Generative Grammar*. Linguistic Inquiry Monograph 1. Cambridge, Mass., M.I.T. Press.

Barnhart, C.L., Steinmetz, S., Barnhart, R.K. 1973. *A Dictionary of New English*. London, Longman

Bauer, L. 1979a. 'Against word-based morphology', *Linguistic Inquiry* 10, 508-509.

Bauer, L. 1979b. 'Patterns of productivity in new formations denoting persons using the suffix *-er* in Modern English', *Cahiers de Lexicologie* 35:2, 26-31.

Brekle, H.E. and D. Kastovsky (eds) 1977. *Perspektiven der Wortbildungsforschung*. Bonn, Bouvier.

Halle, M. 1973.   'Prolegomena to a theory of word-formation', *Linguistic Inquiry* 4, 3-16.

Jackendoff, R. 1975.   'Morphological and semantic regularities in the Lexicon', *Language* 51, 639-671.

Levi, J. 1978.   *The Syntax and Semantics of Complex Nominals.* New York, etc., Academic Press.

Rudnyckyj, J.B. 1959.   'Sputnik and -nik derivatives in the present language of North America', *Etudes Slaves et Est-Européennes* 4 (1959-1960), 142-150.

Williams, T. 1965. 'On the -ness peril', *American Speech* 40, 279-286.

NOTES

* This paper was presented at the Third New Zealand Linguistics Conference, held in Auckland, 3-5 September, 1980. I should like to thank my wife, Winifred Bauer, for her comments on draft versions of this paper.   She should not be blamed for the content.

[1] Adams, 1973: 13.6.1

[2] J.K. Jerome, *Three Men in a Boat*, 1889; Penguin edn, 1957: 21.

[3] P. Mann, *Steal Big*, London, etc., Granada Publishing, 1978: 28, 72-3, 123, 134.

[4] Barnhart *et al.*, 1973.

[5] C. Harrison, *How to be a Pom*, Palmerston North, Dunmore Press, 1975: 57.

[6] Lord Mancroft, *A Chinaman in my Bath*, London, Bachman and Turner, 1974: 98.

[7] P. Gallico, *Mrs Harris Goes to New York*, London, Michael Joseph, 1960: 8. Also B. Garfield, *Wild Times*, London, Macmillan, 1979: 96.   Listed in Websters *Third International Dictionary*.

[8] *American Speech* 46 (1971), 292, 294.

[9] Rudnyckyj, 1959.

[10] Williams, 1965.

[11] W. Goldman, *Magic*, London, Macmillan, 1976: 4.