Productivity and competition in morphological change

Stefan Hartmann, HHU Düsseldorf, hartmast@hhu.de
Kristian Berg, Universität Bonn, kberg@uni-bonn.de

The notion of productivity has a long tradition in morphology, but it is only quite recently that explicit corpus-based models have been proposed that allow a more precise quantification of this property of morphological patterns. We present an overview of both qualitative and quantitative approaches to modelling productivity, focusing on the question of how changes in productivity can be measured and interpreted. In addition, we address the potential link between productivity and inflectional morphology that has been postulated in recent work. The issue of productivity change is intimately tied to that of competition among patterns and words. The rise of one pattern is often linked to the demise of another. It is often argued that functionally equivalent patterns cannot coexist peacefully. Thus, competition between patterns often leads to substitution and differentiation processes that also have an impact on the productivity of the patterns involved.

1. Introduction

Morphological patterns differ in the extent to which they can give rise to new instances. Language users are aware of this, as the following examples from the webcorpus ENCOW16AX (Schäfer & Bildhauer 2012; Schäfer 2015) show:

(1)　　a.　　I meant un-impeachable. Is that a word? (https://bit.ly/3GPjUK8, ENCOW)
　　　　b.　　this has not been the outcome from all that ribboning. (Is that a word?)
　　　　　　　(https://bit.ly/3VdepcF, ENCOW)
　　　　c.　　Is unconfirmative a word? (https://bit.ly/3XCh1SY, ENCOW)

While the underlying rationale of these examples is the notion that a word either exists or does not exist, these doubtful cases (doubtful at least to the writers) illustrate that morphological patterns yield products that differ in their acceptability. It is also fairly self-evident that some morphological patterns yield new instances more readily than others. This is what is captured in linguistic morphology by the notion of *productivity*, "the statistically determinable readiness with which an element enters into new combinations" (Bolinger 1948: 18), to quote one of the oldest and least controversial definitions.[1]

---

[1] Whether this is a useful definition, and whether any given operationalization of this definition is useful is another question. This can ultimately be gauged by the usefulness of the concept within a linguistic theory or approach: Does the concept help us to describe the data adequately, economically, elegantly etc.?

The concept of morphological productivity plays a pivotal role in approaches to historical morphology. For instance, Scherer (2005, 2006, 2007) sees changes in productivity as a hallmark of word-formation change. However, there is quite some debate about how the concept should be defined and operationalized. In the late 1980s, Rainer (1987: 187) pointed out that the notion of "productivity" is used extensively in the literature on word-formation but rarely discussed explicitly. Over the last decades, however, this has arguably changed. A host of literature has emerged that discusses both the theoretical foundations and the empirical operationalization of the concept of morphological productivity (see e.g. Plag 1999; Bauer 2001, to mention only two monographs). From a historical perspective, productivity is related to another concept that has arguably been used in a rather pre-theoretical way until recently, namely competition (see e.g. Kempf 2016; Zehentner 2019). Productivity and competition are closely related in the sense that the rise of one pattern is often linked to the decline of another pattern – for example, Aronoff & Cho (2001), discussing the rival English affixes *-dom, -hood,* and *-ship*, argue that the latter has become specialized to distinguish between stage-level (i.e. properties that hold at a given time: *kingship, sponsorship*) and individual-level predicates (i.e. denoting stable and enduring properties: **parentship, *wifeship*; but see e.g. Trips 2009: 223, who argues that the relationship between the three suffixes cannot be characterized in terms of rivalry or competition).

In this chapter, we discuss the concept of morphological productivity and its relationship to competition from a diachronic perspective. We largely focus on examples from English and German derivational morphology, but most of the concepts can easily be applied to inflectional (or even syntactic) patterns, as well as to other languages. The chapter is structured as follows: In Section 2, we give a general overview of the concept of productivity. Section 3 focuses on the question of how productivity can be adequately measured. Section 4 discusses an alternative approach to productivity, namely Yang's tolerance principle. Section 5 delves into the question of how productivity and word-formation change are connected. Section 6 explores the relationship between productivity and competition. Section 7 very briefly discusses the impact of language-external, cultural factors on linguistic productivity, before Section 8 concludes the chapter.

2. Productivity

The term 'productivity' plays a key role in explaining the dynamics of language at different levels, but it means slightly different things in different contexts. Hockett (1960) famously lists productivity among his 13 "design features" that set human language apart from other communication systems, defining it as "the capacity to say things that have never been said or heard before and yet to be understood by other speakers of the language." In a similar vein, the term is used in the literature on child language acquisition, referring to a child's developing capability to go beyond the input they receive in their use of a language (see e.g. Rowland 2014: 104). In research on (derivational) morphology, the term *productivity* usually refers to the degree to which a pattern can be extended (see e.g. Booij 2012: 18). These different concepts of

productivity, which we will try to disentangle in more detail below, have in common that they conceive of productivity as a feature of linguistic patterns which are *used* by speakers or signers to produce instances of the pattern. For instance, when children form the past tense **goed*, they productively use the default English past tense construction. In the same way, they can produce grammatical past tense forms like *laughed, waited, googled* even without having heard them before. When investigating language acquisition, the analyst's focus is usually on the individual learner. Hence, the crucial question is to what extent one person makes use of a pattern. From the perspective of historical morphology, the global, community-wide perspective is usually more relevant. Here, the key question is whether and to what extent the language community can extend a pattern.

Importantly, productivity can be understood as a categorical or as a gradual phenomenon. For example, Bauer (2001: 49), following Corbin (1987), makes a distinction between availability (*disponibilité*) and profitability (*rentabilité*; the English terms have been proposed by Carstairs-McCarthy 1992). The former notion is a binary one (Plag 2021: 485), as it refers to the question of whether or not a morphological process can be used to produce new words. The latter one, by contrast, is more gradual and refers to "the possibility of being applied to a large number of bases and/or to produce a large number of attested derivatives" (Corbin 1987: 177, our translation[2]). In other words, profitability is "the *extent* to which a morphological process may be employed to create new pertinent forms" (Plag 2021: 485, emphasis added). For example, *-ness* and *-ity* can both be used to derive nouns in present-day English, but *-ness* is used more often to form neologisms and can thus be said to exhibit a higher degree of profitability (Bauer et al. 2013: 32). While much quantitative work on productivity draws on an explicitly gradual understanding of the concept, Gorman & Yang (2019: 174) argue that "[t]he child language acquisition evidence [...] unambiguously favors the categorical view" on productivity. A few considerations can potentially reconcile the binary and the gradual view on productivity. Firstly, and fairly obviously, the two views do not exclude each other, especially if we see productivity as a multidimensional phenomenon that can be broken down into different subconcepts such as Corbin's aforementioned availability and profitability. Secondly, the different views can be complementary when applied to different levels of analysis. For example, in inflectional morphology, which can be conceived of in terms of the filling of paradigmatic cells (see Aronoff 2019: 56), processes like, say, past tense formation for English verbs are fully productive, but specific inflectional patterns such as strong and weak inflection differ in their productivity (see e.g. Gorman & Yang 2019 and Yang 2016 on productivity in inflection). And thirdly, as pointed out above, it makes a difference whether we focus on productivity in individuals over the course of language acquisition, or on productivity on a community level, especially when taking a diachronic perspective.

What these considerations show, however, is that it makes sense to distinguish between several concepts of productivity. Barðdal (2008: 10–11), for example, distinguishes no less than

---

[2] French original: "la possibilité de s'appliquer à un grand nombre de bases et/ou de produire un grand nombre de dérivés attestés."

19 ways in which the term *productivity* can be used. Focusing on productivity in word-formation, Rainer (1987: 188–190) distinguishes six different concepts of word-formation productivity:

1. The productivity of a word-formation rule is the function of the quantity of the formations that have been realized on the basis of this rule at a given point in time.
2. The productivity of a word-formation rule is the function of the amount of its possible bases (e.g. Lieber 1981).
3. Productivity as the ratio of actual to possible formations (e.g. Aronoff 1976; Al & Booij 1981).
4. Productivity as the possibility of new formations (e.g. Schultink 1961).
5. Productivity as the probability of new formations (e.g. Harris 1951, Aronoff 1980).
6. Productivity as the number of new formations in a given time span (e.g. Neuhaus 1973).

It seems worthwhile to discuss those six concepts in more detail, as they all address aspects that have played an important role in the discussion on morphological productivity. The first definition is the simplest one, as it assumes a direct correlation between (synchronic) type frequency and productivity. But from quite early on, this approach has been regarded as too simple as it does not take the potential of a morphological pattern to yield new words beyond the existing ones into account, and neglects its diachronic dynamics (see e.g. Aronoff 1980: 72; Rainer 1987: 188). The second and third definition put the concept of "possible word" central stage, with the third definition resembling the index proposed by Aronoff (1976: 36), who suggested to operationalize morphological productivity as "the ratio of possible to actual words". The concept of "possible formations" itself, however, is difficult to operationalize: If, for example, we are interested in the productivity of *hood*-suffixation in English, which takes nouns and adjectives as bases (*neighborhood, falsehood*), can *all* nouns and adjectives be considered potential base lexemes for the pattern, and if not, how can the subset of nouns that are eligible for *hood*-derivation be properly described? For fairly obvious reasons, it is debatable what exactly counts as a possible word (see e.g. Bauer 2014). For instance, can blocked words like English **studier* (as in *carrier*, blocked by *student*) be considered possible words (see e.g. Gardani et al. 2019: 20f.; Fonteyn, Klein & Hartmann, this volume)? According to Plag (1999: 7),

> A possible, or potential, word can be defined as a word, existing or non-existing, whose morphological or phonological structure is in accordance with the rules of the language. It is obvious that before one can assign the status of 'possible word' to a given form, these rules need to be stated as clearly as possible.

Note, however, that such an understanding of possible words hinges upon the definition of "rule", as Bauer (2014: 97) points out: "the moment we move away from the standard notion of 'rule' with a defined input and output, and instead consider something that has an aleatoric component in the creation of the output, the notion of a possible word is changed". The concept

of "actual word", or "usual word", is no less problematic, as ever-larger corpora allow us to find increasing numbers of attested words, but not every word-formation product that is attested is necessarily accepted in a linguistic community (see Bauer 2014). Especially from the point of view of historical linguistics, where it is even harder to determine "possible" or "actual" words, the last three of the six definitions cited above are therefore much more feasible.

Note that the sixth definition takes an explicitly diachronic perspective, while the first three are explicitly synchronic, and the fourth and fifth are neither specifically synchronic nor specifically diachronic. The fourth and fifth definitions are close in spirit to Corbin's notions of availability and profitability, as well as to Bolinger's (1948) classic definition cited above. The sixth, diachronic, definition adds an important aspect that is crucial to any definition of productivity: perhaps somewhat paradoxically, it reminds us that productivity is an inherently synchronic notion. As the productivity of each pattern is subject to change, productivity can only be assessed for a specific point in time. At the same time, productivity is an inherently dynamic notion as it relies on the fact that new words come into being, which can only be assessed by comparing a word-formation pattern's output at different points in time (see e.g. Scherer 2015: 1784).

The six concepts of productivity mentioned by Rainer (1987) mainly focus on existing ("usual") and new formations, as well as on theoretically possible formations. In other words, they focus on types, not taking token frequency into account (also see Fonteyn, Klein & Hartmann, this volume). But token frequency has been hypothesized to have an impact on productivity. For instance, according to one hypothesis, forms with high token frequency "tend to be accessed whole, are not easily decomposed, and so do not contribute to the productivity of the affixes they contain" (Hay 2001: 1041). This assumption seems intuitively plausible, especially from a diachronic point of view and given that high-frequency words often undergo lexicalization, thus departing from other word-formation products not only in terms of frequency but also semantically (see e.g. Fonteyn 2019). Hay's (2001) experimental results, however, do not support a direct link between absolute token frequency and non-decomposability/non-transparency. Instead, relative frequency, i.e. the frequency of the derived form relative to its base, seems to be the more decisive factor. This contradicts Baayen & Lieber's (1991: 807) assumption that "the relation between the frequencies of base and derivative are not particularly relevant to the study of productivity in derivation." But again, those different perspectives can partly be reconciled if we conceive of productivity as a multidimensional phenomenon that emerges from multiple interacting factors such as "parsability, relative frequency, semantic and phonological transparency" (Plag 2021: 497).

Arndt-Lappe (2023) proposes yet another conceptualization of productivity, starting from the observation of individual differences between language users: in the Analogical Modeling approach she adopts, the productivity of an affix can be defined as

the probability of (a group of) stored complex words with that affix to serve as analogues for a novel form. This probability may be different for each individual novel form, because it

depends on how the Mental Lexicon is populated with stored complex words that are similar to that form. Productivity in that sense, then, is a lexical profile across different novel forms sharing an affix rather than a property of an affix that can be captured in terms of a global productivity value. (Arndt-Lappe 2023: 27)

Note that this view is highly compatible with the idea of conceptualizing productivity as a multidimensional concept, as hinted at above. It also has implications for measuring productivity, as it entails that measuring productivity globally for a given pattern may yield misleading results, and that it might, in many cases, make sense to assess productivity for specific subgroups instead.

The conceptualization of productivity also depends on the researcher's theoretical assumptions regarding the representation of language in the mind. O'Donnell (2016: 22) points out that most cognitive theories assume that a trade-off between storage and computation is at the heart of productivity, but there are two ways of how this trade-off is construed: some approaches see it as a trade-off between the memory space used by stored items on the one hand and the amount of computation necessary for language production and comprehension on the other, while others construe the trade-off "in terms of the ability to succinctly represent the input data and thereby accurately predict new forms" (O'Donnell 2016: 22).

The different aspects that can be seen as constitutive of productivity become even clearer if we take a closer look at the empirical operationalization of the concept, which is the topic to which we now turn.

3. Measuring productivity

If we conceive of productivity as a gradual concept, it stands to reason that linguistic patterns differ in their degree of productivity. This entails the challenge of measuring productivity in a quantitative way. A number of different ways of measuring productivity have been proposed. They are partly complementary in that they focus on different aspects, or dimensions, of productivity. In this section, we will first take a closer look at the widely-used measures of morphological productivity developed by Harald Baayen, and then discuss some approaches that have been proposed in the more recent literature.

3.1 "Classic" measures of productvity

Baayen (2009) chooses a vivid and intuitively plausible comparison for explaining both the theoretical background of approaches to productivity and the way these theoretical concepts are operationalized in a quantitative way: He compares morphological rules to companies, thus taking up the economic metaphor that permeates the terminology around morphological productivity (cf. terms like "profitability", as mentioned above). A company, or a morphological pattern, can be successful because it has a large share of the market. In this case, there are many extant word-formation products that are formed using this pattern. Thus, the pattern will have a high type frequency. But just because a company has a large share of the market doesn't

necessarily mean that it will remain successful. It is also conceivable that the market is saturated one day, and while the company's products are still widespread, it doesn't sell many new products. In this case, the pattern may have a high token frequency and maybe even a high type frequency, but it will not yield many new formations. Conversely, the company could also be a successful start-up whose products have only just begun to conquer the market. In this case, both the type and the token frequency of the pattern would be rather low, but we would be able to find many new coinages.

These ideas are operationalized in the widely-used measures of productivity developed by Baayen (e.g. 1992, 1993a, 1994; also see Baayen & Lieber 1991), summarized in Table 1 (following the overviews in Baayen 2009 and Hilpert 2013: 132). For each of the different measures, different terms have been proposed – for example, the measure called potential productivity in Table 1, which can arguably be considered the most important measure of productivity given its broad application in the literature (see e.g. Plag et al. 1999; Scherer 2006; 2007; Trips 2009; Hartmann 2016, 2018a, among many others), has been called "productivity in the narrow sense", "category-conditioned degree of productivity", "Baayen's $\mathcal{P}$", or simply P. All of the measures mentioned in Table 1 work with three different ways of counting frequencies: firstly, we can count the number of tokens, i.e. the running words in the text (n); secondly, the number of types, i.e. the number of different lexemes (v); and thirdly, the hapax legomena, or hapaxes for short, i.e. words occurring only once in the corpus (n_1).

Tab. 1: Established measures of morphological productivity.

Measure of productivity	Operationalization	Formula	Example
Realized productivity	Sum of all types of a construction (type frequency)	$P = \sum v_c$	Sum of all *hood*-nouns in the corpus
Type frequency change	Sum of all new types in a given time span	$P = \sum v_{c,new}$	Sum of all *hood*-nouns not attested in a (sub)corpus covering a previous time span

Normalized frequency change	Sum of all new types in a given time span, divided by corpus size	$$P = \frac{\sum v_{c,new}}{N}$$	Sum of all *hood*-nouns not attested in a (sub)corpus covering a previous time span, divided by the total number of tokens
Potential productivity	Sum of all hapax legomena of a pattern, divided by the number of all tokens belonging to the same pattern	$$P = \frac{\sum n_{c,1}}{N_c}$$	Sum of all *hood*-nouns attested exactly once in the corpus, divided by the total number of *hood*-nouns in the corpus
Expanding productivity	Sum of all hapax legomena belonging to the construction, divided by the sum total of hapax legomena in the corpus	$$P = \frac{\sum n_{c,1}}{\sum N_1}$$	Sum of all *hood*-nouns attested exactly once in the corpus, divided by the total number of tokens attested exactly once in the corpus

Fig. 1 illustrates the application of Baayen's measures of productivity using the example of the suffix -*hood* in the Corpus of Late Modern English Texts (CLMET, De Smet 2005, De Smet et al. n.d.). Note that the plots were generated using the lemmatization of the corpus without manual correction, which is often not sufficient as spelling variants or wrongly tagged items can distort the results (Evert & Lüdeling 2001) – for instance, the current dataset contains deviant spellings like *falshood* and *child-hood* (with a hyphen) as well as the place name *Robinhood*; but for expository purposes, the numbers obtained in this way can provide a good starting point for discussing the practical application of Baayen's productivity measures. Starting with realized productivity, the type frequency of -*hood* derivatives is 59 in the first time slice, 62 in the second, and 57 in the third. The raw type frequency is of course only informative if roughly equally-sized subcorpora are used. In the case of CLMET, the differences between the three time slices are not as grave as in other corpora, but still, the second time slice comprises more tokens than the first and the third more than the second (around 10, 11, and 12 million tokens, respectively). In the upper-left panel of Fig. 1, the realized productivity has therefore been

normalized by dividing the number of *hood*-types by the number of tokens in each time slice. (An alternative approach would have been to divide it by the number of types; the results would be very similar.) As Panel A in Fig. 1 shows, the measure of realized productivity suggests a very slight decrease in the productivity of *hood*-derivation. For example, *tutorhood, squirehood,* or *novicehood* are only attested in the first of the three time slices.

Panel B in Fig. 1 shows the development of the pattern's potential productivity, i.e. the ratio of hapax legomena in *-hood* to the token frequency of *hood*-derivatives. Here, we see a relatively steep decline that might indicate a decrease in the number of new formations but that may at least partly also be an artefact; we will return to this issue below when we discuss the limitations of potential productivity in more detail. The general idea, however, is that as we see a smaller share of hapaxes, we can assume that during this time slice, the pattern has yielded less neologisms than in the period before. The logic behind this is that the number of hapax legomena correlates with the number of new formations, even though in some cases, it might be a coincidence that a lexeme occurs only once in the corpus (see e.g. van Marle 1992 for a critical discussion of hapax-based measures). Turning to expanding productivity, panel C in Fig. 1 shows that the ratio of hapaxes in *-hood* to the total number of hapaxes decreases, which suggests that the pattern's contribution to the overall growth of the vocabulary is getting lower. Finally, panel D in Fig. 1 shows the so-called global productivity of the pattern. Global productivity is not a measure in its own right,[3] but rather an x-y plot of type frequency and potential productivity (see Baayen & Lieber 1991: 818; Bauer 2001: 154; Hilpert 2013: 132). Taking both realized and potential productivity into account at the same time reminds us that the different productivity measures should not be used or interpreted in isolation – as mentioned above, they cover different aspects of productivity. As Hilpert (2013: 132) puts it, "[g]lobal productivity thus takes into account that morphological categories may be productive in different ways, either in terms of many types, or in terms of a high ratio of hapaxes".

---

[3] However, Baayen & Lieber (1991: 836–838) also offer a refined definition of global productivity as the ratio of the number of types in the sample to (an estimate of) the number of types in the population.
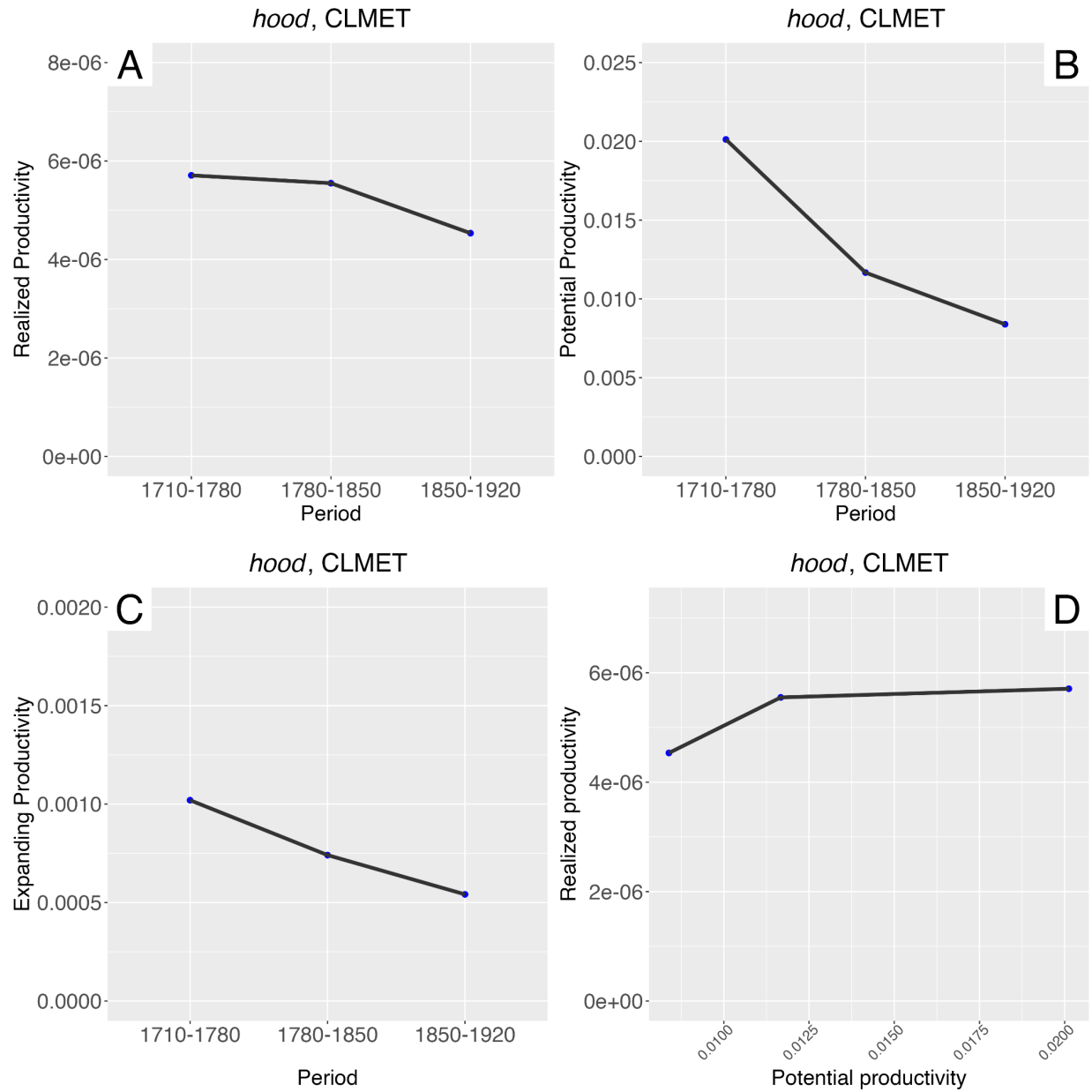
Fig. 1: An example for the application of Baayen's productivity measures. The four plots show the potential productivity (A), the realized productity, normalized by dividing the number of types by the overall number of tokens (B), the expanding productivity (C), and the global productivity (D) of *-hood* in CLMET.

3.2 Problems of potential productivity, and possible solutions

As mentioned above, potential productivity is probably the most widely-used measure of productivity. This has various reasons. Firstly, if we conceive of productivity as the degree to which a word-formation pattern yields new instances, potential productivity offers an easy-to-operationalize measure that gauges the amount of new formations from the number of hapax legomena. Secondly, it enables us to compare different time slices in diachronic corpora, while a neologism-based measure of productivity as proposed by e.g. Cowie (1999) or Berg (2021b) suffers from the disadvantage that the number of first attestations is almost necessarily skewed to the left as the size of the inventory of attested types increases cumulatively with each time slice. Hapax legomena, by contrast, are usually determined on the basis of the entire corpus, and their distribution across different time slices can give clues to changes in morphological productivity.

However, such a comparison only really makes sense if the subcorpora are of roughly equal size. This is intuitively plausible: The smaller our sample is, the higher the probability that a lexeme that is not a neologism occurs only once in the dataset. But not only different sample sizes can be problematic. Even if the subcorpora are equally-sized but there are large differences in the number of attestations of the pattern, this can distort the results (see e.g. Gaeta & Ricca 2006). This is because the number of hapaxes does not grow in a linear fashion with the number of tokens (see e.g. Säily 2011: 123; Kempf 2016: 117). Recall that the measure of potential productivity only takes tokens belonging to the pertinent pattern into account, as we divide the hapaxes belonging to the pattern by the total number of tokens belonging to the pattern. In general, potential productivity tends to overestimate the productivity of a pattern in the case of subcorpora with fewer tokens, while it tends to underestimate productivity in the case of subcorpora with many tokens.

A number of solutions have been proposed for this problem. Gaeta & Ricca (2006) propose what they call a variable-corpus approach, where different subcorpora are used in such a way that the number of tokens belonging to the word-formation pattern in question is held constant. A more widespread approach draws on cross-linguistically well-attested laws of word frequency distributions (Baayen 2001) which allow for extrapolating the vocabulary size (i.e., the number of types) and the number of hapaxes for different Ns (where N denotes the total number of running tokens in a given subcorpus). Finite Zipf-Mandelbrot models are often used as an extrapolation method (e.g. Kempf 2016, Schneider-Wiejowski 2011). These models make use of the fact that word frequency distributions have been found to be strikingly similar to the distribution described by Zipf's law (see e.g. Evert 2004). According to Zipf's law, the frequency of words is inversely proportional to their rank order (see e.g. Baayen 2001: 13–16). Thus, the most frequently attested word in a corpus tends to be twice as frequent as the second-most frequent one, three times as frequent as the third-most frequent one, and so on. The Zipf-Mandelbrot law is a generalization of Zipf's law and can be formalized as in (1) (Zipf's law is a special case of (1) with $b = 0$).

(1) $\pi_z = \dfrac{C}{(z+b)^a}$

In (1), π_z denotes the population probility of a word with the Zipf rank z. C is a normalizing constant that ensures that the probabilities sum up to 1 (see Baayen 1993b: 348; Baayen 2001: 14f.). The distribution parameters a and b modify the slope of the rank-frequency graph (a) and introduce a downward curvature that is particularly useful for describing high-frequency ranks (b), respectively (see Baayen 2001: 101f.). Evert (2004) re-formulates (1) in terms of a type density function. This allows for calculating expected values of V (i.e., the number of different types) and V_m (i.e., types occurring m times) for a specific sample size N. Fig. 2 shows the results of such an approach for -*hood* in the CLMET corpus. The results were obtained using Evert & Baroni's (2007) R package *zipfR*. A finite Zipf-Mandelbrot model was fit to each of the three periods. The left panel in Fig. 2 shows the vocabulary growth curve, i.e. the number of -*hood* types at a given corpus size, while the right panel shows the extrapolated number of hapaxes (which can in turn be used to gauge the potential productivity for an arbitrary corpus size).
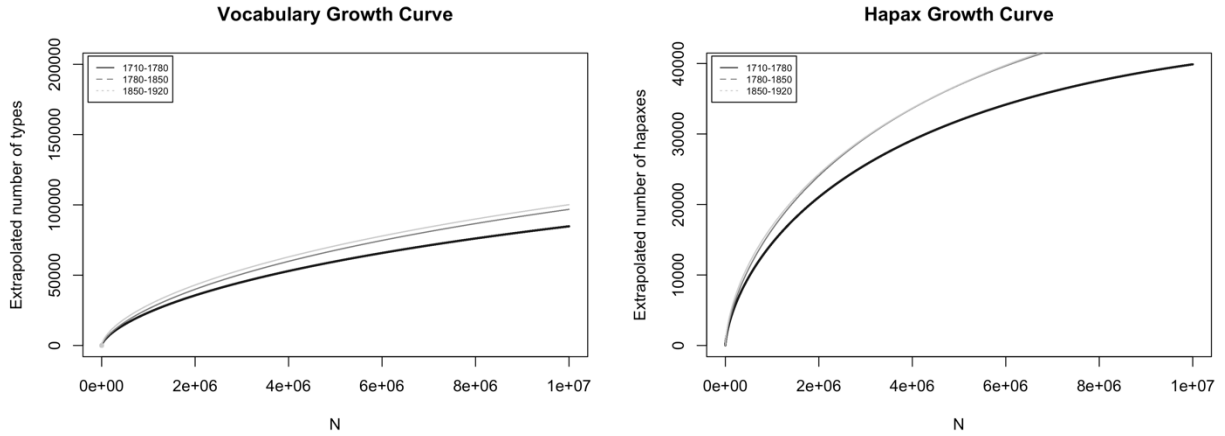

Fig. 2: Vocabulary growth curve (left) and hapax growth curve (right) for -*hood* in CLMET, extrapolated with the help of a finite Zipf-Mandelbrot model.

Note that the results differ from the ones in Fig. 1 that indicate a decline in productivity; in Fig. 2, this is not as clear. The pattern is consistently less productive in the last time span when compared to the time span before – but it is least productive in the first time span according to the fZM model. However, these results must not be overinterpreted: The goodness-of-fit of all three models, assessed using multivariate chi-squared tests that compare the extrapolated values to the observed ones, is decidedly bad, with very high chi-squared values and very low p-values – the opposite would indicate a good fit of the model to the data (Baayen 2001: 118–122). One reason for this is probably the relatively low sample size, which is a pertinent problem in virtually all areas of historical linguistics. While we cannot fully overcome the fuzziness that low sample sizes introduce into our quantitative assessments, we can make it transparent, e.g. by

using bootstrapping approaches. In bootstrapping approaches, corpus data are randomly reshuffled, thus creating many different virtual corpora that the model encounters. Especially in cases where the order in which individual items enter the model is relevant (as is the case for fZM models), this can help get a clearer impression of the variation in the data. Hartmann (2018a), for example, uses a bootstrapping approach to gauge differences between fZM models fit to data of nominalizations with the suffix *-ung* from earlier stages of New High German.

Säily (2016) uses a different approach to overcome the problem of the non-linear relationship between types and tokens: A corpus is divided into samples "large enough to preserve discourse structure" (Säily 2016: 132), e.g. samples the size of individual texts; then, an algorithm picks the samples in a random order to construct a type accumulation curve for the pattern in question. This is then repeated a large number of times (1 million times in Säily 2016). This bootstrapping approach allows for gauging upper and lower boundaries of the type accumulation curve. According to Säily (2016: 131), an important advantage of this method is that it "does not make simplifying assumptions like 'words occur randomly in texts'." Figure 3 shows the results of such an approach for the aforementioned *-hood* data from CLMET (but with only 1000 iterations). On average, the results suggest a decrease in type frequency, as did the simple measure of realized productivity in Fig. 1 above. However, the results also show that there is much variation. Unsurprisingly, the degree of variation is inversely correlated to the sample size – the larger the sample, the less variation.
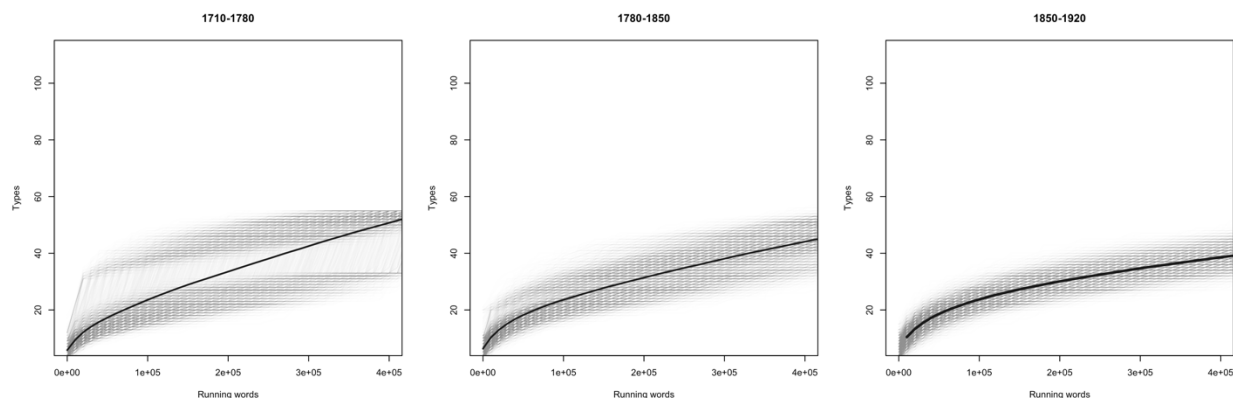

Fig. 3: Results of a bootstrapping approach à la Säily (2016). The thick lines in the middle show the mean across all growth curves.

While Säily (2016) uses type frequency as an indicator of productivity, the same method could of course also be applied for determining e.g. the number of hapax legomena. Berg (2021a) criticizes the choice of type frequency as a measure of productivity, as "with this measure, we can only determine whether a pattern is productive at all, not whether the level of productivity changes over time" (Berg 2021a: 178). Quantifying productivity in a reliable way arguably always requires the combination of different measures (Kempf 2021: 47).

One advantage of working with historical data is that the number of actual new words in a corpus in any given time span can easily be determined. For synchronic data, this is much harder – this observation led to the development of hapax-based measures (as indicators for new words) in the first place. The absolute number of new words for a given pattern in a given time span can be used as one dimension of the pattern's productivity (V_{neo} in Berg 2020). For example, in CLMET, in the second period (1780–1850), there are 35 new *-hood* nouns (i.e. *-hood* nouns not attested in the first period from 1710 to 1780), but no less than 582 new *-ness* nouns. Another dimension is the ratio of new words for a given pattern among all words of that pattern in a time span (R_1 in Bolozky 1999, P_{neo} in Berg 2020). This measure captures the renewal rate of a pattern. If patterns A and B yield the same absolute number of new words (V_{neo}), but pattern B has twice as many established formations compared to A, then its P_{neo} value is only half that of pattern A; pattern A has a higher ratio of new formations, a higher degree of regeneration. One challenge these measures face is their dependence on corpus size. For the results to be comparable, each time span should contain the same amount of material. That of course is wishful thinking for most diachronic corpora; texts from earlier periods are notoriously sparse. A workaround is a) to discard periods with too little material and b) to use Säily's (2016) approach of re-sampling the corpus (cf. Berg 2020 for an application). One advantage of using new words is that the lists of new words for each time span can also be used for qualitative analyses.

Similar to Cowie (1999) and Berg (e.g. 2021a,b), Kempf (2016) and Kempf & Hartmann (2018) use a neologism-based measure of productivity alongside other measures (such as Baayen's realized and potential productivity). But instead of gauging the emergence of a new word from its first attestation in the corpus, they opt for a manual approach in which for each single lexeme, several sources (e.g. a corpus, GoogleBooks, and an etymological dictionary) are checked, and the oldest attestation is used as a proxy for determining when the word-formation product must have been coined. This method is of course only well-suited for relatively small samples as it requires painstaking manual work. In addition, it might also be affected by some of the problems that it tries to solve: As mentioned above, neologism-based measures that determine neologisms corpus-internally have the disadvantage that they are skewed towards earlier periods as the amount of text taken into account grows with each period. However, a case could be made that this does not only affect individual samples (e.g., corpora), but also the full population of extant texts from earlier periods of most languages that have a written tradition. The farther we go back in time, the less texts we have available – which could potentially mean that we tend to underestimate the age especially of mid- and low-frequency words even if we checked all available sources.

4. An alternative approach: Yang's tolerance principle

A fundamentally different approach to inflectional and derivational productivity is Yang (2016). He focuses on child language acquisition, and consequently, his way of assessing productivity is

more synchronically and more psycholinguistically oriented. The key insights of Yang's work are a) that learners' input can be used to derive predictions about their output and b) that for rules to be "discovered", exceptions must not exceed a certain threshold. In this vein, Yang analyzes child-directed speech, and he proposes a "tolerance principle": The number of exceptions for a putative rule is always below the total number of cases divided by the natural logarithm of the number of cases. For example, there are 1022 past tense forms in Yang's corpus of child-directed speech. 995 of them consist of adding -ed to the stem form (e.g. *played*, *picked*, *fitted*), while 127 involved some other form of (non-)modification (e.g. *brought*, *wrote*, *put*). The tolerance principle predicts that this rule is learnable because the actual number of exceptions is smaller than the maximum number of exceptions that learners can tolerate (147 exceptions ≈ 1022/ln1022), and this is precisely what happens.

It is maybe no coincidence that the most frequently-cited examples for the tolerance principle stem from inflectional morphology. To cover derivational morphology, the tolerance principle must be slightly altered. For example, Yang (2016: 108ff.) suggests semantic plausibility as a criterion for exceptions to a rule. The pattern -*ness* is attested 23 times in Yang's corpus of child-directed speech, and 21 of these cases are semantically regular in that they nominalize adjectives and denote the abstract quality (e.g. *coolness*, *sadness*). Two -*ness* words are exceptional in this regard, the interjection *goodness* and the noun *witness*; both are not related to the base in the same way. But the pattern -*ness* with 23 cases can easily accommodate for these two exceptions; the tolerance principle allows for a total of eight exceptions. The pattern -*ness*, it follows, is used productively because it is sufficiently regular semantically for this regularity to be uncovered by language learners. In a similar vein, we could take a closer look at the -*hood* data from CLMET discussed above. Depending on how we define the semantic scope of -*hood*, the only true exceptions among the 126 *hood*-types are *childhood* and *boyhood* as well as *neighborhood*, as they have temporal and locative meanings, respectively, rather than referring to a state. However, this also shows an inherent challenge of operationalizing this approach: It requires us to make a decision as to what plausibly counts as an exception. Does *falsehood* count as an exception as it refers to a more or less reified piece of false information, rather than a 'state of being false', or can it still be viewed as a regular coinage? Given the semantic heterogeneity of word-formation products, especially those instantiating highly productive patterns (a well-known case being agentive -*er* in various languages, see e.g. Meibauer 1995, Panther & Thornburg 2001), it can be problematic to draw a clear boundary between 'regular' and 'exceptional' cases.

5. Productivity and word-formation change

The previous sections have shown that productivity has often been conceived of as an inherently diachronic concept, and/or has been operationalized in a diachronic way by comparing productivity values across time slices using quantitative measures of productivity. This raises the question of how exactly productivity is related to morphological change. As large parts of the

literature on operationalizing and measuring productivity focus on derivational morphology, we will discuss the diachronic change of word-formation patterns in this section, even though changes in productivity also affect inflectional patterns (see e.g. Dressler 2003, Gardani 2013).

According to Scherer (2007: 258), the productivity of a word-formation pattern depends on its formal and semantic constraints. As such, changes in morphological productivity reflect changes in the restrictions that word-formation patterns are subject to. Scherer's notion of restrictions (or constraints) is rather broad, comprising both input and output restrictions. Input restrictions can pertain to the word class of potential bases or to their (formal and semantic) features; output restrictions, on the other hand, concern the derivatives themselves. For instance, Scherer (2005, 2006, 2007) shows that in the case of German agentive nouns in *-er* like *Bäcker* 'baker', the output of the pattern has changed in that the object reading (as in e.g. *Wasserkocher* 'water heater') has become more frequent over time. However, it could also be argued that it might be misleading to treat the preferences that a pattern shows with regard to its input and the output it produces on a par and to frame both as restrictions or constraints. Especially in the case of a pattern's output, which is also strongly influenced by lexicalization, it might make more sense to speak of the output potential of a pattern instead (see e.g. Hartmann 2018b: 155).

6. The relationship between productivity and competition

As mentioned by Berg (2021a: 178), many of the measures of productivity surveyed above have been developed for synchronic investigation, especially with an eye on competition between different word-formation patterns. According to Bauer et al. (2013: 33), "[t]wo processes compete when they both have the potential to be used in the coining of new synonymous forms from the same base". While this definition is more synchronically-oriented, competition also plays a role from a diachronic perspective – for instance, in inflectional morphology, strong and weak inflection patterns compete with each other in many Germanic languages (see Fonteyn, Klein & Hartmann, this volume). And in word-formation, we also find a number of so-called alternations (see Pijpops 2020 for a discussion of this concept), i.e. "highly systematic correspondences between functionally near-identical patterns" (Zehentner 2019: 234). In English, for example, the nominal suffixes *-hood, -dom* and *-ship* fulfill very similar functions and can therefore be seen as "rivals in word-formation processes" (Trips 2009: 165). A similar case is the set of German adjectival suffixes *-mäßig* and *-haft*, which both denote a similarity to certain qualities of the base (e.g., *Bud-Spencer-mäßig, Bud-Spencer-haft* 'like Bud Spencer'). Scherer (2015: 1785) argues that "[t]he most important language-internal factor causing change in productivity is change in competition". Consequently, over historical time, we can often observe that the rise of one pattern is linked to the demise of another. In inflectional morphology, for instance, the expansion of the English gerund *-ing* can be seen as the reason why the present participle suffix *-end* fell out of use (Aronoff 2019: 51). In word-formation, similar processes can sometimes be observed, even though they tend to be less categorical: For instance, in German, infinitive conversion seems to fulfill functions that used to be fulfilled by nominalization with

the suffix *-ung* (Demske 2000, Hartmann 2016): Many *-ung*-nominals with a processual reading that are attested in earlier stages of German (such as *Lesung* 'reading' referring to the process of reading, rather than an event) are no longer possible in present-day German, and would instead be expressed by means of a nominalized infinitive (infinitive *lesen* 'to read' > nominalized infinitive *das Lesen* 'the (process of) reading').

In recent years, a number of competition models have been suggested that draw on evolutionary theory, either by borrowing concepts from biological evolution, as in Croft's (2000) approach to language change, or by explicitly adopting a generalized theory of evolution in the spirit of e.g. Richerson & Boyd (2005), as advocated by e.g. Aronoff (2019). In evolutionary models, competition plays a crucial role, along with variation and selection. Zehentner (2019), for instance, develops a model of competition that is couched in a Construction Grammar framework. She sees constructions, i.e. form-meaning pairings at various levels of abstraction, as units of replication (replicators). Constructions change on a micro-level as language users introduce variation, which can lead to competition. Importantly, newly emerged constructions can also enter competition with patterns that were previously unrelated, especially if two constructions overlap in meaning (Zehentner 2019: 301). In constructionist approaches to morphology (e.g. Booij 2010), inflection and word-formation patterns are viewed as constructions as well. Constructions in turn are organized in a network, the "construct-i-con" (Goldberg 2019). Thus, Zehentner's model can also be applied to word-formation patterns. Functionally similar patterns can be conceived of as being connected via "synonymy links" (Goldberg 1995: 91), and they can be seen as "allostructions" (Cappelle 2006). Once a new competitor enters the network, the links to other constructions can gradually become stronger or weaker, depending on various factors like the contexts in which they occur, their usage frequency, and potentially also the degree to which they "stand out" in comparison to other constructions, e.g. via the use of "extravagant" formal means such as repetition, unusual phoneme combinations, or (apparent) violation of grammatical rules (for examples of such "extravagant morphology", see Eitelmann & Haumann [eds.] 2022).

A key assumption of Construction Grammar is that no two constructions cover exactly the same functional domain – this is also called the "principle of no synonymy" (Goldberg 1995, but see Uhrig 2015 for critical remarks, and Leclercq & Morin forthc. for a response).[4] Especially if we adopt a broad notion of meaning, as is common in Construction Grammar (see e.g. Croft 2001), this makes intuitive sense – as it is highly unlikely that two patterns are used in exactly the same contexts, they will almost inevitably be subject to functional differentiation. De Smet et al. (2018) distinguish a number of ways how competing forms change their functions: in the case of substitution, only one form survives – for instance, they argue that *-ing*-clauses may currently be in the process of substituting *to*-infinitives after *begin* in American English (*begin to work* > *begin working;* De Smet et al. 2018: 207). Differentiation, by contrast, refers to the

---

[4] This does not mean that Construction Grammar rejects synonymy per se; Goldberg (1995) actually assumes synonymy links between constructions that fulfill similar functions. But she assumes that quasi-synonymous constructions differ in at least some semantic, pragmatic, and/or distributional aspects.

phenomenon of competing constructions coming to occupy different functional niches. For instance, they show that [*start* + *-ing*-clause] and [*start* + *to*-infinitive] hardly combined with non-agentive subjects until the 1990s in American English; then, however, both constructions opened up for non-agentive subjects, but this trend is much more pronounced for [*start* + *to*-infinitive]. Finally, attraction refers to the phenomenon of two constructions becoming functionally more, rather than less, alike. Again, [*begin* + *-ing*-clause] and [*begin* + *to*-infinitive] can be seen as an example of this, according to De Smet et al. (2018: 211): While a purely frequency-based comparison suggests a relation of substitution, a closer inspection of *how* each form is used also suggests that the two constructions grow more similar – while [*become* + *-ing*-clause] becomes more permissive of non-agentive subjects, [*begin* + *to*-infinitive] starts to combine with agentive subjects more frequently. Note that the rise of one variant co-occurs with the demise of another one; there currently is no method to determine the causality involved (did the demise of one pattern shift users to the other? Or did the rise of one pattern lead to the demise of the other?).

While De Smet et al.'s (2018) examples can be placed at the morphology-syntax interface (and also show that morphological change virtually always interacts with changes in other domains), a textbook example of competition between word-formation patterns is discussed by Kempf (2016), who also takes a Construction Grammar approach. She investigates the development of German adjectival word-formation patterns, especially those with the suffixes *-lich, -ig*, and *-isch*, but also taking other, competing suffixes like *-bar* and *-sam* into account, showing that they partly cover the same functional territory. Her results suggest that there are push- and pull-chain effects especially between *-bar* and *-lich*, with the former becoming the default suffix for expressing the passive-like semantics 'can be X-ed' and the latter becoming gradually less frequent as it is ousted by *-bar*. For *-lich, -ig*, and *-isch*, she shows that *-ig* becomes dominant for a variety of reasons: firstly, it is used more frequently in "possessive-ornative" function (as in *kurzhaarig* 'short-haired'). Secondly, it shows a high degree of affinity to complex derivation, which is even more important given the increased tendency towards compression of complex contexts in single word-formation products. With *-ig*, it is possible to derive adjectives from syntactic constructions: the base *vier Türen* 'four doors' can be turned into the adjective *viertürig* 'having four doors'. And thirdly, it takes over the functional domain of *-icht* (*sumpficht* → *sumpfig* 'swampy'), which largely falls out of use (see Kempf 2016: 297). While the three suffixes *-lich, -ig*, and *-isch* seem to occupy complementary niches, they do influence each other, e.g. via blocking patterns (*rechtlich*, but **rechtisch* 'legal', with a native base; *juristisch*, but **juristlich* 'legal', with a non-native base; see Kempf 2016: 300). Taking up De Smet et al.'s (2018) typology, the functional development of German adjectival suffixes as sketched by Kempf (2016) can partly be characterized as attraction, partly as differentiation – for instance, *-bar* becomes more similar to *-lich* by taking over its "passive" function, leading to doublets (*veränderlich/veränderbar* 'changeable'), but the subsequent decline of *-lich* in this passive function (see Kempf 2016: 191) can be seen as a case of differentiation. This also shows

that processes like substitution, attraction, and differentiation cannot be distinguished in a clear-cut way – instead, they are often closely interwoven.

As the aforementioned example of German *rechtisch* and *juristlich* shows, which have been argued to be blocked by *rechtlich* and *juristisch*, respectively, the concepts of competition and blocking are closely connected. Aronoff (2023) compares three ways in which blocking and morphological rivalry can be looked at: Firstly, Yang's Tolerance Principle postulates a clear maximum to the exceptions that a putative rule may have if it is to be productive (see above). Secondly, the Elsewhere Principle that can be traced back to Pāṇini's grammar of Sanskrit. According to this principle, a more specific rule takes precedence over a more general one – for instance, *-t* blocks *-ed* in the case of *bent* (Rainer 2012: 169). Thirdly, Gause's principle of competitive exclusion from the domain of evolutionary biology posits that no two species can coexist in a stable equilibrium. Applied to affix rivalry, this entails that either one of two competing affixes gets ousted or the two affixes come to occupy distinct niches. Aronoff (2023: 59) argues that competitive exclusion is the overarching property that governs both blocking and elsewhere distribution, and that Gause's principle can fruitfully be combined with the Tolerance Principle, which presumes competitive exclusion and explains individual gaps in otherweise general patterns. On a more controversial note, he also argues that the notion of blocking and the Elsewhere Principle are observationally useful but have no explanatory value.

Note that the evolutionary approach suggested by Aronoff is highly compatible not only with the Construction Grammar view sketched above but also with approaches that construe morphological rivalry as a gradient, rather than categorical, phenomenon. For instance, Guzmán Naranjo & Bonami (2023) argue that while most literature makes a rather categorical distinction between rival and non-rival processes, there are situations in which the meanings of two word-formation processes can overlap without being identical. French *-eur*, for example, derives agents or instruments, while *-oir* derives instruments or locations (Guzmán Naranjo & Bonami 2023: 88). While there seem to be no attested doublets, they argue with Roché (1997) that the French term for 'truck driver', *camionneur*, uses *-eur* for phonological reasons, while the expected suffix from a semantic point of view would have been the denominal agent suffix *-ier*, which shows the degree of rivalry between the two suffixes is low but not zero (Guzmán Naranjo & Bonami 2023: 90). This gradualistic view is very much in line with the approaches from usage-based linguistics and Construction Grammar mentioned above, which tend to emphasize that many aspects of language are gradual rather than categorical.

Summing up, then, there is an intricate relationship between productivity and competition, as the productivity of morphological patterns partly depends on the productivity of its competitors. Given the complexity of the relationships that exist between different patterns, many questions regarding the details of the interaction between morphological "rivalry" and productivity, and regarding the role of competition patterns in morphological change, have not been exhaustively answered yet, which is why this topic will certainly remain an important topic in research on morphology and morphological change in the coming years.

7. Cultural factors determining productivity

An evolutionary approach like the one defended by Aronoff, and inspired by Gause, comes with the additional advantage that it allows for explicitly taking cultural factors into account. Apart from the intra-linguistic factors mentioned above, cultural factors can also determine the rise or the demise of morphological patterns. For example, scientific or technical developments may produce an need for new words – and if the words are not needed anymore, the pattern becomes unproductive. A case in point is the English nominal suffix *-age*. During feudalism, this suffix was very popular for coining words that denote taxes and related notions. With the end of feudalism, however, it fell out of use in this sense (see e.g. Fleischman 1976).

Prestige and fashion arguably also play a role for the tides of morphological patterns. Certain patterns can come into fashion and fall out of it rather quickly, e.g. non-medical *-itis* in German (cf. Evert & Lüdeling 2001, e.g. *Telefonitis* 'excessively using the telephone'). A writer's age may play a role, too (as is often the case in matters of fashion): Berg (2022) suggests that the German nominal suffix *-nis* became unproductive in the 20th century because it became associated with older writers.

8. Conclusion

In this chapter, we have given an overview of the current state-of-the-art of research on morphological productivity, focusing on its relevance for morphological change. We have shown that especially from a historical perspective, there is a close link between productivity and competition: Over time, newly emerging patterns can gradually supersede older competitors; processes like substitution, differentiation, and attraction can influence the degree of productivity of competing patterns. Productivity is thus a highly dynamic notion and as such a key concept for historical morphology. Measuring productivity, however, can be a challenge, and while there is still some debate regarding the question of how productivity should be measured, there is arguably a growing consensus that productivity should be treated as a multidimensional concept (see e.g. Barðdal et al. forthc.). As such, a single measure of productivity cannot capture all relevant aspects. Instead, multiple measures should be combined to get a coherent picture of a pattern's synchronic productivity, as well as of the development of its productivity over time. Especially in the domain of word-formation, the diachronic development of morphological productivity is arguably the most important indicator of morphological change. According to Scherer (e.g. 2006), changes in word-formation productivity reflect changes regarding the constraints individual word-formation patterns are subject to.

While diachronic changes in productivity have been investigated in quite some detail recently, taking previously neglected aspects (e.g. sociolinguistic ones, see Säily 2016) into account, some questions remain open. While there are valuable models that account for the relationship between productivity and competition, the exact ways in which these concepts are interconnected should be explored in more detail in future research. Also, causal factors that determine changes in the productivity of individual patterns should be teased apart in more detail, taking complex interactions between intra- and extra-linguistic factors into account (see

e.g. Scherer 2015: 1785). This could also help us to tease apart aspects of productivity that can be considered universal from potential language-specific tendencies, which in turn could help us assess which observations made on the basis of well-documented languages can be generalized to smaller, less well-documented ones. After all, the methodological approaches discussed above require a considerable amount of data, and cannot be easily applied to underdocumented languages or so-called corpus languages like Latin (see Panagl 1982 for some considerations).

Supplementary material

The datasets and scripts used for the example analysis in the present chapter are available at https://osf.io/zqp6k/.

Acknowledgments

References

Al, Bernard & Geert E. Booij. 1981. De produktiviteit van woordvormings-regels: Enige kwantitatieve verkenningen op het gebied van de nomina actionis. *Forum der Letteren* 22. 26–38.

Arndt-Lappe, Sabine. 2023. Different lexicons make different rivals. *Word Structure* 16(1). 24–48.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge: MIT Press.

Aronoff, Mark. 1980. The relevance of productivity in a synchronic description of word formation. In Jacek Fisiak (ed.), *Historical morphology* (Trends in Linguistics 17). The Hague, New York: Mouton.

Aronoff, Mark. 2019. Competitors and alternants in linguistic morphology. In Franz Rainer, Francesco Gardani, Wolfgang U. Dressler & Hans Christian Luschützky (eds.), *Competition in Inflection and Word-Formation*, 39–66. Cham: Springer International Publishing.

Aronoff, Mark. 2023. Three ways of looking at morphological rivalry. *Word Structure* 16(1). 48–61. https://doi.org/10.3366/word.2023.0220.

Aronoff, Mark & Sungeun Cho. 2001. The semantics of -ship suffixation. *Linguistic Inquiry* 32(1). 167–173.

Baayen, Harald. 1992. Quantitative Aspects of Morphological Productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1991*, 109–149. Dordrecht: Kluwer.

Baayen, R. Harald. 1993a. On Frequency, Transparency, and Productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1992*, 181–208. Dordrecht: Kluwer.

Baayen, Harald. 1993b. Statistical models for word frequency distributions: A linguistic evaluation. Computers and the Humanities 26. 347–363.

Baayen, R. Harald. 1994. Productivity in Language Production. *Language and Cognitive Processes* 9(3). 447–469.

Baayen, R. Harald. 2001. *Word Frequency Distributions* (Text, Speech and Language Technology). Dordrecht: Kluwer.

Baayen, R. Harald. 2009. Corpus Linguistics in Morphology: Morphological Productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics* (HSK 29.2), 899–919. Berlin, New York: De Gruyter.

Baayen, Harald & Rochelle Lieber. 1991. Productivity and English Derivation: A Corpus-Based Study. *Linguistics* 29. 801–843.

Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic* (Constructional Approaches to Language 8). Amsterdam/Philadelphia: John Benjamins.

Barðdal, Jóhanna, Renata Enghels, Quentin Feltgen, Sven Van Hulle & Peter Lauwers. forthc. Productivity in diachrony. In *The Wiley-Blackwell Companion to Diachronic Linguistics.*

Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2014. Grammaticality, acceptability, possible words and large corpora. *Morphology* 24(2). 83–103. https://doi.org/10.1007/s11525-014-9234-z.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.

Berg, Kristian. 2020. Changes in the productivity of word-formation patterns: Some methodological remarks. *Linguistics* 58.4. 1117–1150.

Berg, Kristian. 2021a. Productivity, vocabulary size, and new words. A response to Säily (2016). *Corpus Linguistics and Linguistic Theory* 17(1). 177–187. https://doi.org/10.1515/cllt-2017-0075.

Berg, Kristian. 2021b. Produktivität in der historischen Wortbildung: Neubelege als Produktivitätsmaß. *Germanistische Linguistik* 252–254. 53–79.

Berg, Kristian. 2022. Produktivität, Lexikalisierung und Rückbau. *-nis* im Neuhochdeutschen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 144(4): 545–572.

Bolinger, Dwight. 1948. On defining the morpheme. *Word* 4. 18–23.

Bolozky, Shmuel. 1999. Measuring productivity in word formation: The case of Israeli hebrew. Leiden: Brill.

Booij, Geert E. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Booij, Geert E. 2012. The Grammar of Words: An Introduction to Linguistic Morphology. 3rd ed. Oxford: Oxford University Press.

Cappelle, Bert. 2006. Particle placement and the case for "allostructions." *Constructions* 7(SI1). doi 10.24338/cons-381.

Carstairs-McCarthy, Andrew. 1992. *Current morphology*. London, New York: Routledge.

Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique* (Linguistische Arbeiten 193). Tübingen: Niemeyer.

Cowie, Claire. 1999. *Diachronic word-formation: a corpus-based study of derived nominalizations in the history of English (PhD Dissertation, University of Cambridge)*. Cambridge. https://www.repository.cam.ac.uk/handle/1810/251674. doi https://doi.org/10.17863/CAM.16202

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow: Pearson Longman.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Demske, Ulrike. 2000. Zur Geschichte der ung-Nominalisierung im Deutschen: Ein Wandel morphologischer Produktivität. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 122. 365–411. https://doi.org/10.1515/bgsl.2000.122.3.365.

De Smet, Hendrik. 2005. A corpus of Late Modern English texts. *ICAME Journal* 29. 69–82.

De Smet, Hendrik, Frauke D'hoedt, Lauren Fonteyn & Kristel Van Goethem. 2018. The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics* 29(2). 197–234.

De Smet, Hendrik, Susanne Flach, Hans-Jürgen Diller & Jukka Tyrkkö. n.d. *The Corpus of Late Modern English Texts, version 3.1*. https://fedora.clarin-d.uni-saarland.de/clmet/clmet.html. (last checked 05/05/2022)

Dressler, Wolfgang U. (2003). Degrees of grammatical productivity in inflectional morphology. Italian Journal of Linguistics, Vol. 15, No. 1, pp. 31 — 62.

Eitelmann, Matthias & Dagmar Haumann (eds.). 2022. *Extravagant morphology*. Amsterdam, Philadelphia: John Benjamins.

Evert, Stefan. 2004. A simple LNRE model for random character sequences. In *Proceedings of JADT 2004*, 411–422.

Evert, Stefan & Marco Baroni. 2007. zipfR: Word frequency distributions in R. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Posters and Demonstrations Sessions, Prague*, 29–32.

Evert, Stefan & Anke Lüdeling. 2001. Measuring morphological productivity. In Paul Rayson, Andrew Wilson, Tony McEnery, Andrew Hardie & Shereen Khoja (eds.), *Corpus linguistics 2001*, 167–175. http://dx.doi.org/10.18452/13417.

Fleischman, Suzanne. 1976. *Cultural and linguistic factors in word formation: an integrated approach to the development of the suffix -age* (University of California Publications in Linguistics ; v. 86). Berkeley: University of California Press.

Fonteyn, Lauren. 2019. *Categoriality in language change: the case of the English gerund* (Oxford Studies in the History of English). New York, NY: Oxford University Press.

Gaeta, Livio & Davide Ricca. 2006. Productivity in Italian word-formation. *Linguistics* 44(1). 57–89.

Gardani, Francesco. 2013. *Dynamics of morphological productivity: the evolution of noun classes from Latin to Italian* (Empirical Approaches to Linguistic Theory volume 4). Leiden, Boston: Brill.

Gardani, Francesco, Franz Rainer & Hans Christian Luschützky. 2019. Competition in morphology: A historical outline. In Franz Rainer, Francesco Gardani, Wolfgang U. Dressler & Hans Christian Luschützky (eds.), *Competition in Inflection and Word-Formation*, 1–36. Cham: Springer International Publishing.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, London: The University of Chicago Press.

Goldberg, Adele. 2019. *Explain me this: creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press.

Gorman, Kyle & Charles Yang. 2019. When Nobody Wins. In Franz Rainer, Francesco Gardani, Wolfgang U. Dressler & Hans Christian Luschützky (eds.), *Competition in Inflection and Word-Formation* (Studies in Morphology), 169–193. Cham: Springer.

Guzmán Naranjo, Matías & Olivier Bonami. 2023. A distributional assessment of rivalry in word formation. *Word Structure* 16(1). 87–114.

Harris, Zelig S. 1951. *Structural linguistics*. Chicago, London: University of Chicago Press.

Hartmann, Stefan. 2016. *Wortbildungswandel. Eine diachrone Studie zu deutschen Nominalisierungsmustern*. Berlin, Boston: De Gruyter.

Hartmann, Stefan. 2018a. Derivational morphology in flux: a case study of word-formation change in German. *Cognitive Linguistics* 29(1). 77–119. https://doi.org/10.1515/cog-2016-0146.

Hartmann, Stefan. 2018b. *Deutsche Sprachgeschichte. Grundzüge und Methoden*. Tübingen: Francke.

Hay, Jennifer. 2001. Lexical frequency in morphology: is everything relative? Linguistics 39(6). https://doi.org/10.1515/ling.2001.041.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.

Hockett, Charles F. 1960. The Origins of Speech. *Scientific American* 203. 88–96.

Kempf, Luise. 2016. Adjektivsuffixe in Konkurrenz. Wortbildungswandel vom Frühneuhochdeutschen zum Neuhochdeutschen. Berlin, Boston: de Gruyter.

Kempf, Luise. 2021. Methoden der Produktivitätsmessung in diachronen Korpusstudien. In Christine Ganslmayer & Christian Schwarz (eds.), *Historische Wortbildung: Theorien-Methoden-Perspektiven* (Germanistische Linguistik 252–254), 23–54. Hildesheim: Georg Ollms Verlag.

Kempf, Luise & Stefan Hartmann. 2018. Schema unification and morphological productivity: A diachronic perspective. In Geert E. Booij (ed.), *The Construction of Words. Advances in Construction Morphology*, 441–474. Berlin: Springer.

Leclercq, Benoît & Cameron Morin. forthc. Submission Title: No equivalence: A new principle of no synonymy. *Constructions*. doi 10.24338/cons-535.

Lieber, Rochelle. 1981. *On the organization of the lexicon*. PhD thesis. Boston: MIT. http://www.ai.mit.edu/projects/dm/theses/lieber80.pdf (last checked 27/12/2022)

Meibauer, Jörg. 1995. Wortbildung und Kognition: Überlegungen zum deutschen -er-Suffix. *Deutsche Sprache* 23. 97–123.

Neuhaus, H. Joachim. 1973. Zur Theorie der Produktivität von Wortbildungssystemen. In Abraham P. ten Cate & Peter Jordens (eds.), *Linguistische Perspektiven*, 305–318. Berlin, New York: De Gruyter.

van Marle, Jaap. 1992. The relationship between morphological productivity and frequency: A comment on Baayen's performance-oriented conception of morphological productivity. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1991*, 151–163. Dordrecht: Kluwer.

O'Donnell, Timothy J. 2016. *Productivity and reuse in language: a theory of linguistic computation and storage*. Cambridge, Massachusetts: The MIT Press.

Panagl, Oswald. 1982. Produktivität in der Wortbildung von Corpussprachen: Möglichkeiten und Grenzen der Heuristik. *Folia Linguistica* 16. 225–239. https://doi.org/10.1515/flin.1982.16.1-4.225.

Panther, Klaus-Uwe & Linda Thornburg. 2001. A Conceptual Analysis of English -er Nominals. In Martin Pütz, Susanne Niemeier & René Dirven (eds.), *Applied Cognitive Linguistics* (Cognitive Linguistics Research 19.2), 149–200. Berlin, New York: De Gruyter.

Pijpops, Dirk. 2020. What is an alternation? Six answers. *Belgian Journal of Linguistics* 34. 283–294. https://doi.org/10.1075/bjl.00052.pij.

Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. (Topics in English Linguistics 28). Berlin, New York: De Gruyter.

Plag, Ingo. 2021. Productivity. In Bas Aarts & April M. S. McMahon (eds.), *The handbook of English linguistics* (Blackwell Handbooks in Linguistics), 483–499. 2nd edn. Malden, MA, Oxford: Blackwell.

Plag, Ingo, Christiane Dalton-Puffer & Harald Baayen. 1999. Morphological productivity across speech and writing. *English Language and Linguistics* 3(2). 209–228.

Rainer, Franz. 1987. Produktivitätsbegriffe in der Wortbildungstheorie. In Hort Geckeler, Hans-Martin Gauger & Wolf Dietrich (eds.), *Grammatik und Wortbildung romanischer Sprachen*, 187–202. Tübingen: Narr.

Rainer, Franz. 2012. Morphological Metaphysics: Virtual, Potential, and Actual Words. *Word Structure* 5(2). 165–182.

Richerson, Peter J. & Robert Boyd. 2005. *Not by genes alone: how culture transformed human evolution*. Chicago: University of Chicago Press.

Roché, Michel. 1997. Briard, bougeoir et camionneur: Dérivés aberrants, dérivés possibles. In Corbin, Danielle & Fradin, Bernard & Habert, Benoît & Kerleroux, Françoise & Plénat, Marc (eds.), Silexicales 1: Mots possibles et mots existants, 241–250. Lille, France: Université Lille 3. URL https://stl.univ-lille.fr/fileadmin/user_upload/laboratoires/stl/Roche_Reperes_morphologie_p42-50.pdf (last checked 19/11/2023)

Rowland, Caroline. 2014. *Understanding child language acquisition*. London, New York: Routledge.

Säily, Tanja. 2011. Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus Linguistics and Linguistic Theory* 7(1). 119–141. https://doi.org/10.1515/cllt.2011.006.

Schäfer, Roland. 2015. Processing and querying large web corpora with the COW14 architecture. In Piotr Bański, Hanno Biber, Evelyn Breiteneder, Marc Kupietz, Harald Lüngen & Andreas Witt (eds.), *Proceedings of challenges in the management of large corpora 3 (CMLC-3)*. Lancaster: IDS. http://rolandschaefer.net/?p=749.

Schäfer, Roland & Felix Bildhauer. 2012. Building Large Corpora from the Web Using a New Efficient Tool Chain. In Nicoletta Calzolari, Khalid Choukri, Terry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of LREC 2012*, 486–493.

Scherer, Carmen. 2005. *Wortbildungswandel und Produktivität: Eine empirische Studie zur nominalen er-Derivation im Deutschen*. . Vol. 497. (Linguistische Arbeiten). Tübingen: Max Niemeyer.

Scherer, Carmen. 2006. Was ist Wortbildungswandel? *Linguistische Berichte* 205. 3–28.

Scherer, Carmen. 2007. The Role of Productivity in Word-Formation Change. In Joseph C. Salmons & Shannon Dubenion-Smith (eds.), *Historical Linguistics 2005*, 257–271. (Current Issues in Linguistic Theory 284). Amsterdam and Philadelphia: John Benjamins.

Scherer, Carmen. 2015. Change in productivity. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word Formation: An International Handbook of the Languages of Europe* (HSK), vol. 3, 1781–1793. Berlin, New York: De Gruyter.

Schneider-Wiejowski, Karina. 2011. *Produktivität in der deutschen Derivationsmorphologie*. Bielefeld: (Diss.).

Schultink, Hendrik. 1961. produktiviteit als morfologisch fenomeen. *Forum der Letteren* 2. 110–125.

Trips, Carola. 2009. *Lexical Semantics and Diachronic Morphology: The Development of -hood, -dom and -ship in the History of English*. (Linguistische Arbeiten 527). Tübingen: Niemeyer.

Uhrig, Peter. 2015. Why the Principle of No Synonymy is Overrated. *Zeitschrift für Anglistik und Amerikanistik* 63(3). 323–337. https://doi.org/10.1515/zaa-2015-0030.

Yang, Charles D. 2016. *The price of linguistic productivity: how children learn to break the rules of language*. Cambridge, Massachusetts: MIT Press.

Zehentner, Eva. 2019. Competition in Language Change: The Rise of the English Dative Alternation. Berlin, Boston: De Gruyter.