

PRODUCTIVITY IN WORD FORMATION

Anne Cutler
University of Sussex

Productivity of a particular word formation affix can be assessed in two ways: by its relative frequency of occurrence in the language, and by the frequency with which speakers use it to create nonce-formations. Affixes which attach to a word to form from it another word of a different form class are, as defined by Chomsky and Halle (1968), of two kinds: those which attach with a word boundary (#) and those which attach with a formative boundary (+). The choice of affix boundary is often determined, either because there is only one affix available, or because only one affix is acceptable. For example, a nonce-adverb based on, say, trendy can only be created with #ly; while a noun based on a polysyllabic verb which ends with -ize preceded by an unstressed syllable - say, compartmentalize - must take +ation, never another deverbal nominal suffix such as #al or #ment. In certain instances, however, more than one affix could potentially be used, and it is with such cases, and the factors which determine the choice between alternative possibilities, that the present paper is concerned.

When the choice between acceptable affixes involves both word boundary and formative boundary candidates, it is the word boundary affixes which usually prove to be more productive, in both senses - the language contains more derived words formed with word boundary affixes than with formative boundary affixes (see Aronoff [1976] for a particularly good discussion of this), and speakers prefer to create nonce-words with word boundary affixes rather than with formative boundary affixes. This is shown, for example, by the first of the four investigations which I shall report here, all of them involving relative acceptability of neologisms. In this first study subjects¹ were asked to decide between two alternative forms of each of eight deadjectival nouns, eight deadjectival verbs and eight denominal adjectives. In each case one alternative was formed with a word boundary suffix and the other with a formative boundary suffix: for example, incestuous#ness versus incestuous + ity, ambigu#ize versus ambigu+ify, dowager#ish versus dowager+ial. The subjects were presented with the neologisms embedded in context sentences (for example: "It upset her that the relationship was incestuous. (a) The incestuousity of the relationship upset her. (b) The incestuousness of the relationship upset her."); they selected the version which they felt to be the more felicitous (or less awkward) one. Table I shows the results:

Table I	
Number of choices	
# affixes	+ affixes
168	120

This difference is statistically significant (binomial test, $p < .003$). In other words, subjects in this study clearly preferred neologisms formed with word boundary affixes. However, there is a crucial difference between word and formative boundary affixes, which lies in the phonological relation between the derived word and the base word from which it is formed: whereas words formed with word boundary affixes are transparent, i.e. the phonology of the base word is preserved, words formed with formative boundary affixes can be opaque: for example, stress can fall on a syllable other than the syllable stressed in the base word, a vowel which is tense in the base word can become lax in the derived word, etc. In other words, in choosing neologisms formed with word boundary affixes, subjects in this study were expressing a preference for derived words which were closer to their base words over those which were phonologically further away. In this paper I shall argue that this, and this alone, is the reason why speakers prefer neologisms formed with word boundary affixes.

If it were always, without exception, the case that derived words created with formative boundary affixes were opaque with respect to their base words, then there would be no way of testing whether subjects' preference for word boundary affixes reflected nothing more than a preference for transparency in derived words. However, it is indeed possible to choose base words which will remain transparent with either word boundary or formative boundary affixes. Thus the second study to be reported was a replication of the first in procedure but not in materials; subjects were asked to choose between alternative versions of deadjectival nouns and denominal adjectives formed from base words with final stress (e.g. jejune#ness, jejun+ity; auctioneer#ish, auctioneer+ial), and deadjectival verbs which bore stress on the same syllable irrespective of affix (splendid#ise, splend+ify). If subjects' preferences for word boundary affixes in the preceding study reflected simply their preference for derived words which were phonologically closer to their base, we should expect that in this replication they should show no preferences either way, since the alternative forms are equivalent in phonological transparency; only if something more is involved should we expect the preference for word boundary affixes to be repeated. In fact, as Table II shows, subjects in the second study showed no significant preference for word boundary versus formative boundary derivations:

<u>Table II</u>	
Number of choices	
# affixes	+ affixes
145	143

From this we may conclude that the difference found in the preceding investigation was due to nothing more than the difference in visibility of the base word in the derived form.

In a further study, another group of subjects was given base words and asked to make up their own neologisms - ten deadjectival nouns, ten deadjectival verbs, and ten denominal adjectives. They

were given context sentences, but no ready-made alternatives to choose between (for example: "The tropical heat made them languid. The tropical heat _____ them."). Half the base words would form opaque derivatives with formative boundary affixes (sinister, amorous, villager), whereas the other half could form transparent derivatives with either word or formative boundary affixes (germane, languid, engineer). On the basis of the preceding results, we would predict that subjects would tend to prefer word boundary affixes with the former group and show no preference either way with the latter. The results are shown in Table III:

Table III
Frequency of # and + forms

	Nouns		Verbs		Adjectives	
	#	+	#	+	<i>ft</i>	+
Opaque	100	41	117	33	88	62
Transparent	79	71	62	38	84	66

An analysis of variance performed on the frequency of word boundary choice shows that the overall difference between opaque and transparent base words is statistically significant, but that this effect interacts significantly with word class. Separate t-tests on the data for each word class show the difference in affix preference for opaque and transparent words respectively to be significant in the nouns ($t=3.55$, $p<.01$) and the verbs ($t=8.14$, $p<.001$). There is, however, no significant difference between the two groups of words in the adjectives; this is as it happens due to subjects' consistent preferences for the two forms balloon#ish and spectator+ial. (Indeed, if these two words are omitted from the analysis, the adjectives show an effect in the same direction as the nouns and verbs, and significant at the .01 level.) The results for the verbs also conceal the fact that the suffix +ify can take two forms, one opaque, as in the fluid - fluidify paradigm, and the other transparent, as in the liquid - liquefy paradigm. In the present study subjects chose the transparent form twice as often as they chose the opaque form, exactly as the hypothesis of preferred transparency would predict.

Finally one might look at subjects' judgements of what is and is not an English word. Aronoff and Schvaneveldt (1978) have in fact done just this. In a study using a variant of the lexical decision task they showed that nouns formed from adjectives ending with -ive were more likely to be accepted as English words if they were derived with #ness than if they were derived with +ity. We might speculate that this preference for word boundary affixes would also disappear if the base words were equally transparent with either word or formative boundary affixes. Accordingly, in the fourth investigation to be reported here, Aronoff and Schvaneveldt's study was replicated with an appropriate change in the materials. Subjects were presented with a list of 80 nouns and were asked to judge for each one whether or not it was an English word. Of the 80 nouns, 20 were indisputable real words listed in the Concise Oxford Dictionary (e.g. captivity,

possessiveness, arousal, contusion), 20 were indisputable nonwords (e.g. rubicativity, tulsiveness, combasement, infession), and the remaining 40 were possible words - existing adjectives or verbs expanded by the addition of suffixes into nouns (which did not occur in the dictionary).² Each of the possible words (and in fact each of the nonwords) occurred in two forms, one created with a word boundary suffix, the other with a formative boundary suffix. In half of the possible words the base words became opaque in the formative boundary derivative (e.g. reflective#ness, reflectiv+ity; legislative#ness, legislativ+ity), in half of them both the word boundary and formative boundary forms were transparent (e.g. jejune#ness, jejun+ity; excusc#ment, excus+ion). Two lists were constructed, with type of ending counterbalanced across lists for possible words and nonwords. Fifteen subjects saw each list; thus each subject saw only one derivative of each base word.

We predict that for the group of words which are opaque with formative boundary affixes - which constitutes a direct replication of Aronoff and Schvaneveldt's study - the word boundary forms will prove more acceptable than the formative boundary forms, whereas when transparency is equal with both word boundary affixes and formative boundary affixes, subjects will show no preferences for one over the other. Tables IV and V show the results of this study:

Table IV
Proportion of yes responses

	Words	Possible Words (Opaque)	Possible Words (Transparent)	Nonwords
#	.97	.51	.32	.06
+	.99	.40	.30	.09

Table V
Possible words: frequency of affix choice

	Opaque	Transparent
More yesses to # form	14	10
No difference	1	3
More yesses to + form	5	7
Significance level (Sign test)	P<.05	M.S.

Again the hypothesis is supported: speakers prefer word boundary forms when the formative boundary alternatives are opaque, but show no preferences either way when both word boundary and formative boundary forms are transparent. Thus this series of studies provides strong evidence that speakers' preferences among derived words are determined by transparency or opaqueness of the derivative with respect to the base word.

Yet more support for the preferred transparency hypothesis can be found in a quite different kind of psycholinguistic source material, namely in spontaneous slips of the tongue. Two kinds of speech error prove to be relevant. The first kind involves mistaken word formation, in which the wrong affix is employed to form a derived word, e.g. self-indulgement for self-indulgence. These

are not among the most common types of speech error, and from my own collection, the collection assembled at the University of Illinois by David Fay, and published sources I have gathered only 119 clear examples. It is possible to compare the degree of transparency of the derived word formed with the erroneous suffix with the degree of transparency of the intended word, to determine whether there is any tendency for speakers to produce an error which is more transparent than the intended word would have been. In fact it transpires that the majority of such errors do not allow a test of the hypothesis since they involve errors and intended words with an equal degree of transparency. However, I found 21 cases in which the error resulted in a word more transparent with respect to the base word than the intended utterance would have been, e.g. derival for derivation; dispution for disputation; concedence for concession; expect ion for expectation; professoral for professorial; etc. Against this I found only 8 in which the reverse was true, i.e. the error was less transparent than the intended word would have been, e.g. inconsideration for inconsidcrateness. This difference is statistically reliable (binomial test, $p < .02$), indicating that speakers prefer transparent over opaque forms in erroneous word formations also.

The second variety of speech error which produces relevant evidence is the lexical stress error (e.g. prosodic). These characteristically show the intrusion of the stress pattern of a morphologically related word (in the example, prosody). A subset of these pairs (of intended word plus "intruder") involves base word plus morphologically complex derivative of it. Among the 53 of these in my collection, 37 (70%) involve a derived word produced with stress on the syllable which bears it in the base word (e.g. advantageous, or syntactic, in which the intruders are presumably advantage and syntax respectively), while only 16 involve base words stressed on a syllable which bears stress in a derivative (e.g. syntax, which presumably shows an intrusion from syntactic). This difference is also statistically reliable (binomial test, $p < .01$). Thus these errors also attest to speakers' preference for transparent over opaque derived words.

A variety of evidence therefore indicates that derivations with word boundary affixes are preferred to derivations with formative boundary affixes because speakers prefer derived words to be transparent, i.e. to exhibit the base word. The source of this preference is without doubt to be found in performance factors; several possible reasons spring to mind. Firstly, transparent words may be preferred because the speaker desires (consciously or unconsciously) to make the origin of the word (and hence the meaning) clear to the hearer. The possibility of the hearer misunderstanding, or failing to make any sense of the word at all, would be particularly great in the case of nonce-formations. By uttering a neologism which has its base word embedded unchanged within it, the speaker offers the hearer at least the chance of accessing the entry of the base word in his internal lexicon. In connection with this, it is noteworthy that transparency would appear to be defined chiefly in terms of the initial portions of the word and/or the

location of primary stress. The liquid-liquefy paradigm is preferred to the fluid-fluidity paradigm for verbs formed from such adjectives as languid and splendid, although the latter paradigm preserves the final consonant whereas the former does not. Thus we could say that speakers are assuming that the hearers' job of lexical access will be easier if the initial portions and stressed syllable of the base word are intact. This would accord well with a model of the internal lexicon such as that proposed by Fay & Cutler (1977), in which words are accessed by left-to-right phonological structure.

A second possible reason for the preference for word boundary derivations might lie in uncertainty on the part of the speaker, as suggested by Aronoff (1976: p.42). Some formative boundary affixes can attach to the base word in more than one way. Nouns ending in +ity, for example, corresponding to adjectives ending with -al sometimes include the -al (external - externality) and sometimes do not (eternal - eternity). In such cases the word boundary forms may be "easier" in the sense that they relieve the speaker of the necessity of choosing, say, which syllable to shift the stress to.

Thirdly, the preference for word boundary derivations may actually reflect the structure of the internal lexicon itself. It may, for instance, be the case that transparent forms are listed in the lexicon in conjunction with the base words from which they are derived, whereas opaque derivatives are accorded separate lexicalisation; in this case transparent forms would be "easier" to produce in the sense that they would be located more rapidly in a lexical search starting from the base word. In fact, exactly this arrangement of the internal lexicon has been suggested independently by Bradley (1979), on the basis of evidence from response time to decide whether or not a derived word is indeed a word. Such response times are very closely correlated with frequency of occurrence of the stimulus words in the language - the more common the word, the faster it is accepted as a word. But Bradley found that lexical decision response time to words ending with #ment is not well predicted by the frequency of occurrence of the derived word alone, but is better predicted by a combined frequency measure which adds the frequency of occurrence of the derived word to the frequency of occurrence of the base word. On the other hand, for words ending with +ion the reverse is true - the frequency of the derived word is a better predictor of response time than the combined frequency of derived word and base. Bradley argued that this result indicates that the lexical entry for the base word is accessed during processing of derived words with word boundary affixes, but not during processing of derived words with formative boundary affixes.

The fact that word boundary forms are more productive than formative boundary forms in the case of nonce-words seems, in conclusion, to be a consequence of the fact that the former will be transparent with respect to the base word whereas the latter might not. It is at least possible that the fact that word boundary forms are more productive than formative boundary forms in the vocabulary of the language might have exactly the same origin.^{4,5}

Footnotes

1. All four investigations used unpaid volunteer subjects drawn from the University of Sussex community. All were native English speakers (47 British, 3 Americans, 2 Australians, 1 Canadian, 1 Jamaican). Twelve subjects took part in each of the first two studies, and 30 in the third and fourth which were administered together; there was no overlap between the three groups.

2. A few of the possible words were listed in the complete Oxford Dictionary (e.g. reflectiveness, surmusal); but none was in the Concise Oxford Dictionary, which I took to represent the outer bounds of my subjects' vocabulary. Aronoff and Schvaneveldt also varied the instructions they gave to their subjects - to judge whether the nouns were (a) in the subjects' vocabulary; (b) English words; (c) meaningful. They found the largest difference between word and formative boundary forms with (b), which was the form of instruction used in the present study. The real words, possible words and nonwords ending with #ness and +ity in the present study were lifted straight from Aronoff and Schvaneveldt's materials.

5. This is not surprising; as already mentioned, in cases where more than one suffix is possible, the word boundary form is more likely to be found in the language than the formative boundary form. Therefore intended words will more often be transparent than opaque; the error forms are in such cases also usually transparent.

4. This research was supported by a grant from the Science Research Council. The author is very grateful to David Fay who kindly presented the paper at the meeting.

5. The author regrets that the paper by Aronoff and Anshen (1979) was not known to her at the time the present paper was prepared.

References

- Aronoff, M. Word Formation in Generative Grammar. Cambridge, Ma.: MIT Press; 1976.
- Aronoff, M. & Anshen, F. Morphological productivity and phonological transparency. Paper presented to the LSA, December 1979.
- Aronoff, M. & Schvaneveldt, R. Testing morphological productivity. Annals of the New York Academy of Sciences, 318, 1978, 106-114.
- Bradley, O.C. Lexical representation of derivational relation. In M. Aronoff & M.-L. Kean (eds.) Juncture. Cambridge, Ma.: MIT Press; 1979.
- Chomsky, N. & Halle, M. The Sound Pattern of English. New York: Harper & Row; 1968.
- Fay, D.A. & Cutler, A. Malapropisms and the structure of the mental lexicon. Linguistic Inquiry, 8, 1977, 505-520.