Blocking as a Function of the Nature of Linguistic Representations: Where Psycholinguistics and Morphology Meet

Versloot, A.P.; Hoekstra, E.

DOI
10.1007/978-3-030-02550-2_6

Publication date
2019

Document Version
Final published version

Published in
Competition in Inflection and Word-Formation

Blocking as a Function of the Nature of Linguistic Representations: Where Psycholinguistics and Morphology Meet

Arjen P. Versloot and Eric Hoekstra

Abstract This paper addresses the question to what extent morphological blocking in language is a rule-based phenomenon. We argue that language users do not operate with a blocking rule, but that a form preference emerges as a result of cognitive selection mechanisms in a neural network of linguistic information. The actual target form develops its own token frequency in a probabilistic process, known as Preferential Attachment. After some time and some generations, one form will develop a nearly absolute dominance with its own local token frequency. This model implies that there is no blocking as an active negative action, but only a local lemma specific frequency, built up by a stochastic Preferential Attachment process, which favours one of the theoretically possible forms and, as a consequence, 'suppresses' the other options.

Keywords Analogy · Blocking · Exemplar memory · Language contact · Neural networks · Probabilistic model · Psycholinguistics · Token frequency

1 Blocking from a Psycholinguistic Perspective

Morphological blocking can be described as the "non-occurrence of one form due to the simple existence of another" (Aronoff 1976). In this article we will explore a psycholinguistic view of morphological blocking phenomena. In this view, language is represented as a neural network. The neural representation of language is a complex issue in which strictly linguistic theories such as traditional grammar or generative grammar are confronted with the experimental results of psycholinguistic

A. P. Versloot (✉)
German and Scandinavian Languages & Cultures, University of Amsterdam, Amsterdam, The Netherlands
e-mail: a.p.versloot@uva.nl

E. Hoekstra
Fryske Akademy, Leeuwarden, The Netherlands
e-mail: ehoekstra@fryske-akademy.nl

F. Rainer et al. (eds.), *Competition in Inflection and Word-Formation*, Studies in Morphology 5, https://doi.org/10.1007/978-3-030-02550-2_6

experiments, and vice versa, for psycholinguistic conceptions of language must also confront various problems which have hitherto been analysed and understood from a purely linguistic point of view. It is our purpose, in this contribution, to take some well-established insights from both psycholinguistics and morphology and present an explanation for the existence of blocking phenomena using a basic statistical learning algorithm applied to a model of grammar, built from exemplars in a neural network (see for similar approaches e.g. Maslen et al. 2004; Beckner and Wedel 2009: 5, *passim*; Bod 2009). Maslen et al. (2004) provide empirical evidence for a frequency-based, rather than rule-based trigger of morphological blocking, although they do not provide a working model of the actual underlying mechanism, as we will provide in Sect. 5.

Our model is empirically supported by data testifying to the rivalry between the two Frisian de-adjectival noun suffixes *-ens* and *-heid*, as in *wurgens/wurchheid* 'tiredness' from *wurch* 'tired' (Versloot and Hoekstra 2016). One of these suffixes is homophonous to its Dutch counterpart *-heid*, whereas the other suffix does not have such a Dutch counterpart. This fact is relevant because speakers of Frisian are all bilingual in Dutch and Frisian. An analysis of corpus data yields the conclusion that the choice between the two suffixes in Frisian is sensitive to the existence of *-heid* formations in Dutch. More specifically, a given *-heid* formation in Frisian gets a boost by means of co-activation, also referred to as secondary activation (see below and see Sect. 3.1), from a corresponding *-heid* formation in Dutch depending on two factors:

1. the degree of phonological similarity existing between the two
2. the frequency of the Dutch *-heid* formation involved

Psycholinguistic experiments have provided independent evidence that these two factors are relevant for the organisation of linguistic representations in the mind (on similarity, e.g. Dijkstra 2003, 2008; Smits et al. 2006, 2009; on frequency, e.g. Bybee 1995; Diessel 2007; Krott et al. 2001). Thus our general approach is strictly mentalistic and probabilistic (compare Aronoff 2019, for a categorical approach inspired by the ecological principle of mutual exclusion).

Our data will suggest that blocking is a tendency that comes into play alongside co-activation. It is well-known that similar items co-activate each other, but it also makes sense to suppose that items may not be too similar, for if they are, it becomes too hard to distinguish them when only one of them needs to be accessed. Our research suggests the following. In case two items surpass a given degree of similarity, then in addition to activating each other, a small frequency difference between them will easily develop from a tendency into a winner-takes-(almost) all situation. This is the central thesis of this paper. The relevance of the *-ens / -heid* data in Frisian is that they show this tendency still at work, before developing into normal cases of blocking. Preferential Attachment is the formal mechanism proposed to be at work in such cases, which are characterised by high semantic and formal co-activation. The computational principle of Preferential Attachment is directly responsible for blocking effects, and we are witnessing it in full operation in the choice between the suffixes *-ens* and *-heid* in Frisian.

2 Rivalry Between *-ens* and *-heid* in a Bilingual Setting

Frisian has two suffixes which turn adjectives into nouns: *-ens* and *-heid* (for an extensive description, see Versloot and Hoekstra 2016 and the references given there). Some examples have been given in (1) and (2):

(1)	*Healwiis* 'foolish'			
	+ − *heid*	=	*healwiisheid*	'foolishness'
	+ − *ens*	=	*healwizens*	'foolishness'
(2)	*Wiis* 'wise'			
	+ − *heid*	=	*wiisheid*	'wisdom'
	+ −*ens*	=	*wizens*	'wisdom'

There is large scale rivalry between these suffixes. They compete without showing clear evidence of blocking, though we will see that the data reveal that blocking is present as a tendency in the frequency distributions per lexical item. Furthermore, the distribution of *-ens* and *-heid* in Frisian is affected by the presence of *-heid* formations in Dutch depending on the similarity of the Dutch *-heid* formation to its Frisian equivalent and depending on the frequency of the Dutch *-heid* formation (Versloot and Hoekstra 2016). This is not surprising since all speakers of Frisian are bilingual, Dutch is heavily encroaching on Frisian (Ytsma 1995; de Haan 1997) and we know from psycholinguistic experiments that words from separate languages are stored closely together if they resemble each other and hence may influence each other by secondary activation.

An example of the rivalry between *-ens* and *-heid* is the rivalry between *wiisheid* (n = 243) and *wizens* (n = 38) 'wisdom' in a corpus of twentieth-century Frisian. Frisian *wiisheid* is supported by the closely similar Dutch word *wijsheid*, by which it is regularly secondarily activated. Frisian *wizens* is not supported by a Dutch item, as Dutch fails to have nominalisations in *-ens*. As a result, we expect that the general drift will be for Frisian *wiisheid* to gain ground at the expense of *wizens*. This expectation is correct, as Table 1 below makes clear.

We investigated all examples of rivalry between these suffixes. Our investigation was set up in the following manner. All *-ens* and *-heid* formations were counted in the Frisian Language Corpus, 1980–2000, which includes ca. 20 million tokens. The summed frequency of *-ens* and *-heid* for a given adjectival stem had to be 4 or more in total in order to guarantee a minimal robustness of the data. This yielded

Table 1 The use of the suffixes *-heid* and *-ens* in the word 'wisdom' in the twentieth century. The observed contrast is statistically significant with p = 0.01 (two-tailed Fisher's Exact Probability Test, http://www.langsrud.com/stat/fisher.htm)

'wisdom'	1900–1980	1980–2000
Wizens	38	0
Wiisheid	208	35

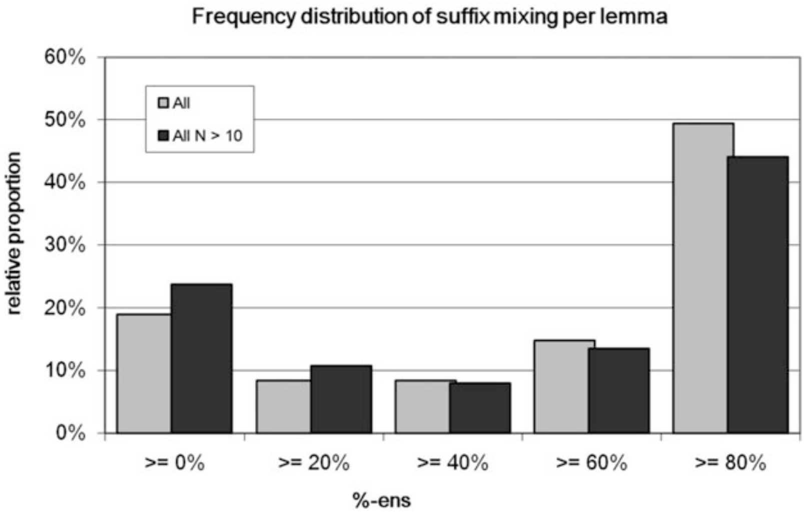

Fig. 1 Bar chart of amount of types for 5 cohorts of percentages of tokens of –ENS

336 adjectival stems which could take either *-ens* or *-heid* (11,167 tokens). For every stem (type), the number of tokens in *-ens* and in *-heid* was counted, and the percentage of tokens in *-ens* was calculated. Trivially, for every item, the percentage of tokens in *-heid* is 100% minus the percentage of tokens in *-ens*. We then counted how many stems (types) had 0–20% tokens in *-ens*, how many items had 20–40% tokens in *-ens*, and so on. The result has been presented in Fig. 1 above (from Versloot and Hoekstra 2016).

The grey bars present the distribution in case n > 3, whereas the black bars show what the distribution is like if we take only types with n > 10. The fact that there is little difference indicates that the distribution does not depend on low frequency *-ens* and *-heid* formations: the characteristic distribution is also found with high frequency formations.

If the choice of suffix was arbitrary, with a probability for either suffix of 0.5 on every occasion a stem is used, we would expect a normal distribution around the average 50%. The middle bars would be highest, and the adjacent bars lower, and the bars at the extremities would almost be zero. But what we see is a bimodal distribution, that is, the extreme categories have most items. Thus, the data already show a tendency to a lexical blocking effect. The endpoint of this tendency would be a situation in which blocking is fully operative, that is, a situation represented by a variant of the bar graph above in which the three middle columns would be zero, and the peripheral columns would be filled by a number of stems (types) taking either 100% *-ens* or 100% *-heid*.

3 Facts About the Neural Representation of Language

3.1 Introduction

Linguistics is a cognitive science, in the sense that the ability to use language is a cognitive ability. Hence research in cognition may well affect linguistic research and vice versa. The interaction between multiple disciplines may sometimes yield an overabundance of terminology. In our contribution, we will first attempt to show that a morphological puzzle can be solved with the help of notions like frequency and analogy (similarity), and that these notions tell us something about the organisation of language in the human mind. This claim is supported not just by linguistic data taken from corpora, but also by psycholinguistic experiments carried out with a completely different methodology and with a different overarching theory. In fact, psycholinguists have made crucial contributions to morphological issues (Rainer 2016: 20, and the references cited there). There are many examples showing that both a purely linguistic (corpus-based or analytical or experimental) approach (e.g. Arndt-Lappe 2014; Bybee 2007, 2010; Haspelmath 2006; Hay and Baayen 2002 and others) and a psycholinguistic approach (cf. the references in Sect. 1) independently arrive at the conclusion, almost a truism, that frequency and analogy are central to understanding how language is organised in the human mind. Such a convergence of results from very different angles provides the strongest possible support for the relevance of frequency and analogy. This view in turn will make it possible for us to see blocking as a tendency resulting from the structure of the information space in which the facts of language are represented.

3.2 Frequency

Probabilistic considerations have not received much attention in generative grammar. Chomsky wrote in the late fifties of the previous century: "I think ... that probabilistic models give no particular insight into *some* (our italics, AV and EH) of the basic problems of syntactic structure" (1957: 17). We can optimistically conclude from this that Chomsky leaves open the possibility that there are *other* basic problems which probabilistic models do give insight into, even though this may not be what Chomsky pragmatically intended. Of course, it is also fair to say that frequency studies have become far more precise and far more interesting with the advent of the computer and the availability of large corpora and the clever use which has been made of them. Perhaps morphology has been better prepared than syntax for the study of frequency, as it has always been clear that there may well be a relation between frequency and productivity when it comes to words.

It is commonplace in psycholinguistics that information in the human brain becomes more prominent as it is more frequently accessed (e.g. Bien et al. 2005). As a result, the representation of frequent items (words, phrases, suffixes) will be more prominent than the representation of infrequent items in the sense that the frequency of an item adds to its competitive strength (Krott et al. 2001). Psycholinguists measure this prominence by means of various tests, such as masked priming, which need not further concern us here. Linguists can measure the prominence of a word represented in the brain by proxy by investigating its frequency in a representative corpus, as noted by Bybee (1995: 452). Thus, frequency is not just a property of the outside world, of speech or text, but it also is by proxy an indication of the prominence of linguistic representations in the inside world of the human mind, in which the iteratively added perceptions of the outside world are represented.

3.3 Analogy (Similarity)

It has long been recognised (e.g. at least since Paul 1909 as noted in Rainer 2016, see also De Vries 1910–1912, and most recently from a general cognitive perspective, e.g. Hofstadter and Sander 2013) that analogy is an important organisational principle of the representation of language in the human mind. However, it used to be difficult to turn analogy into an exact, or even formalised, concept (Gardani et al. 2019, Sect. 2). As a result, questions about the direction and relevance of analogies could not be answered (cf. Sturtevant's Paradox in Fertig 2013: 97). After all, there are many a priori analogies which could be relevant, but never are. As with frequency, psycholinguistics provides us with more insight into the nature of analogy as an organisational principle of the human mind. Thus, it has become clear that words from two different languages are nevertheless stored closely together in the bilingual mind in case they are similar in form and/or meaning (Hulk and Müller 2000; Dijkstra 2003, 2008; Smits et al. 2006, 2009, among others). This makes it clear that analogy is a basic principle in the organisation of linguistic information, overriding even the distinction between two languages. Recently, analyses have been developed in which analogy has been turned into a precise notion from which linguists can derive exact predictions about the phenomena which they study (e.g. Arndt-Lappe 2014). Such analyses presuppose a psycholinguistic conception of analogy in which pieces of linguistic information are stored 'closer' together in case they are more similar. Thus, the more similar two elements are, the smaller the representational distance between them is. Representational distance can be measured in various ways. Phonological similarity is often measured by means of Levenshtein distance (e.g. Heeringa 2004). Arndt-Lappe measures phonological similarity in a simplified manner by means of analogical modelling based on Skousen's work (1989, 2005 and elsewhere). These ways of measuring similarity or analogy are approximations of a psycholinguistically real measure of this concept, which constitute a fairly reliable measure of representational distance (e.g. van Heuven 2011: 11).

4 Strength of Co-activation and Blocking

4.1 Blocking as Differentiation by Means of Bimodal Frequencies

Co-activation is a psycholinguistic notion according to which if the human mind is searching for (trying to primarily activate) a specific piece of information A, then all information B will be secondarily co-activated depending on the following two factors:

1. The more similar (analogous) B is to A, the stronger B will be secondarily activated
2. The more frequent A and B are, the stronger will be the secondary activation

This makes psycholinguistic sense if we translate similarity and frequency into representational distance in an information space:

Analogy:
The smaller the representational distance between B and A, the stronger B will be secondarily activated
Frequency:
The stronger the representational prominence of A and B, the stronger will be the secondary activation

All activation, whether primary or secondary, depends on frequency. For primary activation, only the frequency of the primarily accessed item counts, but secondary activation will involve the frequencies both of the primarily accessed item and of the secondarily accessed item. We have shown elsewhere that for items with the same meaning, the strength of co-activation, or secondary activation (A2), is given by the following formula (Versloot and Hoekstra 2016):

$A2 \sim {}^{2}\log(\text{frequency})/\text{LevenshteinDistance}$

The formula above is concerned with formal dissimilarity only, not semantic dissimilarity. Levenshtein distance is a measure of the dissimilarity between two strings. It counts the minimal number of deletions and insertions needed to transform one string into another.

This formula implies that two pieces of linguistic information co-activate each other more strongly, the more they resemble each other in form and the higher the frequencies are. In the normal case, co-activation will be large enough to be measurable only for a small set of (very similar) words. Secondary activation only affects items which are in the immediate neighbourhood of the primarily activated item, that is, which are very similar in form and meaning. Secondary activation is not a symmetric notion. It is partly defined in terms of frequency, and the frequencies of items involved are not interchangeable. The effect of the formula above is that it predicts the strength of co-activation to correlate with the degree of similarity between two given items and with the frequency of the two items. Thus, a given

Table 2 The impact of secondary activation (A2) from Dutch on the use of the suffix *-heid* in Frisian. The numbers in each cell represent the number of lemmas that prefer the suffix *-heid* per 10 lemmas. Low frequency refers to the bottom 20% in frequency, high frequency to the top 20%. Low similarity to the top 20% in Levenshtein Distance (LD), high similarity to the bottom 20% in LD

#-*heid*/10 lemma's	Low frequency	High frequency
Low similarity	1/10 F: *linigens* D: *soepelheid* 'flexibility'	3/10 F: *ienriedigens* D: *eensgezindheid* 'unanimity'
High similarity	4/10 F: *ûnachtsumheid* D: *onachtzaamheid* 'nonchalance'	8/10 F: *wurklikheid* D: *werkelijkheid* 'reality'

Frisian word in *-heid* will be affected by its semantic equivalent in *-heid* in Dutch depending on the frequency of the Dutch equivalent in *-heid* and on the degree of similarity between the two words. Table 2 presents the outcome of calculating the strength of secondary activation for four categories of items: items of low and high frequency cross-classified with low and high similarity. Strong co-activation yields a greater probability that a given Frisian item ends in *-heid* rather than *-ens*. An example has been added in each category:

The number in each category represents how many items will carry the suffix *-heid* (as against the rival suffix *-ens*). Thus, given 10 instances of *-heid/-ens* items of low frequency and low similarity, our formula correctly predicts that 1 out of 10 times an item will carry the suffix *-heid* and 9 out of 10 times it will carry the rival suffix *-ens*. These numbers can also be read as probabilities of encountering the suffix *-heid*. As predicted by the formula, the effect of secondary activation from Dutch is strongest when both the frequency and the similarity are high. The least use of *-heid* is found among items that are both very different in both languages and have a low frequency of occurrence. The lemma's described above don't have final stress, since final stress is an independent factor in regulating the choice of suffix (see Versloot and Hoekstra 2016).

The question naturally arises what happens in case two items are similar semantically, yet they differ with respect to their frequencies. Take for example the rivalry between *stealer* and *thief*. These words are semantically practically identical, but formally different. The sheer frequency difference between these two is responsible for the fact that *thief* seems to block *stealer*. The inclusion of frequency in the aforementioned formula for secondary activation implies that low-frequency items endure more interference from high-frequency items than the reverse. As a consequence, *thief* will suppress the activation of *stealer* but does not necessarily eliminate it. What seems to be an example of absolute blocking turns out to be relative blocking, that is, a matter of probability, as corpus investigation makes clear. Usually, *thief* is primarily activated, but occasionally *stealer* is primarily activated. This is not only clear from a cursory inspection of the corpus Internet with Google search, but also from dictionaries reporting the word *stealer*. For example, on the Internet we find both the phrase *the stealer of souls* and the phrase *the thief* *of souls*, without any systematic difference in meaning. As *thief* and *stealer* fail to share any formal characteristics, this must be considered as an example of semantic blocking. The example makes it clear that blocking is not absolute, but probabilistic; compare also the example of *wiisheid* versus *wizens* 'wisdom', presented in Sect. 2.

So, when two competing items differ hugely with respect to their frequencies, it is the frequency difference itself which is responsible for the fact that one item blocks the other, or, more exactly, that the probability of one item is much higher than that of the other. Blocking comes out as a tendency to differentiate highly similar items by means of their frequencies: the winner takes (almost) everything, the loser (almost) nothing, as exemplified by the frequencies for *thief* and *stealer*, which show a proportion of 50:1.[1]

4.2 A Rule-Based View of Blocking

A different view is presented in Embick and Marantz (2008: 21–22), who discuss the contrast exemplified by English *gloriosity* and *gloriousness*. They turn high probabilities into 100% probabilities, as a consequence of their rule-based view. They claim that the structure assigned to *glory* (an example of inner-structure suffix assignment) is not available to *gloriosity*. The root *glory* does not appear on the list of roots specified by the suffix *-ity*. Furthermore, they claim the structure assigned to *gloriousness* is not available for *gloriosity* either, since outer-structure suffix assignment (not a listed property) specifies that the nominalisation must be pronounced *-ness*. All in all, two stipulations take care of the supposed observation that *gloriosity* is blocked by *glory*.

This does not give us any specific insight into what is going on. Nevertheless, we agree with Embick and Marantz's general attempt to derive the relevant observations without any explicit blocking principle. However, on closer scrutiny it turns out that Embick and Marantz incorporate a blocking-like principle into their specification of the structures in which the vocabulary items *-ity* and *–ness* are spelled out. They specify that *-ity* is restricted to a list of roots, whereas *-ness* is not. Below, their specifications for *-ity* have been literally reproduced (Embick and Marantz 2008: 21), where capital V is used to introduce roots:

Vocabulary items
N < –> – ity / X_____
X = Roots (VAtroc, VCurious ...); [a, able], [a, al]
n < –> – ness

[1] https://books.google.com/ngrams/graph?content=thief%2Cstealer&year_start=1980&year_end=2000&corpus=0&smoothing=10&share=&direct_url=t1%3B%2Cthief%3B%2Cc0%3B.t1%3B%2Cstealer%3B%2Cc0 (April 14, 2017).

Crucially, the root *curious-* has above been specified as allowing the nominalising head n to be spelled out as *-ity*, deriving *curiosity*. But the root *glorious-* has not been thus specified, being absent from the list of roots allowing the nominalising head n to be spelled out as *-ity*. Furthermore, "With some heads [. . .] the n head defaults to the phonology *-ness*" (Embick and Marantz 2008: 22). The notion default is of course a central ingredient of blocking. Thus, they built in blocking, hidden in their hard and fast rules for the insertion of vocabulary items. For further criticism of Embick and Marantz, see Rainer (2012).

The relevant observations to counter their view concern the difference in probability between the three options *glory, gloriousness* and *gloriosity* (cf. also Arndt-Lappe's 2014 insightful study on the rivalry between *-ity* and *-ness*). On the Internet, in the Corpus of Contemporary American English and in the British National Corpus (Brigham Young University), the difference between *glory* and *gloriousness* is roughly a factor 1000 in favour of *glory* over *gloriousness*. In addition, the Internet is the only corpus to feature a sprinkling of examples of *gloriosity* (as is, by the way, conceded by Embick & Marantz). As unanalysed data tells us nothing, let's have a look at examples involving *gloriosity*. Sure enough, some of the examples are from linguistics articles. But there are also some examples which cannot be thus dismissed, as (1–4) below:

(1) *Oh sure, I put things in the ground, but what they decided to do from there on is out of my control. And then the next thing I know, gloriosity.*(http://plantlust.com/blog/2014/11/i-need-a-thicket/)
(2) *Wrasslin' with the Golden Gloriosity is conspicuously absent from the list.*(http://www.democraticunderground.com/1018248814#post1)
(3) *Withholding information, excessive gloriosity and a lack of realism, all are forgiven because of his aspect as a holy man.* (http://articles.chicagotribune.com/1997-02-23/entertainment/9702230089_1_anthony-bianco-reichmann-family-paul-reichmann)
(4) *Slow down and take the time to enjoy the gloriosity of this season! With love, Fran* (http://www.fransorin.com/inspirational-quotes-for-fall/)

What's more, even if you have never heard or used this word, it is directly clear what is meant by it. So, it seems that it is simply an unwanted idealisation to suppose that the probability of *gloriosity* is zero. Actually, it is a form with a very low probability, deriving from its low frequency (see for the interaction between frequency of occurrence in the input and selective, blocking-like morphological productivity, Maslen et al. 2004). Nevertheless, it is directly understandable because the human mind classifies linguistic information depending on its formal and semantic similarity to other linguistic information, and as a result an L1 speaker of English, and even an L2 speaker of English, will directly analyse *gloriosity* as a nominalisation of *glorious*, even though this person may never have heard this word before.

In order to explain this, Embick and Marantz would have to say that there are some speakers who spontaneously started to spell out the allomorph of the nominalisation as *-ity*, not as *-ness*, as a second option (because these speakers will also have the noun *gloriousness* at their disposal). Since these speakers will choose either suffix depending on some probability measure, it is obvious that probability must play a role in the mental lexicon of these speakers. But if Embick and Marantz need a probability measure anyhow, they can just as well entertain the psycholinguistically motivated view of grammar that is currently being developed in quantitatively oriented literature (as in the present article). Furthermore, there is extensive evidence that the human mind is sensitive to probability, or its linguistic equivalent, relative and absolute frequencies (for example Labov et al. 2006), and numerous examples can be cited from literature on psycholinguistics and on frequency. Frequency can be used to shed light on many linguistic puzzles, as was noticed early on already by Bybee (2007, a collection of articles partly from the nineties). In fact, older literature draws attention to the role of frequency (e.g. Schuchardt 1885/1972). The discussion above illustrates that there seems to be a winner-takes-(almost)-all tendency which is responsible for the empirical effects described as blocking phenomena.

Of course, the question arises why real competition among two forms is a relatively rare phenomenon. It is clear that there is no real competition between *glory, gloriousness* and *gloriosity*, seeing that their frequencies are too different for there to be any such competition. And this holds generally true of the frequency distributions between rival forms, so much so that the blocking principle has been hypothesized to explain this state of affairs. However, the evidence presented in Sect. 2 and in this section provides further support for the claim that blocking is a tendency inherent to the way linguistic information is stored and accessed in the human mind. Furthermore, blocking must be considered a winner-takes-most tendency that differentiates the frequencies of semantically highly similar competing items, with the frequency of the loser dwindling to almost zero. It is almost as if highly similar items must be distinguished by their frequencies, and this is not so strange as it seems, given that human beings have perception and knowledge of frequency information (Labov et al. 2006, among others). A case of a frequency asymmetry is provided by the rivalry between five (!) diminutive suffixes in Italian, where nearly always one of them, the *ino*-diminutive, has by far the highest token frequency (Dressler et al. 2019, see their Tables 1 and 2). Similarly, competition is reported in Italian between expressions of the type *rosso corallo* ‘coral red’ and those of the type *rosso come il corallo* ‘red like a coral’ (Masini 2019, Sect. 4.1). Here again, the frequency of one of the two competitors (*rosso corallo*) dwarfs the frequency of its competitor (*rosso come il corallo*), although Masini argues that these competitors are distinguished by slightly different shades of meaning as well. It seems natural to suppose that as the difference in meaning between two competitors becomes smaller, the difference in frequency must be larger. In addition, many of the examples reported by Masini as non-existent words are found on the internet, with very low frequencies, as was the case with *gloriosity* and with *stealer*. This applies for example to the competition between two types of compounding in Italian, exemplified by *capo di governo / #capogoverno* and *capogruppo / #capo di gruppo*, where the variants marked with # are nonetheless found on the internet, with very low frequencies.

The question now arises which mechanism in the human mind is responsible for blocking, for it is not clear at first sight how blocking, conceived of as a winner-takes-(almost)all principle can be derived from any meaningful interaction between frequency and analogy. On the other hand, we already argued on purely conceptual grounds that two pieces of linguistic information which are very similar cannot occupy the same position, that is, there must be some minimal distance between them in the representational space in which they are stored. The question then arises how this conceptual view can be formalised in an exact manner.

5 Blocking as a Winner-Takes-All Principle of the Organisation of Linguistic Representation

5.1 The Contours of an Explanatory Model

In order to understand the lexical preference for either of two possible forms, we have to model two constellations:

1. The moment the lexical item arose in the history of the language for the first time and speakers developed some form of preference, be it relative or (nearly) absolute, see model 1 (Sect. 5.2)
2. The moment a new speaker acquires the actual lexical item, given an existing bias in the availability in the input of the two lexical/morphological variants, which are competing to express one meaning, see model 2 (Sect. 5.3)

This second situation does not differ from the competition between, for example, strong verb forms and analogically formed weak past tense forms. It is well known that here token frequency is a major factor (Maslen et al. 2004; Lieberman et al. 2007; Strik and Versloot 2015).

We will provide a model of language acquisition, here the acquisition of two or three competing words or word forms for one meaning, using a general learning strategy that can be formalised into a statistical learning procedure, following Beckner and Wedel (2009). We will consider meaning to be monolithic and invariable. Any variability in the semantics may lead to semantic differentiation of the two competing expressions, with an overlap in core meaning, but with different connotations or genre-specific application. This is normal if two words survive: they become specialised and no longer compete.

We assume that linguistic information is stored in the form of exemplars in a neural network, where items are connected on the basis of formal and semantic overlap, where some form of nearness (degree of representational distance) expresses the level of similarity and where the frequency of use of specific linguistic information defines the strength of these connections (for examples of exemplar based grammars see Skousen et al. 2002; Daelemans 2002; Beckner and Wedel 2009; Bod 2009).

The target form attested in the exemplar memory has to compete with other forms that are found in a close semantic and/or formal distance, including forms that are generated through co-activations of adjacent forms. An example of the latter process would be a potential plural form **gooses* for *geese*, based on the co-activation of the semantically very nearby singular concept *goose* and the plural morpheme *-(e)s*, which results from the aggregated evidence from all regular plural noun exemplars in the memory. We do not go further into this look-up mechanism and concentrate on the building of the exemplar memory, taking the probabilities produced by the activations and co-activations that are produced during the look-up procedure as a starting point.

5.2 *Model 1: Acquiring a Preferred Form from Scratch*

In order to understand how something that looks like morphological (or lexical) blocking develops, we assume a language user building a new morphological form (derivation, inflection, etc.) to express a new concept. A child, for example, has neither used nor heard any word for a specific given concept before and it has not been around in the language of any other user so far. To get the model running, we assume, for the sake of the example, that the neural network suggests three options, based on the forms and meanings of other related items/exemplars in the network: two of them with a likelihood $p = 0.45$, the third one with $p = 0.10$ ($\Sigma p = 1$). This is not a mere technicality, for it mirrors the fact that the language user has knowledge of all the other exemplars in her or his memory. In this respect, we follow Skousen (1989) who uses an analogical model that posits such likelihoods to produce a probability distribution to start out with.

The competing options can, for example, be competing suffixes (as in our -*ens*/-*heid*-study) or a new compound.[2] To substantiate these system-induced probabilities for our exemplar-driven simulation model, it must work with three dummy exemplars for which we take the size of their p-values. These dummy exemplars represent the speakers' expectations for new words, based on their knowledge of available exemplars in the rest of the language. Then the speaker produces a form to refer to this new concept. One form is randomly chosen from the three options, using the system-derived probabilities. We now make an assumption for which there is overwhelming psycholinguistic evidence: "[. . .] that usage experience leaves traces in memory that may feed back to future usage behavior [. . .]" (Beckner and Wedel 2009: 6). Thus, the produced form is added as a new exemplar to the exemplar set.

[2]To follow the mentioned example with three variants, it can be mentioned that Frisian has indeed a third alternative, next to *-ens* and *-heid*, namely *-te*: *gruttens, grutheid* and *grutte* all three exist, with substantial overlap in actual use and meaning.

The exemplar set contains 2 exemplars now: one from the first actual choice and the rest comprising the p-values-based dummy content. The distribution of these exemplars is used to compute a new probability for each of the three forms. This means that after having used the form for the first time, the speaker is 50% guided by her/his first choice and 50% guided by the system-derived probabilities. Based on these new probabilities, another form is produced and added to the exemplar set, which again leads to a recalculation of the probabilities. This mechanism is known as Preferential Attachment and was first rigorously described (although not under that name) by Yule (1925).

Various research, such as our own, indicates that there is a (near) linear correlation between preference for a form or resistance to regularisation of irregular forms on the one hand and the logarithm of the frequency, rather than the absolute frequency on the other. For the sake of simplicity, we work with the absolute frequency here, which enhances the contrasts and makes the model more volatile.

When using the meaning (and the various forms) more and more often while adding more and more exemplars to the memory, each new exemplar will add less weight to the probabilities of the forms. Also, the probabilities produced by the neural network will be of less and less importance. The speaker will end up with a fairly stable set of probabilities for the three forms based on the exemplars in her/his memory. Figure 3 shows three different results of this Preferential Attachment process after the first 200 exemplars being added to the exemplar memory.

In the example (I), Variant B becomes fairly dominant, but no variant wins out entirely. The middle graph (II) shows a relatively rare instance where the form with the lowest initial likelihood becomes the most common form, with the other forms on a pretty high secondary level. Graph (III) shows that the absolute dominance of one form is also a potential outcome. A typical feature of this model is that it stabilizes after a couple of runs and very often leaves more than one form as an option. This situation is typical for many of the word pairs in our data set with *-ens/-heid*.

5.3 *Model 2: A Succession of Generations*

We now go on to model the situation in which there is an existing bias in the availability in the input of the two lexical/morphological variants, that is, adult speakers already have knowledge of the competing forms. When a more skewed distribution as in Fig. 2 (I) is used as an input for a new learner with a 'blank slate', the impact on the new learner's exemplar set will be much bigger than on the first speaker in the adult state, who already has many exemplars in her/his memory. A new language learner, confronted with biased input as in Fig. 2 (I), has therefore a bigger chance to end up with one form being dominant (as in Fig. 2 (III)) after building her/his own exemplar set. One way to model this effect is to reset the exemplar memory of the agent after n runs to zero exemplars and using the probabilities as they are after n runs as the starting probabilities for this (new)

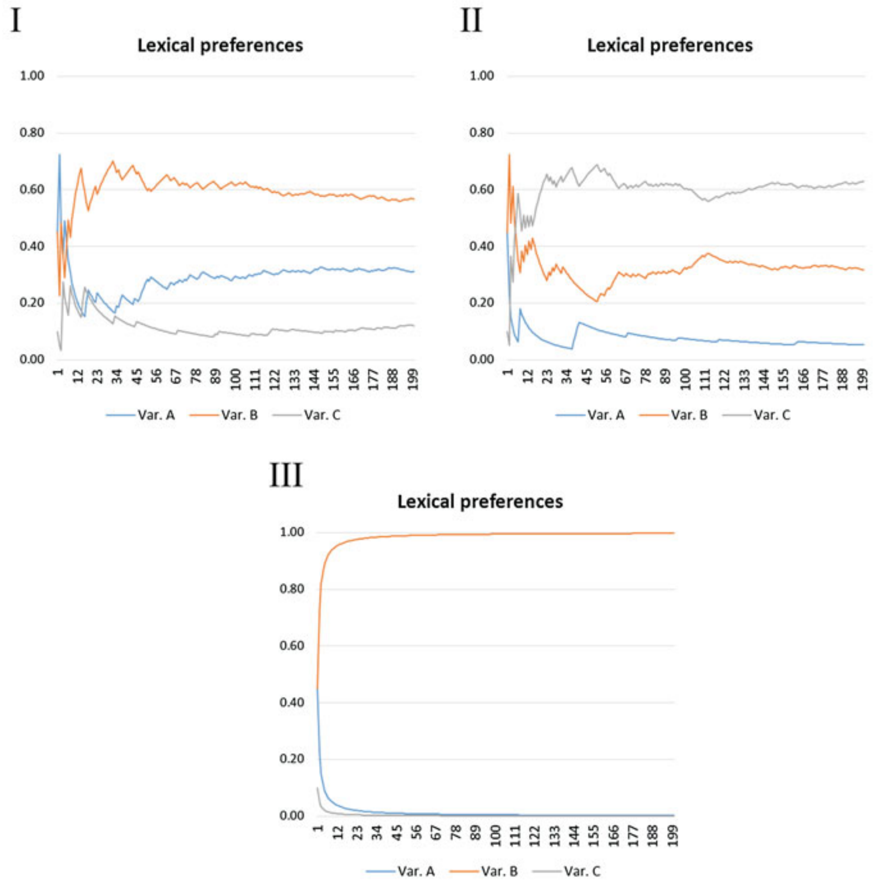

Fig. 2 Choosing a form from three possible variants A, B and C through the process of Preferential Attachment, showing three different runs

agent with an empty exemplar memory. The crucial aspect of this operation is to create a novel learner experiencing much more impact from new exemplars. We chose to model the effect of constantly new language learners added to the speaker population by keeping the total number of exemplars constant at a fairly low level of 20. In this version of the model, every new run is a proxy for a new learner in the population at the stage of 20 exemplars in her/his memory. Figure 3 shows two results of this probabilistic process after 1000 runs.

This variant of the model mostly produces patterns as in Fig. 3 (I) or patterns that resemble the one in Fig. 2 (III), with an even steeper cline and a clear winner. Figure 3 (II) illustrates the rarer instances of a long-standing competition of forms.

The first model represented the language user acquiring new forms. The second model represented the situation of a community in which there is rivalry between

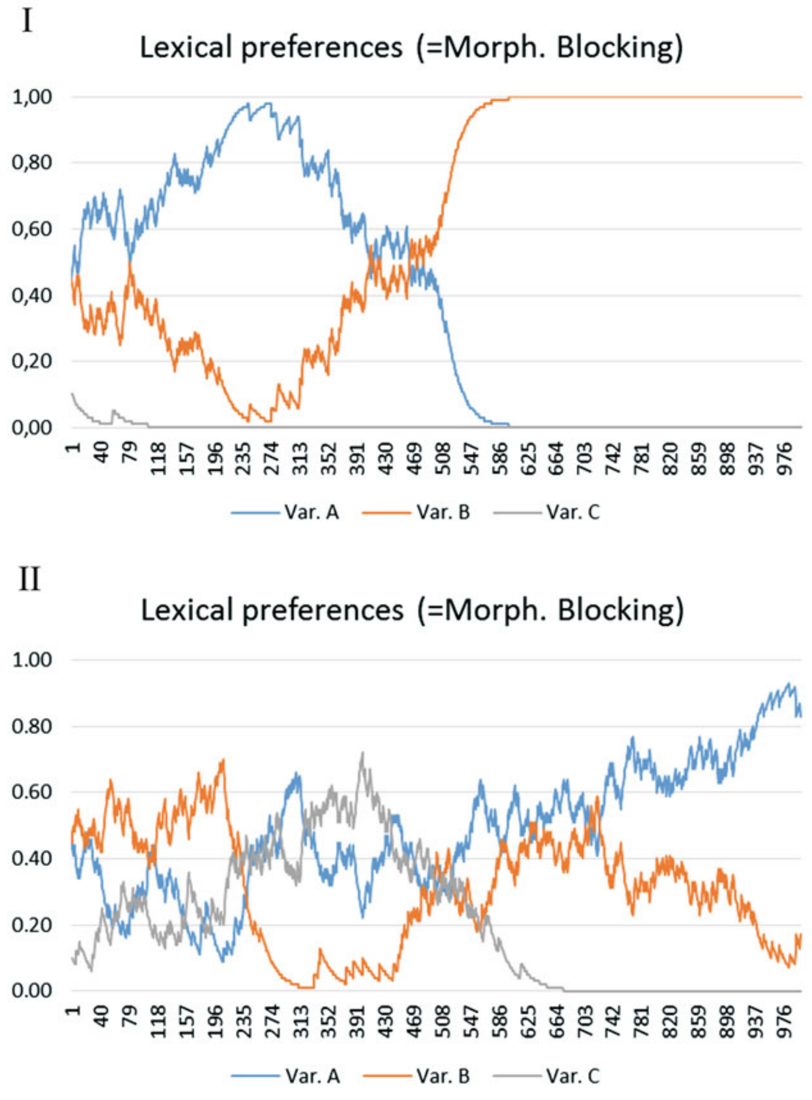

Fig. 3 Choosing a form from three possible variants through Preferential Attachment, implementing a young learner's effect in a generational sequence by keeping the number of exemplars on a constant low level

two competing items. Both models converge in their results in the sense that they tend to produce a situation in which one form develops towards a high frequency whereas its rival develops a low frequency. In this way, two competing, semantically equivalent items will be clearly differentiated by means of their frequency. Of course, sometimes there is no blocking effect, that is, sometimes variation is stable across a longer period of time. The English comparative provides an example of this. Our model includes the possibility for such long-standing rivalry.

The various graphs in Figs. 2 and 3 show that due to mere chance, there can be different winners and, moreover, there can be a situation without a clear winner. Even when the outcome of each individual process is unpredictable, there are clear trends, resulting from the Preferential Attachment mechanism and the initial probabilities. We ran 50 different tests of Model 2 (with 1000 runs per test): 11 times, there was no absolute winner, 39 times there was. This shows the inclination of the learning algorithm to develop an absolute preference for one variant. With every new speaker and new generation, the chances that the speech community will end up with one winner increase. The initial probabilities are also reflected in the chance for one of the variants to end up as the absolute winner. In another series of 50 tests of Model 2, the least likely variant C, with an initial probability of 0.1, came out as the winner six times and once as the dominant form.

Thus, the model clearly develops towards one winner (which looks like a discrete blocking 'rule') and the chances for variants to come out as winner are proportional to their initial probabilities, which fits intuitive expectations. However, the stochastic process presented here also accounts for otherwise unexpected developments and outcomes. The advantage of models like the above is that they incorporate psychological and morphological insights into the nature of linguistic representations.

6 Discussion

In the two versions of the model, we simulated how a probabilistic language learning process, based on an exemplar memory leads to skewed distributions between variants that were initially equally probable. It showed that one generation may end up with skewed distributions of multiple variants, while the accumulated result of multiple generations is more inclined to come up with one winner.

As stated in Sect. 5.1, we want to underline that there are two fundamentally different aspects of morphological blocking:

1. How does the competition work between two morphological variants, be it two equivalent derivational formations as in our *-ens/-heid* example from Frisian or between a regular and an irregular morphological form (such as past tense forms of verbs)?
2. How can it be that even from scratch, even in the hypothetical situation where two variants are assumed to be exactly equally likely from the very beginning, speakers and hence speech communities, develop a preference for either of the two at all?

The key to the first process is to assume an exemplar-memory grammar, where the frequency-defined weight of items in the storage defines their rate of application in speech production and hence their success to survive over generations (see for a similar approach and application of such an exemplar-memory grammar, e.g. Beckner and Wedel 2009). So, given an existing bias in the input, a basic behaviourist approach will lead new speakers to opt for the form that is used in their surroundings. In a more advanced model, based on a neural network formed by the total of the language exemplars, alternative, competing formations are continuously 'suggested' by the network through analogy. Even when we know that such a mechanism exists and works over generations and leads to large-scale regularisations (see, e.g., Lieberman et al. 2007 and the mentioned study by Becker and Wedel), we may expect some kind of clean-up mechanism to 'protect' irregular forms by compensating for basic system noise.

This clean-up mechanism is learning and building an exemplar memory through preferential attachment. It is this learning process that not only supports and continuously enhances the preference for the most frequent form, but it also accounts for the second phenomenon, i.e. the question why biases in frequency of occurrence of two variants arise at all. Thus, preferential attachment not only produces biases towards one of possible variants, it is also used to clean up 'noise', i.e. spontaneously generated analogical variants, errors, and so on.

In our model, we operated with one agent, who was sensitive to her/his own production. This does not differ fundamentally from a model with two or more agents, who are *additionally* sensitive to each other's production. An important step in this model is the relative unimportance of the weight of the suggestions made by the language system on the basis of analogy. In the first model, the added weight of the three suggestions was equivalent to only one exemplar. The century-long resistance of morphologically irregular high frequency items in languages implies that a high frequency of occurrence can indeed easily overrule analogical pressure (e.g. Bybee 1995; Strik and Versloot 2015; Versloot and Adamczyk 2018 and many others).

It seems that the impact from system analogy is not that easily overruled for items with a lower frequency of occurrence. Especially when introducing something like decay of exemplars into the model, less frequently used items will be prone to analogical pressure. When the two competing forms are a regular versus an irregular form, where the latter can never re-appear once it failed to be transmitted, it may eventually lead to regularization of the morphology and exclusion of the variation. Of course, irregularity can be reintroduced as the result of phonological processes or incidentally through analogy with other irregular forms.

When the system suggests multiple options through analogy with other exemplars in the system, a lower frequency of occurrence will lead to weakening of the lexical preference and a (re)introduction of variation. In addition, language contact, as in our Frisian example, can cause system noise or continuous availability of variants in the input, which cannot be easily eliminated by the speech community. An example comes from West Frisian Breaking, a vowel alternation pattern, that can e.g. occur in plural forms and diminutives of nouns with the root vowels *ie, oe, oa, ea* [iˑə, uˑə, oˑə, ɪˑə], in a similar vein as *Umlaut* in Modern High German morphology: it is phonologically regular, but unpredictable in its lexical distribution (Tiersma 1979: 17–20). An example is *beam – beammen – beamke* [bɪˑəm] – [bjɛmn̩] –

[bjɛmkə] ‘tree – trees – little tree’ with breaking, against *stream* – *streamen* – *streamke* [strɪˑəm] – [strɪˑəmn̩] – [strɪˑəmkə] ‘stream – streams – little stream’ without breaking. Breaking is non-existent in Dutch. It is observed that in the case of (nearly-)full Frisian-Dutch cognates, such as *dier* ‘animal’ (Dutch idem) or *roas* [roˑəs] ‘rose’ ∼ Dutch *roos* [roˑᵘs], breaking is more often absent in Frisian than in words that differ more strongly from their Dutch semantic counterpart (Tiersma 1979: 59). The reason is that the available Dutch input continuously poses a competing unbroken form as an alternative for the broken Frisian form, which cannot be eliminated by the bilingual language user, because the speakers of Frisian do not control the availability of variants from Dutch.

Whether an item has a high or a low frequency of occurrence, language history shows that analogically produced suggestions remain always around, ready to ‘attack’ especially irregular forms, or simply to be used in a creative way to build new word forms, especially for special purposes such as additional connotations. A typical example from English is the plural of *mouse* in the meaning of a ‘computer device’, which is often rendered as *mouses*.[3] Despite the fact that the irregular plural *mice* is firmly rooted and seems entirely stable in present-day English, an analogical form can be built by analogy with the type frequent plural forms of other nouns when such an occasion occurs due to a semantic extension of the concept.

7 Concluding Remarks

We modelled the emergence of a form preference among various theoretically possible variants (indirectly implying the blocking of other forms) in a neural network, built on an exemplar memory. This neural network produces initial likelihoods for competing variants through analogy with other exemplars for related items (based on similarity and frequency). The actual target form develops its own token frequency in a probabilistic process, known as Preferential Attachment. After some time and some generations, one form will develop a nearly absolute dominance with its own local token frequency.

This model implies that there is no *blocking* as an active negative action, but only a local lemma specific frequency, built up by a stochastic Preferential Attachment process, which favours one of the theoretically possible forms and, as a consequence, ‘suppresses’ the other options. As the neural network will still provide the basic likelihoods by analogy, secondary or alternative forms will in reality never become entirely 0%, but always be an option: either by chance or in intentional language use. Also, language contact can provide such competing forms, which remain part of the neural network. To sum, our approach combines insights from morphology and from psycholinguistics to posit a model which explains blocking as a tendency. Conceptually, blocking was viewed as an example of creating a distinction between two competing highly similar items by means of their frequency.

[3] See: https://www.merriam-webster.com/dictionary/mouse (visited March 30, 2017).

References

Arndt-Lappe, Sabine. 2014. Analogy in Suffix Rivalry – The Case of English *-ity* and *-ness*. *English Language and Linguistics* 18 (3): 497–548.

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

———. 2019. Competitors and Alternants in Linguistic Morphology. In *Competition in Inflection and Word-Formation*, ed. Franz Rainer, Francesco Gardani, Wolfgang U. Dressler, and Hans Christian Luschützky, 39–66. Cham: Springer.

Beckner, Clay, and Andrew Wedel. 2009. The Roles of Acquisition and Usage in Morphological Change. *Annual Meeting of the Berkeley Linguistics Society* 35 (1): 1–12. https://doi.org/10.3765/bls.v35i1.3593.

Bien, Heidrun, Willem Levelt, and Harald Baayen. 2005. Frequency Effects in Compound Production. *Proceedings of the National Academy of Sciences of the USA (PNAS)* 102 (49): 17876–17881. https://doi.org/10.1073/pnas.0508431102.

Bod, Rens. 2009. From Exemplar to Grammar: A Probabilistic Analogy-Based Model of Language Learning. *Cognitive Science* 33 (5): 752–793.

Bybee, Joan. 1995. Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10: 425–455.

———. 2007. *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

———. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Daelemans, Walter. 2002. A Comparison of Analogical Modeling of Language to Memory-Based Language Processing. In *Analogical Modeling: An Exemplar-Based Approach to Language*, ed. Royal Skousen, Deryle Lonsdale, and Dilworth B. Parkinson, 157–179. Amsterdam/Philadelphia: John Benjamins.

de Haan, Germen J. 1997. Contact-Induced Changes in Modern West Frisian. *Us Wurk* 1997: 61–89.

De Vries, Wobbe. 1910–1912. *Dysmelie: opmerkingen Over syntaxis. Verhandeling behoordende bij het Programma van het Gymnasium der gemeente Groningen 1910–1911 en 1911–1912*. Groningen: B. Jacobs.

Diessel, Holger. 2007. Frequency Effects in Language Acquisition, Language Use, and Diachronic Change. *New Ideas in Psychology* 25 (2): 104–123.

Dijkstra, Anton. 2003. Lexical Storage and Retrieval in Bilinguals. In *The Interface Between Syntax and the Lexicon in Second Language Acquisition*, ed. Roeland van Hout, Aafke Hulk, Folkert Kuiken, and Richard Towell, 129–150. Amsterdam/Philadelphia: John Benjamins.

———. 2008. *Met andere woorden: over taal en meertaligheid* [In Other Words: On Language and Multilingualism]. Inaugural address. Nijmegen: Radboud Universiteit.

Dressler, Wolfgang, Lavinia Merlini Barbaresi, Sonja Schwaiger, Jutta Ransmayr, Sabine Sommer-Lolei, and Katharina Korecky-Kröll. 2019. Rivalry and Lack of Blocking Among Italian and German Diminutives in Adult and Child Language. In *Competition in Inflection and Word-Formation*, ed. Franz Rainer, Francesco Gardani, Wolfgang U. Dressler, and Hans Christian Luschützky, 123–143. Cham: Springer.

Embick, D., and A. Marantz. 2008. Architecture and Blocking. *Linguistic Inquiry* 39 (1): 1–53.

Fertig, David. 2013. *Analogy and Morphological Change*. Edinburgh: Edinburgh University Press.

Gardani, Francesco, Franz Rainer, and Hans Christian Luschützky. 2019. Competition in Morphology: A Historical Outline. In *Competition in Inflection and Word-Formation*, ed. Franz Rainer, Francesco Gardani, Wolfgang U. Dressler, and Hans Christian Luschützky, 3–36. Cham: Springer.

Haspelmath, Martin. 2006. Against Markedness (and What to Replace it with). *Journal of Linguistics* 42 (1): 25–70.

Hay, Jennifer, and Harald Baayen. 2002. Parsing and Productivity. In *Yearbook of Morphology 2001*, ed. Geert Booij and Jaap van Marle, 203–235. Berlin: Springer.

Heeringa, Wilbert. 2004. *Measuring Dialect Pronunciation Differences Using Levenshtein Distance*. Groningen: University of Groningen dissertation.

Hofstadter, Douglas, and Emmanuel Sander. 2013. *Surfaces and Essences: Analogy As the Fuel and Fire of Thinking*. New York: Basic Books.

Hulk, Aafke, and Natascha Müller. 2000. Bilingual First Language Acquisition at the Interface Between Syntax and Pragmatics. *Bilingualism: Language and Cognition* 3 (3): 227–244.

Krott, Andrea, Harald Baayen, and Robert Schreuder. 2001. Analogy in Morphology: Modeling the Choice of Linking Morphemes in Dutch. *Linguistics* 39: 51–93.

Labov, William, Sharon Ash, Maciej Baranowski, Naomi Nagy, Maya Ravindranath, and Tracey Weldon. 2006. Listeners' Sensitivity to the Frequency of Sociolinguistic Variables. *University of Pennsylvania Working Papers in Linguistics* 12 (2): 10. http://repository.upenn.edu/pwpl/vol12/iss2/10.

Lieberman, Erez, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A. Nowak. 2007. Quantifying the Evolutionary Dynamics of Language. *Nature* 449 (7163): 713–716. https://doi.org/10.1038/nature06137.

Masini, Francesca. 2019. Competition Between Morphological Words and Multiword Expressions. In *Competition in Inflection and Word-Formation*, ed. Franz Rainer, Francesco Gardani, Wolfgang U. Dressler, and Hans Christian Luschützky, 281–305. Cham: Springer.

Maslen, Robert J.C., Anna L. Theakson, Elena V.M. Lieven, and Michael Tomasello. 2004. A Dense Corpus Study of Past Tense and Plural Overregularization in English. *Journal of Speech, Language, and Hearing Research* 47(6): 1319–1333.

Paul, Hermann. 1909. *Prinzipien der Sprachgeschichte*. Vierte Auflage. Max Niemeyer: Halle a.S.

Rainer, Franz. 2012. Morphological Metaphysics: Virtual, Potential, and Actual Words. *Word Structure* 5 (2): 165–182. https://doi.org/10.3366/word.2012.0025.

———. 2016. Blocking. *Oxford Research Encyclopedia of Linguistics*. http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-33.

Schuchardt, Hugo. 1885/1972. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Oppenheim. Reprint and translation in Theo Vennemann and Terence H. Wilbur (eds) (1972). *Schuchardt, the Neogrammarians, and the Transformational Theory of Phonological Change,* 1–72. Frankfurt: Athenäum.

Skousen, Royal. 1989. *Analogical Modeling of Language*. Dordrecht/Boston: Kluwer Academic Publishers.

———. 2005. Analogical Modelling. In *Quantitative Linguistics: An International Handbook*, ed. Reinhard Köhler, Gabriel Altmann, and Rajmund Piotrowski, 705–716. Berlin: Walter de Gruyter.

Skousen, Royal, Deryle Lonsdale, and Dilworth B. Parkinson. 2002. *Analogical Modeling: An Exemplar-Based Approach to Language*. Amsterdam/Philadelphia: John Benjamins.

Smits, Erica, Heike Martensen, Ton Dijkstra, and Dominiek Sandra. 2006. Naming Interlingual Homographs: Variable Competition and the Role of the Decision System. *Bilingualism: Language and Cognition* 9 (3): 281–297.

Smits, Erica, Dominiek Sandra, Heike Martensen, and Ton Dijkstra. 2009. Phonological Inconsistency in Word Naming: Determinants of the Interference Effect Between Languages. *Bilingualism: Language and Cognition* 12 (1): 23–39.

Strik, Oscar, and Arjen Versloot. 2015. Frequency and Analogy in Early Modern Frisian Verb Inflection. In *Modelling Analogical Change. A History of Swedish and Frisian Verb Inflection*, Groningen Dissertations in Linguistics 137, ed. Oscar Strik, 165–177. Groningen: University of Groningen.

Tiersma, Piter Meijes. 1979. *Aspects of the Phonology of Frisian Based on the Language of Grou*, Meidielingen Stúdzjerjochting Frysk Oan de Frije Universiteit Yn Amsterdam 4. Amsterdam: Vrije Universiteit, Studierichting Fries.

van Heuven, Walter, Emily Coderre, Taomei Guo and Ton Dijkstra. 2011. The Influence of Cross-Language Similarity on Within- and Between-Language Stroop Effects in Trilinguals. *Frontiers in COGNITION* 2: 374. https://doi.org/10.3389/fpsyg.2011.00374 (26 March, 2012).

Versloot, Arjen Pieter, and Elżbieta Adamczyk. 2018. A Multivariate Analysis of Morphological Variation in Plural Inflection in North Sea Germanic Languages. In *Reorganising Grammatical Variation. Diachronic Studies in the Retention, Redistribution and Refunctionalisation of Linguistic Variants*, ed. Antje Dammel, Matthias Eitelmann, and Mirjam Schmuck, 17–55. Amsterdam/Philadelphia: John Benjamins.

Versloot, Arjen, and Eric Hoekstra. 2016. Attraction Between Words as a Function of Frequency and Representational Distance: Words in the Bilingual Brain. *Linguistics* 54 (6): 1223–1240.

Ytsma, Johannes. 1995. *Frisian as First and Second Language. Sociolinguistic and Socio-psychological Aspects of the Acquisition of Frisian Among Frisian and Dutch Primary School Children.* Ljouwert: Fryske Akademy. PhD-thesis, University Tilburg.

Yule, George Udny. 1925. A Mathematical Theory of Evolution: Based on the Conclusions of Dr. J.C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London*, 21–87. Series B, Containing Papers of a Biological Character 213.