TO APPEAR IN *Anglistentag 1999*. Tübingen: Niemeyer, 2000

INGO PLAG (SIEGEN)

On the mechanisms of morphological rivalry: A new look at competing verb-deriving affixes in English

1. Introduction

In English there are a number of morphological processes by which verbs can be derived from items of a different syntactic category (mostly adjectives and nouns). Disregarding the affixes that add a negative, privative, or reversative element of meaning (such as *de-*, *un-*), the following category-changing verb-forming affixes are attested: the prefixes *en-*, *em-* and *be-* (sometimes accompanied by the suffixation of *-en*), the suffixes *-en*, *-ize*, *-ate*, and *-ify*, and a process of conversion.[1]

Although these processes have been the topic of many studies, the distribution of the different affixes has remained largely unclear. Given a base word, which affix can be applied to derive a new verb? As can be seen from the data in (1), this is not a trivial matter. Only certain combinations are possible, which raises the question how the affixes select their bases (or the bases select their affixes). Obviously, there must be certain restrictions at work, whose nature, however, is not quite clear.

(1)	academic-ize	*academic-ify	?academic-$Ø_V$
	computer-ize	*computer-ate	*computer-en
	art-ify	*art-ate	?art-$Ø_V$
	Nazi-fy	*en-nazi	*Nazi-ate
	passiv-ate	passiv-ize	passiv-ify

Apart from this descriptive empirical problem the situation in English pertains to a more general theoretical problem, that of morphological rivalry. Given a set of rival affixes, which general mechanisms govern their distribution? Basically there are two hypotheses concerning this problem, the separation hypothesis and the sign-based hypothesis. Very briefly stated, the separation hypothesis claims that the different affixes are different phonological spell-outs of one underlying lexical derivation. The affixes should therefore show absolute synonymy, and with individual affixes no polysemy effects

[1] In what follows I will refer to the latter process also as a kind of affixation or 'zero-derivation', but only as a matter of terminological convenience and not out of theoretical conviction.

should be observable (see Beard 1995, Don 1993, Gussmann 1987, Szymanek 1985 for details and further elaboration). The sign-based hypothesis, on the other hand, claims that the distribution of verbalising affixes is governed by the individual properties of each affix, which behave like other signs (e.g. lexemes) in this respect. The different affixes should therefore not be completely synonymous.

There is yet another mechanism that has been proposed to influence morphological rivalry and that cuts across the two models just outlined, viz. blocking. The concept of blocking can be reformulated as what I call the blocking hypothesis, which says that the distribution of the different affixes is governed by type-blocking (e.g. *-ity* blocks *-ness*) and token-blocking (e.g. *thief* blocks **stealer*). The central claim is that the (more) special case blocks the more general case (see e.g. van Marle 1985, 1986, Rainer 1988 for elaboration and discussion).

In this paper, I will summarise the analysis of verb-deriving processes as put forward in Plag (1999) and argue on the basis of the analysis of a large number of 20th century neologisms that only the sign-based hypothesis makes correct predictions about the distribution of rival verb-deriving affixes in English. It is shown that the separation hypothesis cannot account for the data and that type-blocking is not a relevant mechanism. Hence, I argue for a sign-based model of morphological rivalry, in which, apart from local analogy and token-blocking, no additional mechanisms are at work. Given the obvious limitations of space, many details cannot be discussed in this article; readers interested in such details are referred to Plag (1999).

I will proceed as follows. First, I will briefly look at the productivity of the verb-deriving processes. Since I am concerned with present-day English, processes that are no longer productive are outside the scope of my investigation. Then, given that the term 'rivalry' implies phonological distinctness but semantic identity, i.e. the synonymy of the affixes in question, I will first present an analysis of the meaning of each affix to see whether this assumption is justified. It will turn out that the affixes are not completely synonymous and that their domains of application only partially overlap. Section 4 will then turn to the phonological properties of derived verbs in order to investigate whether the distribution of affixes in the truly overlapping domains can at least partially be explained as the result of phonological restrictions. Finally, the competing hypotheses introduced above will be evaluated against the attested data. The data for this study were in their vast majority 20th century neologisms extracted from the *OED on CD-ROM* and *hapax legomena* extracted from the Cobuild corpus (c. 20 million words, July 1995). More than 1000 derivatives were systematically investigated. As it turned out, the idea to concentrate on neologisms proved to be a crucial methodological innovation. Earlier studies, like e.g. Schneider (1987), had suffered from the fact that many of the words investigated were lexicalised and had adopted all kinds of idiosyncratic properties, which newly created words, formed on the basis of a productive word-formation rule, lack. It was thus possible to find generalisations that earlier studies necessarily had missed.

2. Productivity

Productivity is generally loosely defined as the possibility to coin new complex words according to the word formation rules of a given language. In Plag (1999) the productivity of *-ate*, *be-*, *en-*, *-en*, *em-*, *-ify*, *-ize*, and conversion is measured using both corpus-based and dictionary-based methods. According to the traditional dictionary-based measure, the number of neologisms with a given affix in a certain period is counted and taken as a direct reflection of the productivity of that affix. Under the more sophisticated statistical measures developed by Harald Baayen and his co-workers, we compute the probability of encountering a newly-formed word of the relevant morphological category when sampling through a large text corpus (see Plag 1999: chapters 2 and 5 for detailed discussion of quantitative measures of productivity). In our case, the results of both dictionary-based and corpus-based methods did not yield any spectacular results but rather confirmed and quantified more precisely what previous authors like Marchand (1969: 148, 164, 364), Bauer (1983: 223), Gussmann (1987: 96) had already stated before. Consider first the figures derived from the *OED* in Table 1:

Table 1: Number of attested 20th century neologisms in the OED

affix	*N* (number of types)	N_{PART} (number of neologisms only attested as participles)	N_{TOTAL} (total)
Ø	488	0	488
-ize	284	62	346
-ate	72	15	87
-ify	23	7	30
eN-	7	0	7
-en	2	0	2
be-	0	0	0

(adapted from Plag 1999: 104)

The *OED* figures shows that conversion, *-ize*, *-ate*, and *-ify* (in that order) are most frequently used to derive new verbs in English. Furthermore, these processes seem to be the only ones that are at least to some extent productive. *-ify* might even be considered a doubtful candidate for a productive suffix, but the corpus-based measure, to which we now turn, shows that it is in fact more productive than *-ate*.

The main methodological problem with measuring the degree of productivity of a given affix is to operationalise the notion of 'possibility' mentioned in the above definition of productivity. Apart from truly unproductive derivational processes like e.g.

nominalising *-th* (as in *length*), productivity seems to be a scalar concept. In other words, with some affixes one is more likely to encounter newly-formed words than with others, a fact that makes productivity a probabilistic notion which is susceptible to statistical analysis. Baayen and co-workers (Baayen 1989, 1992, 1993, Chitashvili and Baayen 1993, Baayen and Lieber 1991, Baayen and Renouf 1996) have developed a number of corpus-based statistical measures of productivity, which all rely on the existence of more or less representative and sufficiently large samples of computerised texts.

There are three principal statistical measures available. The first of these measures is the number of tokens *N* of a given morphological category, which is calculated by counting how often words of a given morphological category are used in the corpus (number of tokens = *N*). The second measure is the number of types *V* of a given morphological category, which is calculated by counting how many different words belonging to the category occur in the text (number of types = *V*). *V* is also referred to as 'extent of use'. The third important measure is the number of words of the given category that occur only once in the corpus (so-called *hapax legomena*, or hapaxes for short), which can be interpreted as an indication of how often a suffix is used to coin a hitherto unattested word, i.e. a neologism. Why should hapaxes, i.e. the new, unobserved types, tell us anything about productivity? After all, the new, unobserved types could only be rare words, and not neologisms. There are, however, strong arguments for claiming that hapaxes are significant for productivity. Thus, it has been shown (e.g. by Baayen and Renouf 1996) that the number of hapaxes of a given morphological category correlates with the number of neologisms of that category, so that the number of hapaxes can be taken as an indicator of productivity. Note that it is not claimed that a *hapax legomenon* ***is*** a neologism. A *hapax legomenon* is defined with respect to a given corpus. When this corpus is small, most *hapax legomena* will be well-known words of the language. However, as the corpus size increases, the proportion of neologisms among the *hapax legomena* increases, and it has been shown that it is precisely among the *hapax legomena* that the greatest number of neologisms appear (Baayen and Renouf 1996). From a statistical viewpoint, the *hapax legomena* play an essential role for gauging the probability that new forms will be encountered that have not been observed before in the corpus.

This approach to measuring morphological productivity receives strong support from the fact that high-frequency words (e.g. *happiness*) are more likely to be stored in the mental lexicon than are low-frequency words (see e.g. Rubenstein and Pollack 1963, Scarborough *et al*. 1977, Whaley 1978). Baayen and Renouf write that

> [i]f a word-formation pattern is unproductive, no rule is available for the perception and production of novel forms. All existing forms will depend on storage in the mental lexicon. Thus, unproductive morphological categories will be characterised by a preponderance of high-frequency types, by low numbers of low-frequency types, and by very few, if any, *hapax legomena*, especially as the size of the corpus increases. Conversely the availability of a productive word-formation rule for a given affix in the mental lexicon guarantees that even the lowest frequency complex words with that affix can be produced and understood. Thus large

> numbers of *hapax legomena* are a sure sign that an affix is productive. (Baayen and Renouf 1996:74)

Having established the significant role of hapaxes in the determination of productivity, we can use the number of hapaxes and the number of tokens to calculate a derived measure of productivity known as 'productivity in the narrow sense', defined as the quotient of the number of *hapax legomena* n_1 with a given affix and the total number of tokens *N* of all words with that affix:

(2) $P = n_1^{aff} / N^{aff}$

Baayen and Lieber (1991: 809-810) explain the idea behind *P* as follows.:

> Broadly speaking, *P* expresses the rate at which new types are to be expected to appear when *N* tokens have been sampled. In other words, *P* estimates the probability of coming across new, unobserved types, given that the size of the sample of relevant observed types equals *N*.

Although there are certain problems involved in the sampling of relevant tokens and types (see Plag 1999: 106-115 for discussion), the productivity *P* of an affix can be calculated and interpreted in a rather straightforward fashion. A large number of hapaxes leads to a high value of *P*, thus indicating a productive morphological process. Conversely, larger numbers of high frequency items lead to a high value of *N*, hence to a decrease of *P*, indicating low productivity. These results seem to be exactly in accordance with our intuitive notion of productivity, since high frequencies are indicative of the less-productive word-formation processes (Anshen and Aronoff 1988, Baayen and Renouf 1996, Plag 1999: chapter 5). Having clarified the reasoning behind corpus-based measures of productivity, I will now return to the problem of derived verbs. In view of the uncontroversial non-productivity of *eN-*, *-en*, and *be-*, and the impossibility to search for zero-derived verbs (a serious drawback of corpus-based productivity measures), the analysis was limited to the only three productive overt affixes, *-ize*, *-ate* and *-ify*. The results are summarised in Table 2 (cf. next page).

Table 2. Types, hapaxes, tokens, and productivity P of derived verbs in the Cobuild corpus

suffix	types (*V*)	hapaxes (n_1)	tokens (*N*)	productivity *P* (*P*=n_1/*N*)
-ate	481	69	41561	0.0017
-ify	88	18	7236	0.0025
-ize	347	80	20865	0.0038

(adapted from Plag 1999:111)

The following picture emerges from table 2. The suffix *-ate* has the highest extent of use *V*, followed closely by *-ize*, whereas *-ify* occurs in much fewer types. However, *-ate* is clearly the least productive in the narrow sense, since in relation to its extremely high number of tokens (*N*=41561) there are very few hapaxes, i.e. new types. The suffix *-ate* can therefore be characterised as a suffix that occurs in many different existing words, but which is not used very often to coin new verbs.

By far the most productive process in the narrow sense is *-ize*, which occurs in almost as many different types as *-ate*, but gives rise to many more new words in relation to the overall number of tokens. Thus the probability of encountering a neologism among all *-ize* derivatives is much greater than with *-ate* verbs. The suffix *-ify* occupies the medial position in terms of *P*, although only few existing verbs contain this suffix, i.e. *V* and n_1 are low.

To summarise, we can say that only conversion, *-ize*, *-ify*, and *-ate* are productive verb-deriving processes in English. This excludes *be-*, *en-*, *-en*, *em-* from further consideration, they are not rivals of the other four affixes. This does of course not exclude the possibility that an existing derivative with one of the unproductive affixes token-blocks a new formation with one of the productive affixes. For example, *blacken* may block zero-derived **to black*. Let us now turn to the semantics of the productive affixes to see whether they are indeed semantically identical, i.e. true rivals.

3. The semantics of derived verbs

In this section we will look at the meaning of the different kinds of derived verbs. It will turn out that *-ize* and *-ify* have the same meaning, which differs from that of *-ate*, which in turn differs from that of conversion. There is however, a certain amount of overlap between *-ize/-ify* and the other two suffixes. This overlap will be discerned as precisely as possible.

It has long been observed that *-ize* derivatives can adopt a number of different meanings that are traditionally labelled as in (3):

(3)

locative	'put (in)to X'	*hospitalise*
ornative	'provide with X'	*patinize*
causative/factitive	'make (more) X'	*randomize*
resultative	'make into X'	*peasantize*
inchoative	'become X'	*aerosolize*
performative	'perform X'	*anthropologize*
similative	'act like X'	*Powellize*

(cf. Jespersen 1942: 319, Marchand 1969: 320)

In Plag (1998) I have shown that these meanings are systematically related to each other and that *-ize* is therefore a clear case of a polysemous suffix. Using Jackendoff's lexical conceptual semantics as the framework of analysis, the following underspecified lexical conceptual structure was proposed (see Plag 1998, 1999 for details):

(4) Lexical Conceptual Structure (LCS) of *-ize* verbs

[[ ]$_{BASE}$ *-ize*]$_V$
{ NP$_i$ ___ NP $_{Theme}$, NP $_{Theme}$ ___ , NP$_i$ ___ }
CAUSE ([ ]$_j$, [GO ([$_{Property, Thing}$] $_{Theme / Base}$; [TO [$_{Property, Thing}$] $_{Base / Theme}$])])

Due to the underspecification of the argument positions in (4), one derivative can adopt different meanings, as illustrated in (5):

(5) LCS of *primitivize*

'*trans. and intr*. To render primitive; to impute primitiveness to; to simplify; to return to an earlier stage' (*OED*)

a. 'ornative' = 'impute primitiveness to'
CAUSE ([ ]$_i$, [GO ([*primitive*] $_{Base}$; [TO [ ] $_{Theme}$])])

b. 'causative' = 'render primitive'
CAUSE ([ ]$_i$, [GO ([ ] $_{Theme}$; [TO [*primitive*] $_{Base}$])])

c. 'inchoative' = 'become primitive', cf. 'return to an earlier stage'.
GO ([ ] $_{Theme}$; [TO [*primitive*] $_{Base}$])

Depending on which argument slot is filled, different interpretations of one and the same derivative emerge. If, for example, the base word *primitive* is interpreted as a Property and as the first argument of the GO function, the ornative meaning results. If, however, *primitive* is assigned to the argument position of the TO function under elimination of the optional CAUSE function, the inchoative meaning is computed (note that the dashed underline in (4) signals optionality, following Jackendoff's conventions (1990: 73).

Turning to *-ify* an almost identical state of affairs can be observed, although the two most marginal meanings 'performative' and 'similative' are not attested among the 23 *OED* neologisms. However, the fact that the similative meaning is attested once in the Cobuild corpus (among the 16 hapaxes) indicates that the rarity of examples of that kind is probably due to the overall small number of *-ify* derivatives in the data.

(6)

locative	'put (in)to X'	*tubify*
ornative	'provide with X'	*youthify*
causative/factitive	'make (more) X'	*aridify*
resultative	'make into X'	*trustify*
inchoative	'become X'	*mucify*
performative	'perform X'	-
similative	'act like X'	*nannify*

It can therefore be assumed that *-ify* and *-ize* can express the same range of meanings, i.e. they are synonymous. Hence I propose the same LCS for *-ify* as for *-ize*.

In contrast to *-ize/-ify* derivatives, *-ate* derivatives are much more constrained semantically. Among the productively-formed derivatives, only ornative and resultative meanings are attested, with the additional restriction that the base word must denote a chemical substance:

(7) ornative-resultative *-ate* (one third of the data)

a. [['chemical substance']$_{\text{Base}}$ *-ate*]$_{V}$

b.

ornative	'provide with X'	*fluorinate*
resultative	'make into X'	*methanate*

(Plag 1999, cf. also Marchand 1969: 258)

A large portion of the data does, however, not conform to this pattern, but shows various kinds of idiosyncrasies, with *-ate* being no more than an indicator of verbhood (see Marchand 1969: 258, Plank 1981: 214 for a similar point). Some examples of such non-canonical formations are given in (8):

(8) *-ate* as indicator of verbhood (unpredictable semantics, two thirds of the data):
- back-formation (*formate* < *formation*),
- local analogy (*stereoregular : stereoregulate :: regular : regulate*),
- conversion (*citrate*),
- idiosyncratic formations (*dissonate, fidate, pathosticate, tambourinate vagulate, Coventrate*)

In sum, we can state that apart from idiosyncratic formations there is a productive ornative-resultative suffix *-ate* that can express a subset of the possible meanings expressed by the other two overt suffixes.

Conversion has been the most popular of all verb-deriving processes as a subject of linguistic inquiry. Accounts of the meaning of the zero-affix are extremely numerous and diverse, but all researchers agree that it is an extremely productive process. Verbs derived by conversion exhibit the widest range of meanings of all verb-deriving processes. In (9) I list a number of meaning categories (with examples) that have been proposed in the pertinent studies (e.g. Kulak 1964, Marchand 1964, 1969, Rose 1973, Karius 1985):

(9)

locative 'put (in)to X'	*jail*	
ornative 'provide with X'	*staff*	
causative	'make (more) X'	*yellow*
resultative	'make into X'	*bundle*
inchoative	'become X'	*cool*
performative	'perform X'	*counterattack*
similative	'act like X'	*chauffeur, pelican*
instrumental	'use X'	*hammer*
privative	'remove X'	*bark*
stative	'be X'	*hostess*

(e.g. Kulak 1964, Marchand 1964, 1969, Rose 1973, Karius 1985)

Sometimes semantic categories have been suggested that cut across the ones proposed above, such as 'movement in time and space' (*jet, winter*), 'typical action of base' (*hammer*), 'typical function of base' (*cripple*), see Karius (1985). The relevance of these labels and their empirical and theoretical justification need not concern us here. What is important, however, is the growing consensus in the linguistic literature that the variety of meanings that can be expressed by zero-affixation is so large that there is no specific meaning that can be attached to the process of conversion.

This position has been argued for in detail in Clark and Clark (1979) and Aronoff (1980) who claim that zero-derivation is a semantically impoverished morphological process.[2] Being a verb, the derived form must denote an Event, State, or Process. Being derived from another word, the verb must denote something that has to do with this base word. The diversity of meanings then "follows directly from the fact that the meaning of the verb is limited only to an activity which has some connection with the noun" (Aronoff 1980: 747). The correct interpretation of the derived verb crucially involves non-linguistic knowledge (in a way we need not discuss here, see Clark and Clark 1979 for some discussion). The semantic diversity and unpredictability of zero-derived verbs can be further illustrated with some examples from my neologism corpus:

(10)

eel	'fish for eel' or 'to move ... like an eel'
premature	'of a shell or other projectile: to explode prematurely'
crew	'act as a (member of a) crew' or 'assign to a crew'
young	'to present the apparently younger side'

(paraphrases from *OED*).

We can conclude that the meaning of conversion is indeterminate and simply verbal in nature.

---

[2] See Lieber and Baayen (1993) for a similar account of zero-derived verbs in Dutch.

To summarise the semantic analysis of the productive verb-deriving processes in English we can state the following facts. Derivatives in *-ize* and *-ify* are synonymous. *-ate* formations are semantically strongly restricted, to the effect that their possible meanings are a subset of those of *-ize/-ify* derivatives. Conversion is semantically indeterminate, converted verbs can express all those meanings that overtly affixed verbs can and many more. It could therefore be considered the most general case in terms of type-blocking. With regard to the question of rivalry, there is only partial semantic overlap between the affixes, so that the domains where there is true competition between them is smaller than previously conceived. We will see in the next section that these domains are further curtailed by phonological restrictions.

4. The phonology of verbal derivatives

In Plag (1999) I have put forward a detailed phonological analysis of derived verbs in English. The results of this analysis can be summarised as follows. The investigation of the prosodic and segmental properties of derivatives in *-ize* and *-ify* has shown that we are dealing with a case of phonologically conditioned suppletion. The two suffixes are (almost) complementarily distributed, with trochaic and dactylic bases taking *-ize* (e.g. *rándom - rándomìze, féderal - féderalìze,* but *tube - *tubize*), and iambic and monosyllabic bases taking *-ify* (*tube - tubify, random - *rándomify*). Stress shift is not possible (cf. **randómify*). The only, but systematic, exception to the complementarity of *-ize/-ify* can be observed with trochaic base words ending in /I/, which take *-ify* under loss of that segment (as in *nazify*), or take *-ize* (with no accompanying segmental changes apart from optional glide insertion, as in *toddyize*).

Similar to *-ize*, the suffix *-ate* only attaches to trochaic and dactylic bases. With conversion no phonological restrictions are observable. However, conversion is not totally unconstrained, since it does not apply to certain kinds of derived adjectives nor to derived nouns, which is no problem for *-ize* (cf. *conventionalize* vs. **to conventional*). We are now in a position to investigate those specific domains where both semantic and phonological constraints allow more than one affix.

5. Where have all the rivals gone?

As has become clear in the previous sections, the domains where more than one affix is possible is much more curtailed than previously conceived. Due to their individual semantic and phonological properties, the domains of application of verbal affixes overlap only partially. Apart from the suppletive suffixes *-ize/-ify*, and contrary to the claims made by the separationists (e.g. Beard 1995), the productive verb-deriving processes of English are not completely synonymous. Furthermore, individual overt suffixes show strong polysemy effects in a way that is typical of lexemes. Coming back to the two initial hypotheses introduced above, both facts are unexplained and unexpected under the separation hypothesis, but are predicted by the sign-based hypothesis. But even if we can reject the separation hypothesis at this stage, we still have to answer the question of which mechanisms govern the distribution of affixes in the truly overlapping domains.

The data allow us to test the following prediction: If no type-blocking occurs and no other mechanisms are operative, the sign-based hypothesis predicts the occurrence of more than one affix where they are semantically and phonologically licensed. If, to the contrary, type-blocking were at work, we would expect that the more specific cases block the more general cases. In other words, *-ate* would block *-ize* in its more restricted domain, and *-ize/-ify* and *-ate* would block the most general case of conversion in their respective domains. As we will see, the latter prediction is not born out by the facts, i.e. type-blocking does not play a role.

First, let us consider *-ize* and *-ify*. As already hinted at above, there is a small overlap in the prosodic structures of possible base words for the two suffixes. This overlap further extends to bases that occur either as monosyllabic bound roots or as disyllabic words. In such cases, doublets are attested, with *-ify* attaching to the monosyllabic bases and *-ize* attaching to the disyllabic bases:

(11) sinicize/sinify, plastify/plasticize, technicize/technify

In other words, no extra restrictions are at work. Of course, given the (otherwise) complementary distribution, it is hard or impossible to decide which of the two suppletive suffixes should be regarded as the more special case (in van Marle's terms). Therefore the examples in (11) do not really argue against type-blocking itself. Let us therefore turn to the more straightforward case of *-ize/-ify* vs. *-ate*. Here, *-ate* should pre-empt the possibility of attaching *-ize* in its own domain. Crucially, this is not the case.

Phonologically, the domains of *-ize* and *-ate* overlap considerably, making *-ate* derivatives and *-ize* derivatives much more similar to each other than to verbs in *-ify*. Formations involving *-ize* and *-ate* have in common that both suffixes take secondary stress, both need an unstressed syllable preceding it, and both may involve the deletion of base-final segments. They differ in that formations in *-ate* have strictly alternating stress, while *-ize* may tolerate two preceding adjacent unstressed syllables under the specific conditions discussed in Plag (1999: 170-179, 185-187, 213-217). The stress patterns of the two kinds of derivatives may therefore contrast in such a way that *-ate* formations are always primarily stressed on the antepenult, while *-ize* verbs may also be stressed on the pre-antepenult. With trochaic bases the suffixes behave similarly.

With these semantic and phonological similarities in mind, let us look at potentially rival formations, i.e. those potential and attested forms that express the ornative-resultative meaning and conform to the semantic restriction on the base words ('chemical substance'). (12) lists those forms where *-ize* formations are attested. In all cases, the corresponding *-ate* derivative would be a possible, that is well-formed, derived verb:

(12)

not attested but possible	**attested**	**not attested but possible**	**attested**
!chemic(al)ate	chemicalize	!spheroidate	spheroidize
!radiumate	radiumize	!strychninate	trychninize
!saccharinate	saccharinize	!trypsinate	trypsinize
!silanate	silanize	!vitam(in)ate	vitaminize

The next set of forms is characterised by the fact that now it is the *-ate* derivatives that are attested, but not their corresponding *-ize* words, which seem equally possible and well-formed.

(13)

attested	**not attested but possible**
cannulate	!cannulize
fluorinate	!fluorinize
formylate	!formylize
hydroxylate	!hydroxylize
mercurate	!mercurize
protonate	!protonize

Finally, even two doublets can be found:

(14)

attested	**attested**
fluoridate	fluoridize
nitrogenize	nitrogenate

The patterning of the data in (12) to (14) reveal that type-blocking does not play any role in the distribution of *-ize* and *-ate*. Rather, where both suffixes are semantically and phonologically licensed, both can in principle be attached. Sometimes, only one of the two is attested, sometimes the other, sometimes both. Thus, the distribution can be accounted for without positing restrictions that go beyond the ones stated already for the individual affixes.

With regard to *-ate* vs. *-ify*, a similar picture emerges here as with *-ize* and *-ify*, in that the domains of *-ify* and *-ate* do hardly overlap. Since *-ate* is monosyllabic and subject to almost the same constraints as *-ize* derivatives, the distributional effects are very similar. Recall that *-ify* prefers iambic bases, *-ate* trochaic ones, with both types of derivatives needing strictly alternating stress. The only systematic overlap is (again) with disyllabic bases ending in an unstressed vowel, which, however, do not surface in the data, since the words denoting chemical substances do either end in other segments, or have more than two syllables. This makes, for example, *mercurate* (or !*mercurize*, for that matter) the only possible forms. Both suffixes may in principle form ornative/resultative verbs on the basis of nouns denoting chemical substances, but the actual domain where *-ate* and *-ify* are in competition is extremely small, due to the pertinent phonological restrictions.

We may now turn to the discussion of conversion in relation to the overt suffixes. Semantically, conversion is the most general case in that the meanings of the derivatives with overt suffixes are a subset of the possible meanings of converted verbs. This means that all of the bases attested with the overt suffixes could, in principle, have undergone conversion instead of overt affixation. What made the speakers choose the overt affixes instead? It seems that one reason for this choice lies in the more specific meaning the overt suffixes express in comparison to the completely indeterminate meaning of conversion. As was pointed out above, the interpretation of converted items relies on the linguistic and extra-linguistic context to an even greater extent than the interpretation of, say, -ize-derivatives. Thus, from the view of perception, overtly affixed forms are better than converted items. Furthermore, conversion does not apply to certain kinds of derived adjectives nor to derived nouns, which makes these classes of base words exclusively susceptible to overt suffixation. A look at the neologisms corroborates this, since, at least with *-ize*, a high proportion of base words are morphologically complex (see Plag 1999: appendix 1 for documentation).[3]

In fact, of the 488 converted verbs only 79 (17 deadjectival and 62 denominal ones) actually express meanings that are also associated with the overt suffixes. If we compare this figure with the number of *-ize* derivatives, we see that it is only about one fourth of the number of *-ize* neologisms in the *OED*. This difference can be interpreted in such a way that conversion is certainly not the most productive process in important semantic domains.

Incidentally, not all of the base forms of these 79 converted verbs could have been overtly suffixed, since there are other restrictions at work. One restriction holding for all overt suffixes which I have not yet discussed, is that compounds cannot be suffixed by *-ize*, *-ify* and *-ate*. However, compounds may readily be turned into verbs by conversion. Of the compound-based converted verbs in my corpus the following express meanings that are also found with overtly suffixed verbs: *cobweb, cold-cream, highlight, mothball, pothole, rustproof, scapegoat, streamline, waymark*. A parallel restriction seems to hold for base forms involving the prefixed elements *multi-* (*multiplex*), *cross-* (*cross-reference*) and *super-* (*supercoil*), which do not undergo overt suffixation either. The reason for this impossibility is probably not morphological but phonological. English compounds, as well as the forms with the said prefixes, seem to exhibit a stress pattern that is incompatible with the prosodic restrictions imposed on the derivatives involving overt suffixes.[4] Taking these restrictions into consideration the overlapping domain is further curtailed.

---

[3] The high frequency of complex bases with *-ize* is not only due to its semantics but also due to the toleration of stress lapses. Many suffixed adjectives, for example, are dactyls or end in a dactyl (e.g. *federal*).

[4] Thus, there seem to be no mechanisms available to accommodate compound stress to the stress patterns overtly affixed verbs must conform to. The details of this phenomenon still need to be worked out.

Note that derivatives on the basis of words featuring *photo-* as their first element (*photoisomerize, photosensitize, photosynthesize*) are not compounds. This is evidenced by

With regard to the remaining group of truly competing formations a number of overtly suffixed forms are indeed attested (some of them 20th century forms). Consider the forms in (15), where the converted verbs are listed in the left column, their suffixed rivals in the right column:

(15)

carbon	carbonize	polychrome	polychromatize
dolomite	dolomitize	propadanda	propagandize
dual	dualize	pressure	pressurize
filthy	filthify	pretty	prettify
gas	gasify	quinine	quinize
gel	gelate	romantic	romanticize
indemn	indemnify	satire	satirize
lethal	lethalize	zero	zeroize
phagocyte	phagocytize		

The majority of the derivatives on the left are synonymous to those on the right, which shows that token-blocking is not to be expected with low frequency items. This supports the view that token-blocking is not a structural, but a psycholinguistic mechanism, as put forward in Plag (1999: 54). There it is argued that not only idiosyncratic or simplex words can token-block productive formations, but stored words in general can, irrespective of their morphological make-up. Thus, new formations that are completely synonymous to already existing words, i.e. to words that are attested in the language, can and will arise in all those cases where the individual speaker has not stored - or can momentarily not get access to - the already existing one in his/her lexicon. Hence token-blocking is not "a prophylactic measure to avoid undesired synonymy", as claimed by Plank (1981: 182, my translation, I.P.), but the effect of word storage and word processing mechanisms.

In sum, the data show that the number of actually competing forms is much smaller than previously assumed, and that in the truly rival domains all affixes are applicable. Apart from the restrictions and properties of the individual affixes, type-blocking or other paradigmatic mechanisms are not in operation.

5. Conclusion

In this paper I have summarised the analysis of English derived verbs as put forward in Plag (1999). The semantic and phonological properties of these verbs were described and it was shown that the distribution of the different verb-deriving affixes is governed (only) by the semantic and phonological properties of each individual affix. This result

---

their stress pattern, which runs counter to the compound stress rule: the base words are pronounced with primary stress on the second element. In other words, *photo-* behaves like a combining form (Bauer 1983:213), and not like a noun. This analysis is in accordance with the meaning conveyed by *photo-* in the above derivatives: 'having to do with light', not 'having to do with photographs'.

strongly supports a syntagmatic, sign-based model of (productive) derivational morphology and casts serious doubts on models that want to separate meaning and form in derivational morphology, such as Beard's Lexeme-Morpheme Base Morphology (1995). Thus the semantic analysis of the suffixes provides strong arguments against the separation hypothesis, since separation cannot cope with the sign character of the individual verbal affixes (polysemy of one affix, only partial synonymy of putatively rival suffixes). The complete absence of type-blocking effects in the data indicates that type-blocking can be dismissed as a morphological mechanism. Only a minimum of paradigmatic mechanisms needs to be assumed, namely token-blocking and local analogy.

Works Cited

Anshen, Frank and Mark Aronoff (1988): "Producing Morphologically Complex Words", *Linguistics* 26, 641-655.

Aronoff, Mark (1980): "Contextuals", *Language* 56, 744-758.

Baayen, Harald (1989): *A Corpus-based Study of Morphological Productivity. Statistical Analysis and Psycholinguistic Interpretation.* Doctoral dissertation, Amsterdam: Vrije Universiteit.

Baayen, Harald. "Quantitative Aspects of Morphological Productivity", *Yearbook of Morphology 1991*, Eds. Geert Booij and Jaap van Marle. Dordrecht / Boston / London: Kluwer, 1992. 109-149.

Baayen, Harald (1993): "On Frequency, Transparency and Productivity", in *Yearbook of Morphology 1992*. Ed. by Geert Booij and Jaap van Marle, Dordrecht / Boston / London: Kluwer, 181-208.

Baayen, Harald and Rochelle Lieber (1991): "Productivity and English Word-formation: a Corpus-based Study", *Linguistics* 29, 801-843.

Baayen, Harald and Antoinette Renouf (1996): "Chronicling the Times: Productive Lexical Innovations in an English Newspaper", *Language* 72, 69-96.

Bauer, Laurie (1983): *English Word-formation*, Cambridge: Cambridge University Press.

Beard, Robert (1995): *Morpheme-Lexeme Base Morphology*, Albany: State University of New York Press.

Chitashvili, Revas J. and Harald Baayen (1993): "Word Frequency Distributions", in L. Hrebícek and G. Altmann (eds.), *Quantitative Text Analysis*, Trier: WVT, 54-135.

Clark, Eve and Herbert Clark (1979): "When Nouns Surface as Verbs", *Language* 55, 767-811.

COBUILD (1987): *Collins COBUILD English Language Dictionary*, London—Glasgow: Collins.

Don, Jan (1993): *Morphological Conversion*, Utrecht: OTS, Rijksuniversiteit.

Gussmann, Edmund (1987): "The Lexicon of English De-adjectival Verbs", in Edmund Gussmann (ed.), *Rules and the Lexicon*, Lublin: Catholic University, 79-101.

Jackendoff, Ray (1990): *Semantic Structures*, Cambridge: MIT Press.

Jespersen, Otto (1942): *A Modern English Grammar on Historical Principles*. Part VI *Morphology*, London: Allen and Unwin.

Karius, Ilse (1985): *Die Ableitung der denominalen Verben mit Nullsuffigierung im Englischen*. Tübingen: Niemeyer.

Kulak, Manfred (1964): *Die semantischen Kategorien der mit Nullmorphem abgeleiteten desubstantivischen Verben des heutigen Englischen und Deutschen*, Unveröffentlichte Dissertation: Universität Tübingen,.

Marchand, Hans (1964): "Die Ableitung desubstantivischer Verben mit Nullmorphem im Englischen, Französischen und Deutschen", *Die Neueren Sprachen* 63, 105-118.

Marchand, Hans (1969): *The Categories and Types of Present-day English Word-formation*, 2nd ed., München: Beck.

OED2 on CD (1994): *The Oxford English Dictionary, 2nd Edition, on Compact Disc*, Oxford: Oxford University Press.

Plag, Ingo (1998): "The Polysemy of *-ize* Derivatives. The Role of Semantics in Word Formation", *Yearbook of Morphology 1997*. ed. by Geert Booij and Jaap van Marle, Dordrecht: Kluwer, 219-242.

Plag, Ingo (1999): *Morphological productivity. Structural Constraints in English Derivation*, (Topics in English Linguistics 28), Berlin / New York: Mouton de Gruyter.

Plank, Frans (1981): *Morphologische (Ir-)Regularitäten*, Tübingen: Narr.

Rainer, Franz (1988): "Towards a Theory of Blocking", *Yearbook of Morphology 1988*, ed. by Geert Booij and Jaap van Marle, Dordrecht / Boston / London: Kluwer, 155-185.

Rose, James H. (1973): "Principled Limitations on Productivity in Denominal Verbs", *Foundations of Language* 10, 509-526.

Rubenstein, Herbert and Irwin Pollack (1963): "Word Predictability and Intelligibility", *Journal of Verbal Learning and Verbal Behavior* 2, 147-158.

Scarborough, Don, Charles Cortese and Hollis S. Scarborough (1977): "Frequency and Repetition Effects in Lexical Memory", *Journal of Experimental Psychology: Human Perception and Performance* 3, 1-17.

Szymanek, Bogdan (1985): *English and Polish Adjectives. A Study in Lexicalist Word-formation.* Lublin: Catholic University.

van Marle, Jaap (1985): *On the Paradigmatic Dimension of Morphological Creativity*, Dordrecht: Foris,.

van Marle, Jaap (1986): "The Domain Hypothesis: the Study of Rival Morphological Processes", *Linguistics* 24, 601-627.

Whaley, C. P. (1978): "Word-Nonword Classification Time", *Journal of Verbal Learning and Verbal Behavior* 17, 143-154.