Chapter 23 of *Handbook of English Linguistics*, ed. by Bas Aarts and April McMahon. Oxford: Blackwell.

preprint version of:
Plag, Ingo. 2006. Productivity. In Bas Aarts & April McMahon (eds.), *Handbook of English Linguistics*, 537-556. Oxford: Blackwell.

Productivity

Ingo Plag
University of Siegen

Version of 29 June 2004

1. Introduction

Speakers of English (and of course also of other languages) can coin new words on the basis of other words or word-forming elements. For example, we can turn the adjective *cute* into a noun *cuteness* by adding the suffix *-ness*, or we can form a new compound by joining two existing words, as in *train connection*. A closer analysis of such word-formation processes reveals that much of what happens in this domain is rule-governed, in the sense that there are predictable form-meaning relationships among similar morphologically complex words. For example, we can say that adjectives regularly can take the suffix *-ness* and that *-ness* derivatives regularly express a meaning that can be paraphrased as 'the property of being X', with 'X' standing for the meaning of the base.

Assuming the existence of such morphological rules, patterns or processes according to which complex words are formed, one can easily observe that some rules (or affixes) are quite often used to create new words, whereas others are less often used, or not used at all for this purpose. For example, it seems that no new verb can be formed in Modern English with the help of the prefix *en-* (as in *enlist, enroll, enshrine,* etc.), while the verbal suffix *-ize* happily adjoins to adjectives or nouns to make up new verbs (as in *peripheralize*, first attested 1987 and *Clintonize,* first attested 1992, both according to the *OED*).

In this sense, some morphological rules can be called productive and other rules unproductive or less productive. A number of interesting questions arise from this fact. What makes a given rule productive or unproductive? How we can meas-

ure the productivity of a given rule and which mechanisms are responsible for the variability in the productivity of morphological processes?

Another important theoretical problem is whether productivity should be regarded as a theoretical primitive, i.e. a non-derivable property of word formation rules, or an epiphenomenon, i.e. a property that results from other properties of the rule in question or some yet-to-be-detected mechanisms. It is clear, for example, that the productivity of a rule is never unrestricted in the sense that any given word may serve as its base. In particular, there can be phonological, morphological, syntactic, and semantic conditions on possible bases, or on the derivatives themselves, which may limit the productivity of the process.

The notion of productivity is relevant also for the common distinction between inflection and derivation (see chapter 22). It is commonly assumed (e.g. Haspelmath, 2002, p. 75) that inflectional processes are fully productive, whereas derivational processes are characterized by varying degrees of productivity, with the majority not being fully productive. In other words, inflectional processes apply to all words of a given word class, which is not the case for derivational processes. For example, all verbs in English can take the past tense morpheme, but not all verbs take the adjectivizing suffix *-ive* (*invent - invented - inventive*, *associate - associated - associative*, but *call - called - *callive*, *cite - cited - *citive*). Though intuitively appealing, there are some problems with the idea that inflection is fully productive. For example, one could argue that though fully productive as a category, the *regular* past tense affix { *-ed* } (with its three allomorphs [d], [t] and [ɪd]) is not fully productive, since there are quite a number of verbs which do not take one of these allomorphs, but use ablaut (e.g. *sang*, *dug*), change their stems (e.g. *brought*), take no overt suffix (e.g. *put*), or use a

combination of different coding strategies (e.g. *kept*). Such ill-behaved verbs are of course well known as 'irregular verbs', and, in order to save productivity as a distinguishing criterion between inflection and derivation, we could simply say that all *regular* inflection is fully productive while derivational morphology is not. This would, however, create the problem that regular derivational processes could also be said to be fully productive. Hence, productivity is an issue that seems not only relevant in word-formation but also in inflection. For reasons of space, we will confine our discussion of productivity in this chapter to derivational morphology.

Most of the more recent discussion on the nature of productivity has focused on English and empirical studies of productivity in other languages are still scarce. The reason for this state of affairs lies primarily in the availability of modern analytical tools, such as such as large electronic text corpora, lexical data bases and electronic dictionaries. English happens to be the language for which the these tools were readily available for the first time. It seems, however, that the findings and concepts developed using English as the sample language can be easily extended and applied to other languages, provided that the necessary methodological tools are available (cf. e.g. Evert and Lüdeling 2001 on German, Gaeta and Ricca 2003 on Italian).

2. Qualitative and quantitative aspects of productivity

One important theoretical question concerning the nature of productivity is whether productivity is a quantitative or a qualitative notion. If productivity is of a qualitative nature, a process or affix could be said to either have this property or not. Alterna-

tively, it has frequently been argued that productivity is a gradual phenomenon, which means that morphological processes are either more or less productive than others, and that completely unproductive or fully productive processes only mark the end-points of a scale. In the following subsection I will lay out the qualitative concept of productivity, which will be followed in section 2.2. by a discussion of approaches that have attempted to devise quantitative measures of productivity.

2.1. Qualitative approaches

Definitions of productivity can be found in any standard morphology textbook. Adams (1973, p. 197), for example, uses "the epithet 'productive' to describe a pattern, meaning that when occasion demands, the pattern may be used as a model for new items." Bauer (1983, p. 18) says that a word formation process is productive "if it can be used synchronically in the production of new forms", Spencer (1991, p. 49) considers a rule productive if it is "regularly and actively used in the creation of totally new words", and Plag (2003, p. 44) defines productivity as "[t]he property of an affix to be used to coin new complex words." These definitions may suggest that productivity is an all-or-nothing property of morphological processes. In one of the most recent monographs on productivity, Bauer (2001) explicitly advocates the all-or-nothing view, when, drawing on earlier work by Corbin (1987), he divides productivity into two distinct phenomena, one of them qualitative, the other quantitative in nature: availability and profitability. A morphological process is defined as available if it can be used to produce new words. "Availability is a yes/no question: either a process is available or it is not." (Bauer, 2001, p. 205). Profitability,

process is available or it is not." (Bauer, 2001, p. 205). Profitability, on the other hand, is the extent to which a morphological process may be employed to create new pertinent forms. This is a quantitative notion, and we will postpone the discussion of profitability until later.

The most problematic point concerning availability is the notion of 'morphological process' (or often called 'word formation rule') itself. Given a set of seemingly related words, on which grounds can one assume the existence of a word-formation rule as being responsible for the creation of these words? In general one would say that we can speak of a rule if there is a sufficient number of regular form-meaning correspondences of individual items, i.e. a recognizable pattern. The theoretical status of such patterns is however controversial. Some scholars believe that what has been traditionally called 'rule' or 'process' is just a larger set of words that are related to one another by the very general mechanism of analogy (e.g. Becker 1990, or, more recently, Skousen et al. 2002). And this analogical mechanism can also be used to coin words on an individual, idiosyncratic basis, which is what earlier, or more traditional, accounts of analogy are more concerned with. The problem now is that in a purely qualitative approach to productivity, an unproductive process would not be able to give rise to new formations at all. Empirically, however, we find that supposedly unproductive processes sometimes do yield new formations, because speakers use existing derivatives to form new words by way of proportional analogy. If this only happens once or twice, we might still say this is an unproductive rule, but where would we draw the line between productive and unproductive processes, if more words are coined? Would we say a process is productive after we have found two, three, five, ten, or twenty new analogical forms?

These considerations lead to the conclusion that even in a qualitative approach to productivity one has to assume the existence of three types of processes: Those that are clearly unproductive (with not even occasional analogical coinages), those that are clearly productive, and those processes that are not easily classified as either productive or unproductive. This is also acknowledged by Bauer, when he writes that "there might be cases of uncertainty" (2001, p. 205) with regard to the availability of a word-formation process.

In view of these problems, many researchers have abandoned the idea of a qualitative notion of productivity and have turned to the exact determination of what was introduced above as 'profitability'. These researchers have sought measures by which the productivity (here: profitability) of processes can be assessed, to the effect that totally unproductive and fully productive processes are conceptualized as end-points on a scale.

2.2. Quantitative approaches

A good starting point for quantitative measures of productivity is the definition by Bolinger (1948), which is based on the idea that productivity can be seen as a kind of probability. In his words, productivity is "the statistical readiness with which an element enters into new combinations" (p. 18). Since the formulation of this definition more than half a century ago, a number of productivity measures have been proposed that try to model the insight behind this definition.

One prominent definition says that the productivity of an affix can be measured by counting the number of attested types (i.e. different words) with that affix at a given point in time, for example by counting the numer of pertinent forms in an unabridged dictionary. The problem with this measure is that there can be many words with a given affix, but nevertheless speakers will not use the suffix very often to make up new words. In other words, the fact that the language has already many words with a given affix indicates that the suffix must have been productive at some period in the past. For example, many words with the nominalizing suffix *-ment* (*entertainment, punishment,* etc.) can be found, but the suffix was mainly productive between the mid-sixteenth and the mid-nineteenth century (e.g. Bauer, 2001, p. 181). Similarly, the verbalizing suffix *-en* (as in *blacken*) is attested in numerous words, but hardly any of them was coined after 1900 (e.g. Plag, 1999, p. 98).

Aronoff (1976) suggests a different productivity measure, the ratio of actual to possible words. 'Actual word' refers to existing established words with a given affix, while 'possible word' (or 'potential word') refers to words which could in principle be formed with that affix. The higher this ratio, the higher the productivity of a given rule. Largely ignored by later authors, this measure had already been proposed earlier by Berschin, who labeled it "Besetzungsgrad" ('degree of exhaustion', 1971, pp. 44-45). Anshen & Aronoff (1981, p. 64) point out the main weakness of this proposal: for extremely productive and for completely unproductive processes it makes wrong predictions. Thus, with highly productive affixes like *-ness* the number of potential words is, in principle, infinite, which necessarily leads to a comparatively low productivity index. With unproductive rules like *-th* nominalization it is unclear how the ratio of actual to possible words should be calculated. If one considers all actual

words with this suffix as possible words, the ratio equals 1, which is the highest possible score and therefore counterintuitive. If, however, the number of possible words with this suffix is considered zero, the index cannot be computed at all.

Another, more general problem of Berschin's and Aronoff's proposals is how to actually count the number of possible words, since the number of possible formations on the basis of a productive rule is, in principle, uncountable, because new potential base words (e.g. new adjectives as bases for *-ness*) may enter the language any time. How can one quantify something that is, in principle, uncountable?

Coming back to the idea of counting the number of derivatives, one can say that this may still be a fruitful way of determining the productivity of an affix, namely if one does not count all derivatives with a certain affix in use at a given point in time, but only those derivatives that were newly coined in a given period, the so-called neologisms. In doing this, one can show that, for instance, an affix may have given rise to many neologisms in the 18th century but not in the 20th century. The number of neologisms in a given period is usually determined with the help of historical dictionaries like the *OED*, which aims at giving thorough and complete information on all words of the language. For example, for the period from 1900 through 1985 we find 284 new verbs in *-ize* (Plag, 1999, chapter 5) in the *OED*, which shows that this is a productive suffix. The power of the *OED* as a tool for measuring productivity should however not be overestimated, because quite a number of new words escape the eyes of the *OED* lexicographers. For instance, the number of *-ness* neologisms listed in the *OED* for the 20th century (*N*=279, Plag, 1999, p. 98) roughly equals the number of *-ize* neologisms, although it is clear from many studies that *-ness* is much more productive than *-ize* (e.g. Plag et al. 1999, Hay & Baayen 2002).

Thus, in those cases where the *OED* does not list many neologisms it may be true that the affix is unproductive, but it is also possible that the pertinent neologisms simply have been overlooked (or not included for some other, unknown reason). Only in those cases where the *OED* lists many neologisms can we be sure that the affix in question must be productive. Given these problems involved with dictionary-based measures (even if a superb dictionary like the *OED* is available), one should also look for other, and perhaps more reliable measures of productivity.

Harald Baayen and his collaborators (1993 et seq.) have developed some corpus-based productivity measures, which all rely on the availability of very large electronic text corpora. Such corpora are, for example, the British National Corpus (BNC) or the Cobuild Corpus, the former containing c. 100 million word tokens, the latter originally containing c. 18 million words, now having been turned into the ever-increasing Bank of English. The word lists that can be extracted from such corpora are the basis for corpus-based productivity research.

The first corpus-based measure to be mentioned here is the number of types, i.e. different words with a given affix. This measure, also known as the type-frequency *V*, has been discussed above, only that it is calculated here not on the basis of a dictionary, but on the basis of a representative language sample.

Two other measures proposed by Baayen rely heavily on the notion of hapax legomenon. Hapax legomena (or 'hapaxes' for short) are words that occur only once in a corpus. Such words are crucial for the determination of the productivity of a morphological process because in very large corpora hapaxes tend to be words that are unlikely to be familiar to the hearer or reader. Complex unknown words can be understood at least in those cases where an available word-formation rule allows the

decomposition of the newly encountered word into its constituent morphemes and thus the computation of the meaning on the basis of the meaning of the parts. The word-formation rule in the mental lexicon guarantees that even complex words with extremely low frequency can be understood. Thus, with regard to productive processes, we expect large numbers of low frequency words and small numbers of high frequency words, with the former keeping the rule alive. In contrast, unproductive morphological categories will be characterized by a preponderance of words with rather high frequencies and by a small number of words with low frequencies.

The crucial point now is that, even if not all of the hapaxes with a given affix may be neologisms, we can be confident that it is among the hapaxes (as against words that have a higher frequency) that we find the highest proportion of neologisms (see, for example, Baayen & Renouf 1996, Plag 2003 for discussion). Given that the number of hapaxes of a given morphological category should correlate with the number of neologisms of that category, the number of hapaxes can be seen 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, and could therefore simply be a rare word of the language (instead of a newly coined derivative) or some weird ad-hoc invention by an imaginative speaker, as sometimes found in poetry or advertisement. The latter kinds of coinages are, however, extremely rare and can be easily weeded out.

The size of the corpus plays an important role in determining the nature of hapaxes. When the corpus is small, most hapax legomena will indeed be well-known words of the language. However, as the corpus size increases, the proportion of neologisms among the hapax legomena increases, and it is precisely among the hapax

legomena that the greatest number of neologisms appear. The number of hapaxes is therefore an important measure for estimating the productivity of a morphological process.

There are, of course, methodological problems that need to be considered. First, as already mentioned, there is the question of corpus size. Small corpora like the 1 million word Wellington Corpus of Written New Zealand English are certainly too small for this kind of approach (cf. Bauer, 2001, pp. 150f). Furthermore, there seem to be some rare cases of morphological categories where the proportion of neologisms among the hapaxes is unexpectedly low (see Plag, 1999, pp. 112f). Other methodological problems concern the determination of pertinent word forms, involving sometimes empirically and theoretically problematic decisions. For example, it is not so easy to develop consistent criteria for or against the inclusion of words such as *entity, quantity, celebrity* as *-ity* derivatives. Such forms occur in abundance in English especially because this language has borrowed a large stock of its vocabulary from other languages (e.g. French, Latin, Greek). Often such words were morphologically complex in the donor languages but were not necessarily decomposed in the borrowing process. If many words with the same affix were borrowed, however, this may have eventually led to the reanalysis of most words of the category and even to a more or less productive derivational process in English, but with a residue of words, whose status as complex words remained questionable (see Dalton-Puffer 1996 for some discussion). In general, the so-called Latinate affixes seem less productive than native affixes (e.g. Plag 2003: chapters 4 and 7). Apart from borrowing, problems of classification can also arise through lexicalization, a process in which a complex word can adopt new and idiosyncratic senses which are no longer identical with the

general meaning of the morphological category. For example, *curiosity* has the predictable meaning of 'property of being curious', but it has also lexicalized the rather idiosyncratic meaning 'curious thing'.

In general the above-mentioned problems of classification are inherent in all work on derivational morphology and not restricted to a particular language or to corpus-based investigations (see Plag, 1999, chapter 5, or Bauer, 2001, section 5.3 for more discussion).

Coming back to the idea of estimating the probability with which new words are coined, we turn to Baayen's 'productivity in the narrow sense'. This measure calculates the ratio of the number of hapaxes with a given affix and the number of all tokens containing that affix. Metaphorically speaking, when calculating this measure we are going through all attested tokens with a given affix and picking out all words that we encounter only once. If we then divide the number of these words (i.e. the number of hapaxes) by the number of all tokens with that affix, we arrive at the probability of finding a hitherto unattested word (i.e. 'new' in terms of the corpus) among all the words of that category. This probability can be expressed by the following formula, where P stands for 'productivity in the narrow sense', n_1^{aff} for the number of hapaxes with a given affix and N^{aff} stands for the number of all tokens with that affix.

(1)
$$P = \frac{n_1^{aff}}{N^{aff}}$$

P can be interpreted in such a way that a large number of hapaxes leads to a high value of *P*, thus indicating a productive morphological process. Conversely, large numbers of high frequency items lead to a high value of N^{aff}, hence to a decrease of *P*, indicating low productivity.

To summarize our review of different productivity measures, we can distinguish between the following methods:

- Using a text corpus or a large dictionary, productivity can be measured by counting the number of attested different words with a particular affix (i.e. the type-frequency *V*). The greater the type-frequency, the higher the productivity of the affix. This measure is, however, indicative of past, rather than present productivity.

- Productivity can be measured by counting the number of neologisms in a given period, using, for instance, a large historical dictionary. The greater the number of neologisms in that period, the higher the productivity of a given affix in that period.

- Productivity can be measured by counting the number of hapaxes with a given affix (n_1) in a large corpus. The higher the number of hapaxes, the greater the productivity.

- Finally, by dividing the number of hapaxes with a given affix by the number of tokens with that affix, we arrive at *P*, which indicates the probability of finding new words among all the tokens of a particular morphological category.

For illustration and discussion of the different productivity measures, let us look at some suffixes for which these measures are readily available, *-ion, -ist, -ity, -ish, -less, -ness* and *wise* (from Plag et al. 1999, Plag 2002, based on data from BNC and *OED*).

Table 1: Productivity measures and token frequencies of some affixes in the BNC and OED

	V	N^{aff}	n_1^{aff}	***P***	**OED neologisms**
-ion	2392	1369116	524	0.00038	625
-ish	491	7745	262	0.0338	101
-ist	1207	98823	354	0.0036	552
-ity	1372	371747	341	0.00092	487
-less	681	28340	272	0.0096	103
-ness	2466	106957	943	0.0088	279
-wise	183	2091	128	0.061	12

The table raises the question of which suffix is most productive. Let us first regroup the table according to each measure in the descending order of their values.

Table 2: Ranking of suffixes according to different measures of productivity

Rank	***V***		***N***		***n_1***		***P***		**OED neologisms**	
1	*-ness*	2466	*-ion*	1369116	*-ness*	943	*-wise*	0.061	*-ion*	625
2	*-ion*	2392	*-ity*	371747	*-ion*	524	*-ish*	0.0338	*-ist*	552
3	*-ity*	1372	*-ness*	106957	*-ist*	354	*-ness*	0.0096	*-ity*	487
4	*-ist*	1207	*-ist*	98823	*-ity*	341	*-less*	0.0088	*-ness*	279
5	*-less*	681	*-less*	28340	*-less*	272	*-ist*	0.0036	*-less*	103
6	*-ish*	491	*-ish*	7745	*-ish*	262	*-ity*	0.00092	*-ish*	101
7	*-wise*	183	*-wise*	2091	*-wise*	128	*-ion*	0.00038	*-wise*	12

Table 2 reveals that each measure establishes a different productivity ranking, such that the different measures seem to contradict each other. However, as we will shortly see, this is not the case, since the different measures highlight different aspects of productivity.

The adverb-forming suffix *-wise* seems to be the most extreme case. While of highest productivity according to *P* it is of extremely low productivity according to the other measures. How can this paradox be solved? The low rank of *-wise* in terms of *V* and n_1 is an indication of the fact that it is a suffix that is used comparatively rarely. Not very many derivatives are used nor are very many newly coined. However, the high value of *P* shows that among all types with the suffix *-wise* the number of new coinages is quite high, such that the proportion of unknown words among all the *-wise* derivatives is high, indicating the suffix's potential to be easily used for the

coinage of new forms, if need be. A look at some forms attested in the BNC supports this impression (cited from Dalton-Puffer & Plag, 2000, 237):

(2) a. Bridhe lifted the baby, slipped a magic coral and rowan-berry necklace over his head and walked **sun-wise** round the bed three times for good fortune,

b. They make no special demands **food-wise**, and tolerate a wide pH range.

The *OED* ranking reflects the fact that *-wise* words are, though easily derivable, not often used. The suffix *-ish* is very similar to *-wise* in this respect.

Turning to *–ion, -ity*, *–ist*, and *-less,* we can state that according to type-frequency, number of hapaxes and number of neologisms the suffixes *–ion, -ity*, and *–ist* must be regarded as quite productive, whereas the suffix *–less* is less productive. However, according to the *P* measure, the situation is exactly the opposite: *–less* must be regarded as more productive, and the suffixes *–ion, -ity*, and *–ist* as ranking very low on the scale. This apparent contradiction can be solved in the following way. The suffix *–less* does not occur in very many different words, and these words are also not so frequently used, hence the lower *V* and *N* figures, and the comparatively small number of hapaxes and *OED* neologisms. If we, however, only consider the words within this morphological category, we find that the *proportion* of hapaxes among all tokens is very high, which means that there is a high probability of finding new forms among all the words with *-less*. And this high probability is expressed by a high *P* measure. In less technical terms, the apparent contradiction can be explained by saying that we obviously don't use *-less* words a lot, but it is very easy to coin new ones. The opposite is the case for the categories of *-ion, -ity*, and *-ist* words. Each of these

categories contains many different words, but these are on average of comparatively high frequency, and the chance of finding a newly coined word among all tokens of one of these categories is comparatively low. In other words, these suffixes are very often used with existing words, but in comparison to the many words we use, we do not so often coin new ones.

Finally, *-ness* scores high in terms of type-frequency and neologisms, but due to the high number of tokens (many *-ness* words are quite frequent, e.g. *happiness*) *P* is lower than that of *-wise* and *-less*. Taking all the different aspects together, *-ness* is the most productive suffix of all. It has a relatively high productivity in the narrow sense and is at the same time also used in a great number of derivatives. The comparatively low number of *OED* neologisms is indicative of the problematic data collection method mentioned already above.

In sum, we can say that researchers have a number of different measures at their disposal to assess the productivity of word-formation processes. Each measure highlights different aspects of productivity and brings with it special methodological problems of data sampling and data analysis. In order to make sound statements about 'the' productivity of a given affix different measures should be taken into account and be interpreted carefully in the light of the methodological problems involved in their computation.

Having clarified the notion of productivity and how productivity can be measured, we may now turn to the problem of how speakers know whether they can use a given affix for the creation of new words. As we will shortly see, this has to do with the question mentioned above whether the productivity of a rule is an inherent,

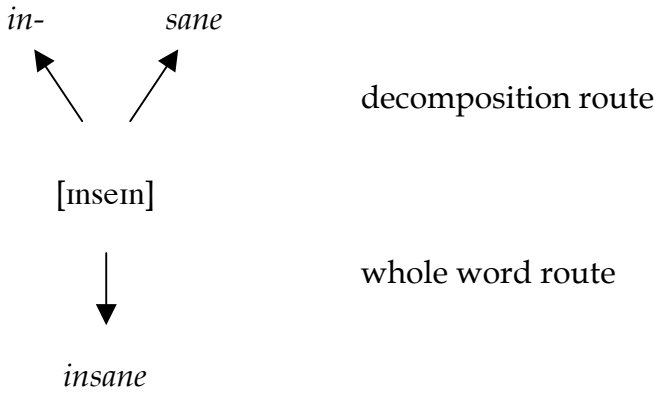
primitive part of that rule or a property derivable on the basis of other properties. We will deal with these issues in the next section.

3. Psycholinguistic aspects: Productivity and the mental lexicon

How can speakers know that a given affix can be used to coin new words? What do productive processes have in common that unproductive processes do not have? Which properties of affixes give rise to different degrees of productivity? In this (and also in the next) section, we will try to answer these questions, making reference to recent psycholinguistic research.

In the previous section we introduced productivity measures that make crucial reference to the frequency of lexical items. The basic reasoning behind the use of frequency in computing productivity is that the frequency of complex words significantly influences the way in which we process and store them. In most current models of morphological processing, access to morphologically complex words in the mental lexicon works in two ways: by direct access to the whole word representation (the so-called **'whole word route'**) and by access to the decomposed elements (the so-called **'decomposition route'**) (see McQueen and Cutler 1998 for an overview). This means that each incoming complex word is simultaneously processed in two ways, with one way of access finally succeeding. On the decomposition route it is decomposed in its parts and the parts are being looked up individually, on the whole word route the word is looked up as a whole in the mental lexicon. The two routes are schematically shown in (3):

(3)



How does frequency come in here? According to Hay (2000, 2001), the degree of decomposability of a given word depends crucially on the relative frequency of the derived word and its base. Relative frequency is defined as the ratio of the frequency of the derived word to the frequency of the base and measures how frequent the derivative is with respect to its base:

(4)
$$f_{relative} = \frac{f_{derivative}}{f_{base}}$$

With most complex words, the base is more frequent than the derived word, so that the relative frequency is smaller than unity. In psycholinguistic terms, the base has a stronger representation, or higher 'resting activation', in the mental lexicon than the derived word. This leads to a preponderance of the decomposed route, since due to its high resting activation, the base will be accessed each time the derivative enters the system. In the opposite case, when the derived word is more frequent than the base, there is a whole word bias in parsing, because the resting activation of the base is lower than the resting activation of the derivative. For example, *business* is much more frequent than its base *busy* (35141 vs. 4879 occurrences in the BNC), so that *business* will have a whole word bias in access. Note that *business* (in the sense of 'company', 'economic transactions' and related meanings) is also semantically and phonologically opaque, which is often the case with derivatives that have strong, i.e. lexicalized, whole word representations. Conversely, *blueness* has a base that is much more frequent than the derived form (10059 vs. 39 in the BNC), so that there will be a strong advantage for the decomposed route. In general, the higher the frequency of the derived word in relation to the base word, the less likely is decomposition. Alternatively, the lower the frequency of the derived word in relation to the base word, the more likely is decomposition.

Hay shows that relative frequency also patterns with other properties of morphological categories: low relative frequency correlates with high productivity and

low relative frequency correlates with high semantic transparency. These correlations do not come as a surprise. As already discussed in the previous section, productive morphological processes are characterized by a high number of low frequency words (i.e. many hapaxes, if we speak in terms of corpora). The lower the frequencies of derived words the lower their relative frequencies (holding the frequency of the base constant). Thus productive processes have a preponderance of words with low relative frequencies, whereas less productive morphological categories are characterized by a preponderance of words with higher relative frequencies. In a detailed study of the relation between parsing and productivity involving 80 affixes of English, Hay & Baayen (2002) demonstrate that the more morphologically decomposable forms containing a given affix are in the lexicon, the more productive that affix will be. Thus, there is a strong relationship between relative frequency, parsing in perception and morphological productivity. Increased rates of parsing lead straightforwardly to increased productivity.

The fact that productive morphological categories are characterized by a high proportion of decomposable words is also responsible for the fact that productive processes exhibit a preponderance of semantically and phonologically transparent formations. This correlation between transparency and productivity has been established in many earlier publications (e.g. Aronoff & Schvaneveldt 1978, Anshen & Aronoff 1981, Cutler 1981).

We can now see that productive categories are semantically transparent as a consequence of processing, since productive processes favor the decomposed route, and decomposed storage strengthens the individual semantic representations of the constituent morphemes. Decomposition and individual storage of the constituent

morphemes thus leaves little room for semantic drift and opacity, which arise easily under whole word access and storage, where the meanings of the parts are less likely to be activated. Hence semantic opacity and low productivity go hand in hand with high relative frequencies.

The relationship between phonological transparency and productivity is further substantiated in Hay & Baayen (in press), who investigate the role of junctural phonotactics with the 80 affixes from the earlier study. The term 'junctural phonotactics' refers to the possible combination of sounds that straddle a morphological boundary or juncture, as for example /n-a/ in the word *combi**n-a**tion*. Hay and Baayen (in press) start out from the assumption that speakers rely on phonotactics for the (pre-)processing of morphologically complex words. In pre-lexical processing, speakers posit morphological boundaries inside phoneme transitions that are unlikely to occur inside mono-morphemic words (see, e.g., Saffran et al. 1996a, 1996b, McQueen 1998). For example, the phoneme transition /pf/ (as in *cu**p-f**ul*) never occurs inside mono-morphemic English words and will therefore strongly facilitate decomposition in speech perception, while the transition /tɪ/ (as in *produc**t-i**ve*) has a much higher probability of occurring morpheme-internally and will therefore not facilitate decomposition. Hay & Baayen now argue that decomposition in speech perception leads to decomposed forms in the lexicon. And, if, as stated above, decomposed forms in the lexicon lead to productivity, it can be predicted that there is a relationship between the junctural phonotactics associated with an affix, and that affix's productivity. This prediction is borne out by the facts. Hay & Baayen find a significant correlation between the kind of junctural phonotactics of an affix and that affix's productivity. Roughly speaking, the more illegal the phonemic transitions cre-

ated by an affix are, the more productive that affix tends to be. Thus, phonotactics contributes probabilistically to the likelihood of decomposition and therefore to the degree of productivity.

To summarize, we can say that, psycholinguistically, productivity can be explained as a syndrome of properties, with parsability, relative frequency, semantic and phonological transparency as important factors. With regard to the question whether productivity is a derived notion or a theoretical primitive, we have seen that the productivity of an affix results in a complex fashion from the above-mentioned processing factors. Among these factors, semantic and phonological transparency are not only psycholinguistically, but also structurally determined in that it is the semantic and phonological structure of affixes and their derivatives that co-determine processing and storage of these forms. In the following, we will see that there are many more structural factors that play a significant role in influencing - and constraining - productivity. It is these factors that are responsible for the fact that Hay & Baayen's findings are not exceptionless principles but strong probabilistic tendencies, which are sometimes overruled by structural restrictions (see Plag 2002 for discussion).

4. Productivity restrictions

One important factor restricting the productivity is of course the usefulness of a newly-coined word for the speakers of the language. No matter which function a particular derivative serves in a particular situation, intended usefulness is a necessary prerequisite for the emergence of productively formed derivatives. But not all

potentially useful words are actually created and used, which means that there must be certain restrictions at work. We must distinguish between, on the one hand, the general possibility to apply a word-formation rule to form a new word and, on the other hand, the opportunity to use such newly coined derivatives in speech. Both aspects are subject to different kinds of restriction, namely those restrictions that originate in problems of language use (so-called pragmatic restrictions) and those restrictions that originate in problems of language structure (so-called structural restrictions). We will discuss each type of restriction in turn.

4.1. Pragmatic restrictions

One of the most obvious usage-based factors influencing productivity is fashion. The rise and fall of affixes like *mega-, giga-, mini-* or *-nik* is an example of the result of extra-linguistic developments in society which make certain words or morphological elements desirable to use and therefore productive.

Another pragmatic requirement new lexemes must meet is that they denote something nameable. Although the nameability requirement is rather ill-defined, it captures a significant insight: the concepts encoded by derivational categories tend to be rather simple and general (e.g. adjectival *un-* 'not X', verbal *-en* 'make X', etc.) and may not be highly specific or complex, as illustrated in the example of an unlikely denominal verb forming category given by Rose (1973, p. 516): "grasp NOUN in the left hand and shake vigorously while standing on the right foot in a 2.5 gallon galvanized pail of corn-meal-mush". This does not mean, however, that more complex

notions cannot be encoded by affixes, but that this requirement seems to be language-specific and is a mere tendency.

The problem with pragmatic restrictions is that, given a seemingly impossible new formation, it is not clear whether it is ruled out on structural grounds or on the basis of pragmatic considerations. Before claiming that a certain form is impossible due to pragmatic restrictions, it is therefore necessary to take a closer look at the structural restrictions involved, which often reveal that a form is impossible because it violates pertinent phonological, morphological, syntactic, or semantic restrictions.

5.2. Structural restrictions

Structural restrictions (or constraints) in word-formation may concern the traditional levels of linguistic analysis, i.e. phonology, morphology, syntax and semantics. A general question that arises from the study of such restrictions is which of these should be considered peculiar to the particular word-formation rule in question and which restrictions are of a more general kind that operate on all (or at least some classes of) morphological processes (see Plag, 1999, chapter 3, or Bauer, 2001, pp. 126-143 for a detailed discussion of both kinds of restrictions) .

Rule-specific constraints may concern the properties of the base or of the derived word. Let us start with phonological constraints, which can make reference to individual sounds or to prosodic phenomena such as syllable structure or stress. For example, suffixation of verbal *-en* (as in *blacken*) is subject to the segmental restriction

that it only attaches to base-final obstruents (cf., e.g., *blacken* vs. **finen*) and to the prosodic restriction that it does not take bases that have more than one syllable.

Apart from being sensitive to phonological constraints, affixation may depend on the morphological structure of the pertinent base words. An example of such a morphological constraint is the suffix combination *-ize-ation*. Virtually every word ending in the suffix *-ize* can be turned into a noun only by adding *-ation*. Other conceivable deverbal nominal suffixes, such as *-ment, -al, -age* etc., are systematically ruled out by this morphological restriction imposed on *-ize* derivatives (cf., for example, *colonization* vs. **colonizement, *colonizal* or **colonizage*).

The suffix *-ee* (as in *employee*) illustrates a semantic restriction. Derivatives with that suffix must denote sentient entities, as shown, for example, by the impossibility to use *amputee* to refer to an amputated limb (see Barker 1998 for detailed discussion).

Finally, productivity restrictions can make reference to syntactic properties. One of the most commonly mentioned ones is the restriction of word-formation rules to members of a certain syntactic category. An example would be the adjectival suffix *-able* which normally attaches to verbs (as in *readable*), or the adjectival suffix *-al*, which attaches to nouns (as in *circumstantial*).

Let us now look at one productivity restriction that is of a more principled kind, blocking. The term 'blocking' has been used in various senses in the literature. Our discussion will be restricted to two kinds of synonymy blocking, token-blocking and type-blocking (Rainer 1988). Token-blocking involves the blocking of a potential regular form by an already existing synonymous word, an example of which is the blocking of **arrivement* by *arrival* or **stealer* by *thief*. In contrast, type-blocking con-

cerns the blocking of the application of one rule by another rival rule (for example *-ness* and *-ity* suffixation).

Token-blocking is a relatively uncontroversial notion and will therefore not be discussed in great detail. One important aspect of token-blocking deserves mentioning, however, namely that it crucially depends on frequency. Contrary to earlier assumptions, Rainer (1988) shows that not only idiosyncratic or simplex words (like *thief*) can block productive formations (such as **stealer*), but that stored words in general can do so. As already discussed above, the storage of words is largely dependent on their frequency. Now, in order to be able to block a potential synonymous formation, the blocking word must be sufficiently frequent. In Rainer's experiment, the higher the frequency of a given word, the more likely it was that the word blocked a rival formation. Both idiosyncratic words and regular complex words are able to block other forms, provided that the blocking word is stored.

That such an account of blocking is on the right track is corroborated by the fact that occasionally really synonymous doublets do occur (which may later develop different meanings, e.g. *passivate/passivize*). Plank (1981, pp. 181-182) already notes that blocking of a newly derived form does not occur in those cases where the speaker fails to activate the already existing alternative form. The likelihood of failing to activate a stored form is negatively correlated to the frequency of the form to be accessed. In other words, the less frequent the stored word is the more likely it is that the speaker will fail to access it (and apply the regular rule instead), and the more frequent the stored word is the more likely it is that the speaker will successfully retrieve it, and the more likely it is, therefore, that it will block the formation of a rival word. With frequency and storage being the decisive factors for token-blocking, the

theory can naturally account for the occasional occurrence even of synonymous doublets.

We may now move on to the notion of type-blocking, which has been said to occur when a certain affix blocks the application of another affix (e.g. Aronoff 1976). The example *decency* vs. *decentness* would be a case in point. The crucial idea underlying the notion of type-blocking is that rival suffixes (such as *-ness, -ity,* and *-cy*) are organized in such a way that each suffix can be applied to a certain domain. In many cases one can distinguish between affixes with an unrestricted domain, the so-called general case (e.g. *-ness* suffixation, which may apply to practically any adjective), and affixes with restricted domains, the so-called special cases (for example *-ity* or *-cy* suffixation). The latter are characterized by the fact that certain constraints limit the applicability of the suffixes to a lexically, phonologically, morphologically, semantically or otherwise governed set of bases. Type-blocking would occur when the more special affix precludes the application of the more general affix.

The problem with this idea of type-blocking is that it cannot account for the patterning of the data. For example, Aronoff (1976, p. 53) regards formations involving nominal *-ness* as ill-formed in all those cases where the base adjective ends in *-ate, -ent* or *-ant*, hence the contrast between *decency* and what he considers an illegal form **decentness*. In his view, the systematic special case *-cy* (*decency*) precludes the general case *-ness*. There are, however, a number of problems with this kind of analysis. The first one is that, on closer inspection, *-ness* and its putative rivals *-ity* or *-cy* are not really synonymous, so that blocking could - if at all - only occur in those cases where the meaning differences would be neutralized. Riddle (1985) shows that there is in

fact a slight but consistent meaning difference observable between rival *-ness* and *-ity* derivatives. Consider, for example, the pair in (5) (from Riddle, 1985, p. 438):

(5) a. The lanterns demonstrated the *ethnicity* of the restaurant.

b. The lanterns demonstrated the *ethnicness* of the restaurant.

In (10a) the lanterns show to which ethnic group the restaurant belongs, whereas in (10b) the lanterns show that the restaurant has an ethnic appeal (as opposed to a non-ethnic appeal). In general, *-ness* formations tend to denote an embodied attribute, property or trait, whereas *-ity* formations refer to an abstract or concrete entity. Hence *-ity* and *-ness* are not completely synonymous, which would be a prerequisite for type-blocking. The second problem of the notion of type-blocking concerns the status of forms like *decentness*, which are in fact attested (a search on the internet yielded 279 occurrences, www.google.com, 28/08/2003) and even listed in dictionaries, hence not at all morphologically ill-formed. Furthermore, the occurrence of many attested doublets rather indicates that the domain of the general case *-ness* is not systematically curtailed by *-ity* or *-cy*: *destructiveness - destructivity, discoursiveness - discoursivity, exclusiveness - exclusivity, impracticalness - impracticality, inventibleness - inventability, naiveness - naivity, ovalness - ovality, prescriptiveness - prescriptivity* (all from the *OED*). The final problem with putative cases of type-blocking is to distinguish them from token-blocking. Thus, putative avoidance of *decentness* could equally well be a case of token-blocking, since one can assume that, for many speakers, the word *decency* is part of their lexicon, and is therefore capable of token-blocking (for a detailed discussion of affixal rivalry, see also Plag, 1999, chapter 8).

To summarize our discussion of blocking, we have seen that type-blocking as a general factor constraining productivity is problematic, while token-blocking restricts the productivity of affixes by preventing the formation of complex rival synonymous forms.

5. Conclusion

In this chapter we have looked at what it means when we say that a word-formation process is productive. The productivity of a given affix can be seen as its general potential to be used to create new words and as the degree to which this potential is exploited by the speakers. This degree can be assessed by various measures, both corpus-based and dictionary-based. We then discussed how complex words are stored and accessed in the mental lexicon, which is crucial for an understanding of the notion of productivity in word-formation. Productivity has been shown to be a derived notion. It emerges from the mental lexicon as the result of different properties, such as parsability, relative frequency, semantic and phonological transparency. Differences in productivity between affixes also raise the question of productivity restrictions. We have seen that apart from constraints on processing and usage, structural constraints also play an important role in restricting productivity. Possible words of a given morphological category need to conform to very specific phonological, morphological, semantic and syntactic requirements. These requirements restrict the set of potential complex words, thus limiting productivity. Finally, token-blocking was discussed, which is a general psycholinguistic mechanism which pre-

vents complex forms from being formed if a synonymous word is already available in the speaker's mental lexicon.

References

Adams, V. (1973). *An introduction to English word-formation*. London: Longman.

Anshen, F. & Aronoff, M. (1981). Morphological productivity and phonological transparency. *Canadian Journal of Linguistics*, 26, 63-72.

Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge: MIT Press.

Aronoff, M. & Schvanefeldt, R. (1978). Testing Morphological Productivity. *Annals of the New York Academy of Science*: *Papers in Anthropology and Linguistics*, 318, 106-114.

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij & J. van Marle (Eds). *Yearbook of morphology 1992.* (pp. 181-208). Dordrecht/Boston/London: Kluwer.

Baayen, H. & Renouf, A. (1996). Chronicling *The Times*: Productive lexical innovations in an English newspaper. *Language,* 72, 69-96.

Barker, C. (1998). Episodic *-ee* in English: A thematic role constraint on a new word formation. *Language*, 74, 695-727.

Bauer, L. (1983). *English word-formation*. Cambridge: Cambridge University Press.

Bauer, L. (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.

Becker, T. (1990). *Analogie und Morphologische Theorie*. München: Wilhelm Fink.

Berschin, H. (1971). Sprachsystem und Sprachnorm bei spanischen lexikalischen Einheiten der Struktur KKVKV. *Linguistische Berichte,* 12, 39-46.

Bolinger, D. (1948). On defining the morpheme. *Word*, 4, 18-23.

Booij, G. E. (1977). *Dutch morphology. A study of word formation in generative grammar.* Lisse: de Ridder.

Corbin, D. (1987). *Morphologie dérivationelle et structuration du lexique* (2 vols.). [Derivational morphology and the structuring of vocabulary]. Tübingen: Niemeyer.

Cutler, A. (1981). Degrees of transparency in word formation. *Canadian Journal of Linguistics,* 26, 73-77.

Dalton-Puffer, C. (1996). *The French Influence on Middle English Morphology: A Corpus-based Study of Derivation*. Berlin/New York: Mouton de Gruyter.

Dalton-Puffer, C. & Plag, I. (2000). Categorywise, some compound-type morphemes seem to be rather suffix-like: On the status of *-ful*, *-type*, and *-wise* in Present Day English. *Folia Linguistica*, XXXIV, 225-244.

Evert, S. & Lüdeling, A. (2001). Measuring morphological productivity: Is automatic preprocessing sufficient? In P. Rayson, A. Wilson, T. McEnery, A. H. & S. Khoja (eds.) *Proceedings of the Corpus Linguistics 2001 conference* (pp. 167-175), Lancaster: University Centre for Computer Research on Language.

Gaeta, L. & Ricca, D. (2003). Frequency and productivity in Italian derivation: A comparison between corpus-based and lexico-graphical data. *Rivista di Linguistica* 15.1.

Haspelmath, M. (2002). *Understanding morphology*. London: Arnold.

Hay, J. (2000). *Causes and consequences of word structure*. Ph. D. thesis, Northwestern University.

Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics*, 39.4, 1041-1070.

Hay, J. & Baayen, H. (2002). Parsing and productivity. In G. Booij & J. van Marle (Eds). *Yearbook of morphology 2001.* (pp. 203-235). Dordrecht/Boston/London: Kluwer.

Hay, J. & Baayen, H. (in press). Phonotactics, parsing and productivity. *Rivista di Linguistica*.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, 39, 21–46.

McQueen, J. M. & Cutler, A. (1998). Morphology in word recognition'. In A. Spencer & A. Zwicky (Eds.) *The Handbook of Morphology*. (pp. 406-427). Oxford: Blackwell.

Plag, I. (1999). *Morphological productivity. Structural constraints in English derivation*. Berlin/New York: Mouton de Gruyter.

Plag, I. (2002). The role of selectional restrictions, phonotactics and parsing in constraining suffix ordering in English, in G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 2001*. (pp. 285-314). Dordrecht: Kluwer.

Plag, I. (2003). *Word-formation in English*. Cambridge: Cambridge University Press.

Plag, I., Dalton-Puffer, C. & Baayen, H. (1999). Morphological productivity across speech and writing. *English Language and Linguistics*, 3, 209-228.

Plank, F. (1981). *Morphologische (Ir-)Regularitäten: Aspekte der Wortstrukturtheorie.* Tübingen: Narr.

Rainer, F. (1988). Towards a theory of blocking. In G. Booij & J. van Marle (Eds). *Yearbook of morphology 1987.* (pp. 155-185). Dordrecht/Boston/London: Kluwer.

Riddle, E. (1985). A historical perspective on the productivity of the suffixes *–ness* and *–ity*. In J. Fisiak (Ed.). *Historical semantics, historical word-formation*. (pp. 435-461). Berlin: Mouton de Gruyter.

Rose, J. H. (1973). Principled limitations on productivity in denominal verbs. *Foundations of Language*, 10, 509-526.

Saffran, J. R., Newport, E. L. & Aslin, R. N. (1996a). Statistical learning by 8-month old infants. *Science*, 274, 1926-1928.

Saffran, J. R., Newport, E. L. & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606-621.

Schneider, K. P. (2003). *Diminutives in English*. Tübingen: Niemeyer.

Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar*. Cambridge: Cambridge University Press.

Skousen, R., Lonsdale D. & Parkinson, D. B. (Eds.) (2002). *Analogical modeling of language*. Amsterdam: Benjamins.

Notes

I would like to thank Maria Braun, Sabine Lappe, Mareile Schramm, two anonymous readers and April McMahon for helpful comments on earlier drafts of this paper.