

Diachronic dynamics of Middle English phonotactics provide evidence for analogy effects among lexical and morphonotactic consonant clusters

ANDREAS BAUMANN
University of Vienna

NIKOLAUS RITT
University of Vienna

CHRISTINA PRÖMER
University of Vienna

Abstract

Consonant clusters that rarely occur lexically (i.e. within morphemes) may function as complexity markers when they span a morpheme boundary, i.e. when they occur morphonotactically. In this study we observe patterns in the diachronic dynamics of Middle English which hint at mutually beneficial effects between morphonotactic and lexical clusters. We suggest that the patterns revealed can be explained by frequency-based analogy effects in language acquisition.

1 Introduction

On the basis of diachronic corpus evidence from Middle and Early Modern English, this paper studies interactions between word-final consonant clusters that occur within morphemes, like /nd/ in *hand*, /lb/ in *bulb*, or /st/ in *fast+est*, and those that span morpheme boundaries, like /nd/ in *quicken+ed* or /mz/ in *seem+s*. The former are by definition phonotactically licensed, and, following in this respect Dressler and Dziubalska-Kořaczyk (2006), we refer to them as ‘phonotactic’ or ‘lexical’. The latter are referred to as ‘morphonotactic clusters’. They may be phonotactically licensed as well, but often they are not. For instance, the /nd/ in *quicken+ed* is, while the /mz/ in *seem+s* is not, because the latter does not occur morpheme-internally.

In English, as in many languages, the sets of lexical and morphonotactic clusters are not identical. This is to be expected, since

phonotactic constraints are known to be tightest at the stem level, i.e. morpheme-internally (Kiparsky 1982; Giegerich 1999; McMahon 2002), and since consonant clusters in general count as phonologically marked, or dispreferred, they are rare within morphemes (Shockey; Berent et al. 2007; Dziubalska-Kołączyk & Zydorowicz 2014). Even when they are not permitted within morphemes, however, they may be produced through morphological or syntactic concatenation, and by virtue of being ruled out morpheme-internally, such clusters then have the potential of signaling syntactic (McQueen 1998) or morphological boundaries (Post et al. 2008; Dressler et al. 2010). Thereby, they serve an important function in the decomposition of speech into meaningful units, and it may be for this reason that they have become stably established at the word or phrase level.

Although the sets of lexical and morphotactic clusters are not identical they often overlap. This is the case, for example, in Polish, French, German, and also in English. Thus, final /nd/, which represents a morphotactic cluster in *quicken+ed*, occurs also morpheme-internally in words like *hand* or *wind*. It is these clusters that our study focuses on.

Specifically, we ask whether—and under what conditions—morphotactic clusters inhibit or promote the emergence of homophonous lexical counterparts. The question is motivated by the following considerations. On the one hand, and as argued by Dressler & Dziubalska-Kołączyk (2006), the signaling function of morphotactic clusters is clearly diminished when there are lexical homophones, so that clusters that span, and thereby indicate, morpheme boundaries should inhibit the emergence of lexical clusters. On the other hand, however, children may acquire highly frequent morphotactic clusters before they recognize the morphological boundaries they signal (Juszyk et al. 2002), which may loosen the constraints that prohibit such clusters within morphemes and thereby promote, rather than inhibit, the establishment of lexical homophones.

We address this question through a quantitative corpus study, in which we chart the development of word-final lexical and morphotactic clusters in Middle English (ME) and Early Modern English (EModE). We show that the pattern in the diachronic dynamics in ME provides evidence of analogy effects by which morphotactic clusters promote rather than inhibit the establishment of lexical homophones.

The paper is structured as follows. Section 2 reviews the relevant aspects of morphotactic theory (2.1), focusing on the diverse relationships that may be established between morphotactic and lexical clusters, and elaborating the research question (2.2). Section 3

introduces the data (3.1), presents an outline of the quantitative approach (3.2), and introduces the analysis and our findings (3.3). Finally, the results are discussed and summarized in the concluding section (4).

2 Consonant clusters, morphotactics, and analogy

2.1 Phonotactic and morphotactic consonant clusters

2.1.1 Inhibitory effects among consonant clusters

As outlined above, morphotactic clusters, which span morpheme boundaries, can signal these boundaries by virtue of their markedness. Clearly, this works best when clusters do not at the same time also occur within morphemes. A good example is ModE /md/. It occurs only when the suffix *-ed* is added to a stem ending in /m/, as in *seem-ed*, thereby facilitating the decomposition of past tense verbs or past participles. When morphotactic clusters have lexical homophones, however, their facilitating effect is diminished (Dressler & Dziubalska-Kořaczyk 2006; Dressler et al. 2010; Calderone et al. 2014). This is the case in English /nd/, which occurs not only in past-tense verbs or past participles, but also in numerous lexical base forms such as *hand*, *band*, *demand*, as well as in highly frequent function words like *behind* or *and*. Thus, clusters like /nd/ are very weak indicators of the morphological structure of the words in which they occur. Hence, lexical clusters would be expected to inhibit the establishment of morphotactic homophones and *vice versa*. Assuming that the inhibitory pressure a cluster exerts on its homophonous counterpart correlates with its frequency, lexical /nd/ should greatly inhibit morphotactic /nd/.

The general prediction that this hypothesis proposes is that cluster types should diachronically tend to become either purely morphotactic or purely phonotactic (Dressler et al. 2010). Such a scenario could come about via selective repair processes such as cluster reduction (cf. Labov 1989, who reports that final coronal deletion more frequently affects /nd/ clusters in simple items such as *find* than in complex forms such as *fine+d*), schwa epenthesis (rare in English, cf. the lexicalized adjective *learned*, /lɜ:nɪd/, but see Schlüter 2005), selective devoicing of /nd/ in past tense or participle forms (e.g. *learn+t*, *burn+t* < *learn+ed*, *burn+ed*), or theoretically also by the avoidance (and eventually the loss) of ambiguous word forms.

2.1.2 Supporting effects among consonant clusters

In addition, and to a certain degree in contrast, to the inhibiting effects outlined in the previous section, mutually supporting effects between morphotactic and lexical clusters have also been suggested. On the one hand, Martin (2007: 99) investigated consonant clusters that occur at the boundary of English noun-noun compounds and concludes that “the categorical phonotactic restrictions that hold within morphemes also hold gradiently across morpheme boundaries”. This provides evidence for a mutually supporting relationship between morphotactic and lexical clusters, since lexical clusters license the presence of their boundary-spanning counterparts. Note that Martin's argument concerns both lexical clusters and morphotactic clusters at constituent boundaries in compounds; the same relationship can be assumed to hold, even more so, between lexical clusters and morphotactic clusters in the prosodically weaker word-final position.

On the other hand, Hogg and McCully (1987: 47) investigated VVCC rhymes and state that “the type of syllable structure found in a word such as *wind* (/waind/) has been protected through analogy with inflected forms such as *weaned*”. Hence, they claim that morphologically produced word-final VVCC rhymes stabilize their lexical counterparts via analogy, thus providing support for the hypothesis that lexical clusters may also benefit from the presence of morphotactic clusters. Hogg and McCully (1987) focus on coda clusters following a long vowel, but in the remainder of this paper, their claim will be extended to coda clusters in general.

2.2 Elaborating the analogy hypothesis

We want to test whether morphotactically produced consonant clusters support their morpheme-internal counterparts. The hypothesized mechanism at work is frequency-based analogy, and the diagnostic method for detecting these analogy effects involves the analysis of the diachronic development of the consonant clusters in question.

2.2.1 Word-internal phonotactics and analogy

We suggest two reasons for analogical transfer from morphotactics to morpheme-internal phonotactics. First, morphotactic and lexical instances of a cluster type obviously share properties such as place or manner of articulation, voicing, or sonority of the respective consonants involved, although in certain articulatory or acoustic features, morphotactic and lexical clusters of a certain type might exhibit

slight differences. For instance, Plag et al. (2015) show that the acoustic duration of word-final /s/ in English is significantly longer if it is non-morphemic (i.e. if it does not represent an inflectional suffix or clitic). Thus, if /s/ is part of a word-final /Cs/ cluster, morphonotactic instances of that cluster are supposed to be shorter than their lexical counterparts. Nevertheless, structural similarity between instances of the two cluster categories should be substantial, so that on the discourse level the production and perception of tokens of one category is supposed to have a priming effect on the other category, along the lines of structural priming theory in syntax (Ferreira & Bock 2006; Pickering & Ferreira 2008). These effects on the level of discourse then facilitate the establishment of certain—in our case phonologically primed—patterns in grammatical knowledge (Gries 2005; Fehér et al. 2016). Clearly, when restricting oneself to phonological structure, this argument in principle goes in both directions: morphonotactic clusters have facilitating effects on lexical clusters, and *vice versa*. However, referring to a study by Shields and Balota (1991) about priming and duration, Jäger and Rosenbach (2008: 97) argue that phonetic priming is asymmetric in that “a phonetic full form has a stronger priming effect on the corresponding reduced form than the other way round”. This suggests that via priming, lexical clusters support their morphonotactic counterparts to a larger extent than the converse, since the former are phonologically less reduced (Plag et al. 2015).

We hypothesize a stronger version of the opposite direction, i.e. that morphonotactic clusters support lexical ones, and that this can be accounted for in terms of language acquisition. It is known that during the first two years of first-language acquisition, learners acquire highly frequent inflected word forms as lexical chunks (cf. Brown 1973; Rumelhart & McClelland 1986). This entails that highly token-frequent, and specifically morphonotactic clusters are acquired before the morphological operations that actually produce them in adult speech (the very same mechanism has been suggested to drive the lexicalization of words (Brinton & Traugott 2005: 91–95)). Crucially, during this first stage these items that were originally produced as morphonotactic clusters are processed as lexical clusters by the learner, which would logically facilitate the acquisition of words containing actual lexical clusters of the same cluster type. These acquired words would, in all likelihood, not be ‘unlearned’ after the onset of the acquisition of morphology, resulting in lexical clusters surfacing more frequently in simplex words. This would entail that morphonotactic clusters promote the acquisition of their lexical counterparts, and this supporting effect is expected to be larger, the more token-frequent the

morphotactic clusters are. We summarize these thoughts in the following hypothesis:

- (1) **Analogy among consonant clusters.** Morphotactic clusters and lexical clusters of the same cluster type mutually support each other via analogy.

2.2.2 Diachronic reflexes of frequency-based analogy effects

If highly token-frequent morphotactic clusters promote the acquisition of words containing lexical clusters, then diachronically the number of instances of the lexical cluster should obviously increase. Impressionistically, this is evident from the developments of the word-final clusters /nd/ and /md/. Through morphological operations, the first one comes about roughly four times as often as the second one in ME. Lexically, /nd/ surfaces in many lexical items such as ME/PDE *and*, *fiend*, *behind*, *wind*, or OE/ME *kalend* ('(first day of a) month') and *healend* ('savior'), whereas /md/ occurs sporadically in items such as ME *fremd* ('foreign'). The crucial point is that while /nd/ even appears in more recently imported loans such as *defend* or *command*, thereby increasing in frequency, /md/ gradually lost its lexical use.²

In the following, this phenomenon will be investigated more systematically. The hypothesis to be tested in this paper thus reads as follows:

- (2) **Diachronic reflexes of analogy.** If the first stages of language acquisition feature analogical transfer from morphotactic to lexical clusters, then, as a diachronic reflex, the lexical counterparts of highly token-frequent morphotactic clusters should in the long run appear in many lexical items and hence become more lexical. Similarly, token-infrequent clusters are expected to become less lexical.

Conversely, albeit strictly speaking not logically accurate, we will see the appearance of such a diachronic reflex as indirect evidence for frequency-based analogy effects between morphotactic and lexical consonant clusters, i.e. hypothesis 1. According to the above hypothesis,

² Clearly, importing loans or disfavoring particular words is not the only way in which a cluster can become more or less acceptable in lexical items. Similarly, fusion (e.g. *whence* < *whenne+s*) or phonological change (e.g. cluster reduction in *bomb*, *knight* or *damn*) provide other sources of variability.

analogy may very well interfere with the pressure of decreasing cluster ambiguity as outlined above. This is the case if a primarily morphonotactic cluster which is also very frequent in terms of morphonotactic tokens by analogy increases in lexical items, and thus becomes more ambiguous.

3 Detecting diachronic reflexes of analogy effects

In this section, we will explain how we tested the previously stated hypothesis that morphonotactically token-frequent clusters should become more lexical by investigating the diachronic development of the ME inventory of word-final consonant clusters. It is structured as follows: first, the ME data are introduced (3.1), then the hypothesis is operationalized in order to investigate it statistically (0), and finally, the data are analyzed and interpreted by means of two modeling approaches (1.1).

3.1 Data description

The dataset used for this study consists of word-final sequences extracted from the Penn Helsinki Corpora of Middle English (PPCME2, Kroch & Taylor 2000) and Early Modern English (PPCEME, Kroch et al. 2004). The compilation dates of the texts included range from 1138 to 1698. All words ending in graphemic C(V)C(V) sequences were extracted with the exception of words labeled as foreign (i.e. cases of code-switching); however, for the present study, only those words which end in a consonant cluster are of interest. Hence, excluded from the data set were sequences for which there is evidence that at least one of the two vowels did not get reduced (such as e.g. *plenty*) and words that are already monosyllabic (e.g. *for*). All other potential clusters were labeled as 'morphonotactic' (e.g. *bann+ed*), 'lexical' (e.g. *hand*) or 'weakly morphonotactic'. The latter intermediate category consists of cases like *concept*, which are not morphonotactically transparent (for etymological reasons, for example), but which might feature an inflectional or derivational operation. Cases labeled as weakly morphonotactic were excluded from the dataset for the following analyses.

In total, 314,158 potential final consonant cluster tokens were included in the dataset, of which 206,427 are lexical and 82,384 are morphonotactic. For each token, the corresponding date (depending on the text it was extracted from) was recorded, and the data cover roughly six centuries. Due to the unequal distribution of texts across this time span, the whole range was divided into sub-periods of 50 years, starting with the period from 1100 to 1150. The short-hand

notation ‘1200’ represents the period from 1200 to 1250, ‘1250’ for the period from 1250 to 1300, etc. Table 1 shows the numbers of potential morphotactic and lexical consonant clusters for each period as well as the sizes of the respective sub-corpora. Due to the small number of word-final consonant clusters in the Early Middle English period and the fact that schwa-loss began to spread no earlier than the 12th century (Brunner 1984; Fisiak 1968; see also Section 3.2.1 below), the first three half-centuries (1100, 1150, 1200) were excluded from the analysis.

period	morphotactic	lexical	total count	sub-corpus size
1250	137 (26.3%)	384 (73.7%)	521 (2.7 pm)	192,086,758
1300	1,892 (25.4%)	5,559 (74.6%)	7451 (32.8 pm)	226,997,791
1350	10,786 (29.0%)	26,378 (71.0%)	37,164 (151.5 pm)	245,362,411
1400	15,409 (29.4%)	36,934 (70.6%)	52,343 (118.6 pm)	441,525,895
1450	13,889 (26.2%)	39,206 (73.8%)	53,095 (80.9 pm)	656,369,953
1500	6,356 (27.5%)	16,742 (72.5%)	23,098 (22.9 pm)	1,009,235,900
1550	6,963 (29.1%)	16,950 (70.9%)	23,913 (14.6 pm)	1,642,395,212
1600	6,525 (28.7%)	16,213 (71.3%)	22,738 (10.9 pm)	2,091,129,356
1650	9,153 (30.3%)	21,088 (69.7%)	30,241 (9.9 pm)	3,065,964,242

Table 1. Frequencies of potential word-final morphotactic and lexical consonant clusters in the half-centuries from 1250 to 1700 together with the sizes of the corresponding sub-corpora. Figures in brackets denote fractions of morphotactic and lexical sequences among all C(V)C(V) sequences (as %), and fractions of the latter sequences among the total number of words in the respective sub-corpora (per million words).

3.2 Operationalization of the hypothesis and its parameters

As described above, we are investigating the impact that the token frequency of morphotactic instances of a given cluster type has on the number of lexical items its phonotactic counterpart occurs in, i.e. the cluster’s lexibility, and we address the question of whether this impact changes diachronically. To this end, three variables have to be operationalized: (a) time, (b) morphotactic token frequency, and (c) lexibility. These variables will be described in the following section, before presenting a more formalized version of our hypothesis.

3.2.1 *Dramatis personae*: time, frequency and lexicality

The first variable, time, simply measures the discrete 50-year periods from 1200 to 1700. The second variable to be covered is morphonotactic token frequency. The goal is to obtain an estimate of the number of morphonotactic tokens of a particular cluster type that a listener was exposed to. To this end, the raw number of morphonotactic tokens ending in a sequence $/C(\ə)C(\ə)/$ was determined for each cluster type and for each text, where $/\ə/$ could be represented by any vowel grapheme. Since, particularly due to schwa-loss, the graphemic representation does not necessarily provide a reliable estimate of its phonological counterpart (think of the graphemic representation of the past tense suffix *-ed*, to name an obvious example), the frequency of the sequences with the above structure was adjusted probabilistically in order to attain more reliably frequencies of actual occurrence.

To explain this step in more detail, the process of schwa-loss in English is actually a combination of two deletion processes, one of which accounts for the loss of word-final schwa while the other deletes inter-consonantal checked schwa. The first process is believed to have initiated at the latest at around 1200 (Fisiak 1968: 36; Minkova 1991; Brunner 1984: 348) and finished no later than sometime in the 15th century (Dobson 1957: 879). The second process, i.e. the loss of checked schwa, started slightly later in the 14th century (Mossé 1991: 35) and was completed in nominal and verbal inflections at around 1600 (Dobson 1957: 883). Accordingly, for the purpose of this analysis, $t_0 = 1200$ and 1300 were taken as rough onset times and $t_1 = 1500$ and 1600 as rough offset times for the respective processes.

Phonological processes, such as schwa-loss, which act on a population of linguistic items often exhibit a sigmoid trajectory (Denison 2003; Wang & Minett 2005; Blythe & Croft 2012; cf. also Kroch 1989). Such a shape arises, in particular, if the process proceeds logistically, i.e. if its growth rate depends (a) on the amount of items that have already been affected by a change and (b) on the amount of items that have not yet been affected. Mathematically, logistic growth is modeled by the logistic function $p(t) = 1/(1 + e^{-c-rt})$, where $p(t)$ measures the proportion of items affected by the change at time t . Given the values of pairs $(t, p(t))$ for two different times t , the values of the constants c and r can be determined by making use of the logit transform of the above equation. Once c and r are known, the proportion p of affected items can be determined for any time t . For a single randomly drawn token at time t , $p(t)$ can now be interpreted as the probability of being affected by the change.

A separate logistic-spread process for each of the two sub-processes of schwa-loss was implemented. The respective proportions

are $p_{\text{final}}(t)$ and $p_{\text{checked}}(t)$. In order to determine the respective constants c_i and r_i (i standing for ‘final’ or ‘checked’), we made use of the above mentioned onset and offset times t_0 and t_1 and defined the onset proportion as $p_i(t_0) = .01$ and the offset proportion as $p_i(t_1) = .99$, i.e. 1% and 99% affected items, respectively. We assume that losing final schwa has no effect on the likelihood of losing interconsonantal schwa, and *vice versa*. In other words, the two sub-processes are regarded as independent. Thus, the probability of a word-final /C(ə)C(ə)/ sequence being a consonant cluster is at least $p_{\text{final}}(t) \times p_{\text{checked}}(t)$ at time t (Figure 1).

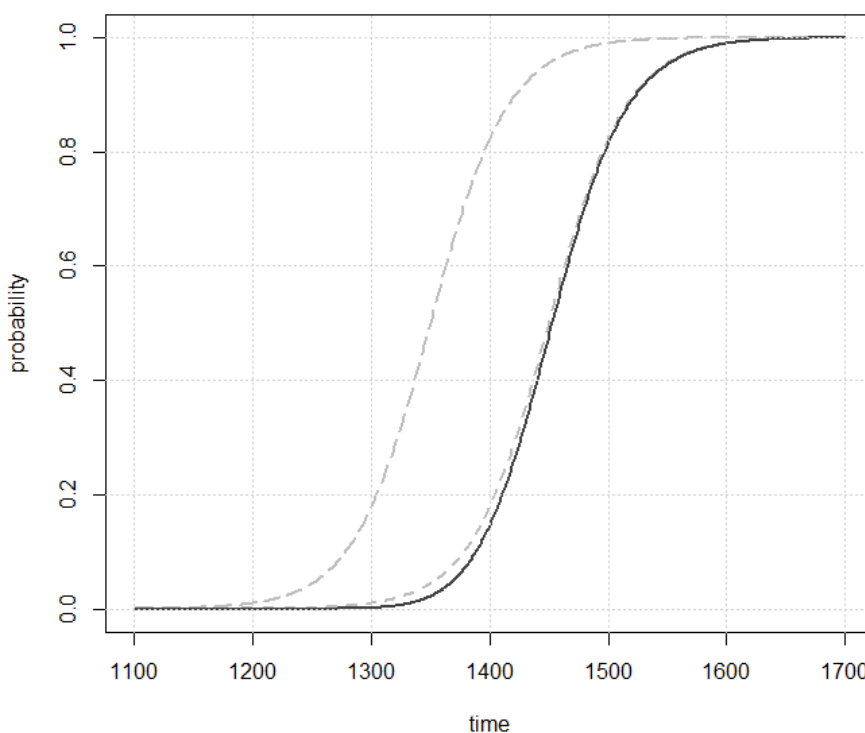


Figure 1. Spread of schwa-loss (dark gray) as an interacting process of the loss of final (light gray, long dashes) and checked schwas (light gray, short dashes). The vertical axis measures the probability of a final /CəCə/ sequence being affected by the change.

We take this product as a conservative estimate of the probability that an item /C(ə)C(ə)/ actually is a cluster, $p_{\text{cluster}}(t)$. It is crucial to note that this estimate constitutes a lower bound in the sense that in a /C(ə)C(ə)/ sequence, one of the two schwas might have been already lost or indeed may have never been present before the onset of schwa-loss. Thus, it ensures that the token frequencies in each period are not

underestimated. We assume that this provides us with a more reliable estimate than just resorting to the problematic graphemic representations, which would result in much lower frequencies of cluster tokens.

The period-wise frequencies of morphonotactic word-final consonant-cluster tokens were then calculated according to the following procedure. The raw token frequencies of the sequences $/C(\text{ə})C(\text{ə})/$ corresponding to a cluster type CC , were determined for each text. These raw frequencies were multiplicatively adjusted by the above-described probability $p_{\text{cluster}}(t)$, where t is the estimated date of the text (see Kroch et al. 2004; Kroch & Taylor 2000). For each cluster type and each period, these adjusted frequencies were summated and subsequently normalized with respect to the period-specific sample size (i.e. the total number of words in all texts that belong to the half-century period). The base of normalization was set at 1 million. The same adjustment and normalization procedure was applied to the lexical instances of the respective cluster types, and the resulting frequencies are denoted as φ_{mpt} (or *mpt.frequency*) and φ_{lex} (or *lex.frequency*), respectively.

Finally, the third variable, lexicality (denoted as λ), is intended to measure in how many instances a cluster type occurs lexically rather than morphonotactically. We simply define it as the fraction of lexical tokens among all instances of that cluster type in a given period, thus, $\lambda = \varphi_{\text{mpt}} / (\varphi_{\text{lex}} + \varphi_{\text{mpt}})$.³ Lexicality, computed in this way, has a straightforward and theoretically relevant interpretation. If for a cluster type the score is close to 0, it is morphonotactic, if it is close to 1, it is lexical, and if it is in-between, the cluster type is ambiguous with respect to its complexity-signaling function. Note that the lexicality scores of a given cluster type in a particular period can range from 0 to 1. Obviously, this score is always strictly smaller than 1, since the present analysis is restricted to cluster types with $\varphi_{\text{mpt}} > 0$ only. According to morphonotactic theory, cluster types are expected to disambiguate, hence approaching either 0 or 1 on the lexicality scale (cf. Section 2 and Dressler et al. 2010). Selecting proportional frequencies allows for a direct application of our findings to morphonotactic theory,

³ Note that the fact that φ_{mpt} appears in the definition of λ is unproblematic, since for large frequencies, the strength of a monotone relationship between morphonotactic frequency and lexicality provides a lower bound for the strength of a monotone relationship between φ_{mpt} and φ_{lex} .

so that we can test whether clusters, under certain circumstances, show the opposite dynamics.

The choice of (adjusted) token frequency over type frequency in the definition of lexicality can be attributed to the lack of lemmatized data in the corpora used (which is obviously a consequence of the spelling variation and inconsistency in ME and, to a lesser extent, EModE).

3.2.2 The changing morphotactic space and how it should evolve

In each period, the inventory of final morphotactic consonant clusters can now be conceptualized by means of the Cartesian product of φ_{mpt} and λ , which we will refer to as the *morphotactic space*. That is, for each cluster type, such as /nd/, /ns/ or /rn/, we determined, first, morphotactic frequency and, second, lexicality. The scores on these two variables determine the cluster's location in the morphotactic space. If this is done for all potentially morphotactic cluster types, the cloud of resulting points in this space constitutes the morphotactic cluster inventory in that period. Figure 2 below shows the cluster inventories in all the half-centuries from 1250 to 1700. In each scatterplot, the horizontal axis measures the morphotactic frequency while the vertical axis measures lexicality. The locations of the cluster types are represented by points in the plot.

For our hypothesis we want to show that those cluster types which are morphotactically frequent, in the long run appear in more and more lexical items, i.e. they should become more lexical. Thus, clusters that are located on the right of the scale should evolve in such a way that they also score high on the lexicality scale. In other words, we hypothesize that diachronically the cluster inventory establishes a positive monotone relationship between φ_{mpt} and λ .

The dynamic component in this hypothesis is crucial. It is not sufficient to show that at a given point in time there is such a monotone relationship. Rather, it has to be shown that frequent clusters evolve in such a way that they become integrated into lexical items via analogy effects. A synchronic view alone fails to shed light on this matter, since the development could just as well move in the opposite direction. In the present analysis, we exploit the fact that in Old English, word-final morphotactic clusters were quite rare, since inflectional endings were typically syllabic. Thus, schwa-loss has given rise to a completely novel inventory of clusters, in which changes, such as those analogy-

correlation (or more precisely: the period-specific correlation coefficients) between frequency and lexicality evolve diachronically (3.3.2).

3.3.1 Fitting a generalized additive model

We are interested in how the interaction between time and the frequency of morphotactic clusters affects the lexicality of a cluster: thus, multidimensional modelling of the dependence of lexicality on the other two variables is required. In this analysis, a generalized additive model (GAM) was selected. In conventional linear regression models, interactions between predictor variables result in multiplicative linear terms. This means, that if one predictor variable—say time, as in the present case—is held constant, the dependent variable (λ) is a linear function of the second variable (φ_{mpt}). However, we are not exclusively interested in linear relationships between the latter two variables. Instead, any monotone (decreasing or increasing) relationship between frequency and lexicality would be of interest according to our hypothesis. Hence, a more flexible modeling technique not restricted to linear dependencies is required, and GAMs fulfill these requirements. In a nutshell, GAMs are models which are composed of linear and nonlinear components (so-called ‘splines’), thus yielding smoothly curved (or ‘wiggly’) surfaces that fit to the data in a statistically satisfying way. GAMs have been used extensively in ecology and evolution, and more recently in linguistics (Wieling et al. 2011; Baayen 2013; Fruehwald 2015).

Before feeding the data into the model, some adjustments had to be made. The φ_{mpt} scores were first normalized with respect to the period-wise maximal scores. This was necessary, since due to the S-shaped spread of schwa-loss, frequency scores were concentrated close to 0 in the earlier periods. By normalizing the data with respect to the maximal scores, a more appropriate model of the relationship between frequency and lexicality was achieved. Second, as the φ_{mpt} scores were strongly skewed to the right (see Figure 2), they were Box-Cox transformed, i.e. put into a shape that resembles normally distributed data (Box & Cox 1964). Figure 3a displays the morphotactic space changing in time. Overall, the dynamics look rather complicated, so that in order to detect diachronic patterns, fitting a model to the data indeed might provide more insights.

In the GAM, lexicality is modeled as being related to the interaction of time and morphotactic frequency. The `mgcv` package in R (Wood 2006a; R Development Core Team 2013) enables us to include so-called *tensor product smooths* into a GAM. While a detailed explanation is not

relevant here, suffice it to note that tensor products provide a simple way of modeling interactions between predictor variables in a GAM (Wood 2006a, 2006b). The GAM computed from the data yields a significant intercept (at 0.224; $p < .0001$) and a tensor-product term ($p = .0079$; estimated $df = 7.597$), which means that the morphonotactic space indeed changes significantly over time rather than staying roughly the same.

In order to interpret the model, it has to be visualized. Figure 3b shows the surface defined by the GAM, in which the following three patterns can be observed. (a) Very rare clusters evolve from medially lexical to predominantly morphonotactic because the surface defined by the GAM heads downwards for low φ_{mpt} values. This fits the frequency-driven analogy effects part of our hypothesis. (b) In accordance with the same hypothesis, medially frequent clusters become more lexical. However, (c) morphonotactically highly frequent clusters evolve from a primarily lexical state to a slightly less lexical one. This contradicts our predictions, since these clusters would be expected to become even more lexical. Looking at the period-wise one-dimensional curves, which depict the dependence of lexicality on frequency by fixing a point on the time axis and moving along the grid on the smooth wiggly surface in the direction of φ_{mpt} , it can be seen that the one which comes closest to an increasing monotone relationship is attained somewhere around the 1450 period. After this date, intermediately frequent clusters (b) overtake highly frequent ones in terms of lexicality (c).

In summary, the data in the first part of the observed time span, which corresponds to a large share of the ME period, seems to provide evidence for frequency effects among morphonotactic and lexical consonant clusters (see hypothesis 2), while the later data do not. This contrast is too interesting to be ignored, and therefore, the following section outlines another approach, which allows for a systematic analysis of this antithetic behavior.

3.3.2 Analysis of the correlation-coefficient trajectory

In the second approach, all periods are dealt with separately. The aim is to investigate whether the monotone relationship between frequency and lexicality increased diachronically. To this end, the corresponding correlation coefficients were determined for each period. This allows the investigation of the trajectory of correlation coefficients, which should increase according to the hypothesis being tested.

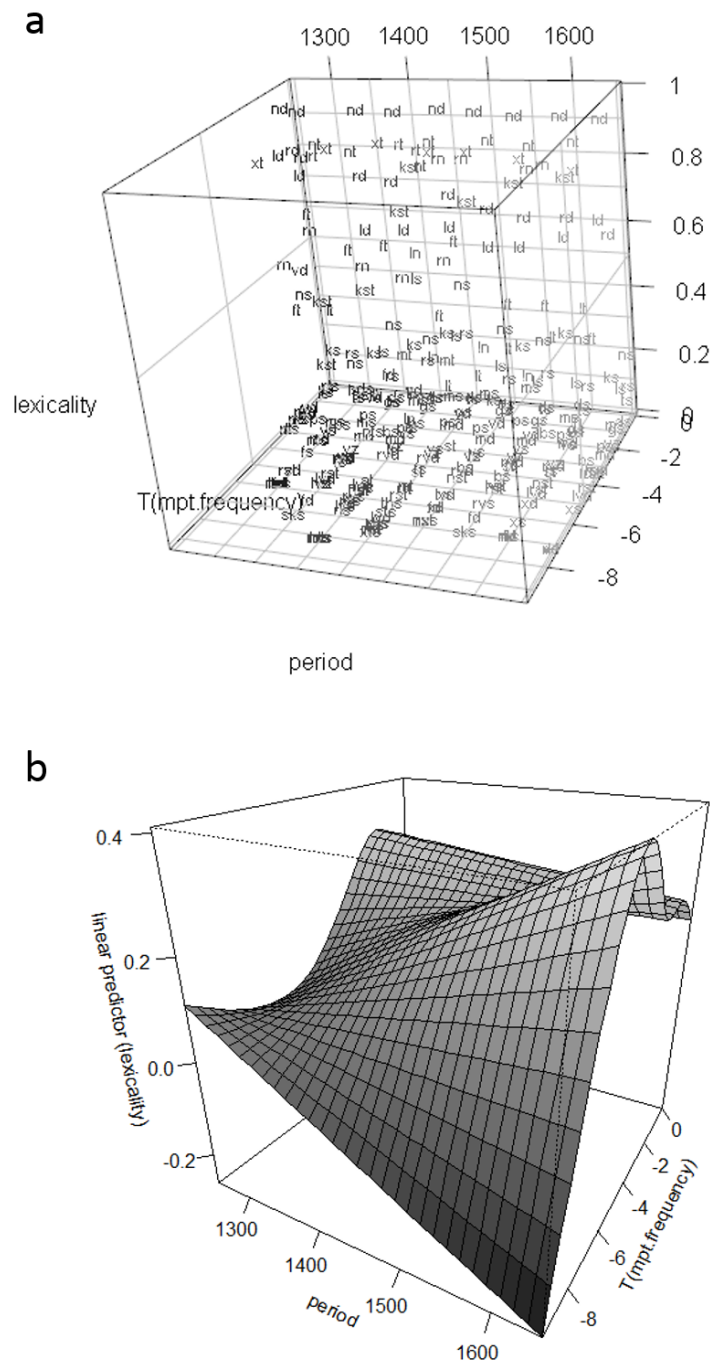


Figure 3. (a) Scatterplot showing the diachronic development of the morphotactic space, where φ_{mpt} is normalized with respect to the period-wise maximal scores and Box-Cox transformed to $T(\varphi_{\text{mpt}})$. (b) Plot of a generalized additive model fitted to the transformed data. The impact of the morphotactic token frequency on cluster lexicality is evidently changing over time.

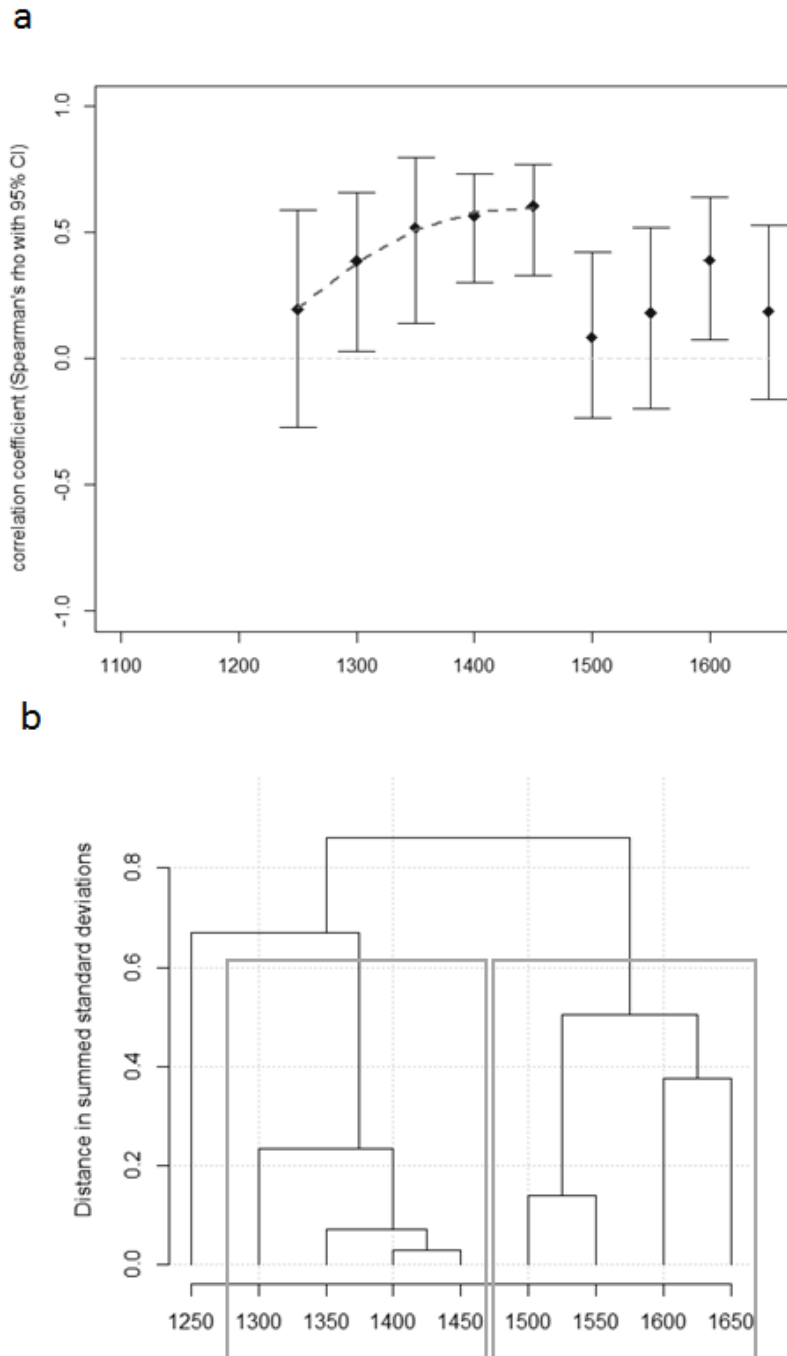


Figure 4. (a) Trajectory of estimated correlation coefficients (Spearman's ρ) together with 95% confidence intervals for all half-centuries from 1250 to 1700. The dashed line corresponds to a fitted quadratic model (adjusted $R^2 = .99$, $p = .004$). From 1250 to 1500 a significant positive correlation is established. (b) VNC based dendrogram of the successive correlation coefficients. Between 1450 and 1500 a break in the diachronic development is clearly observable.

For the present analysis, we selected Spearman's ρ as the correlation coefficient of φ_{mpt} and λ . This is motivated by the skewed distribution of the data and, more importantly, by the fact that we want the correlation measure to be sensitive to any monotone relationship. Due to the non-parametric nature of Spearman's ρ , the data were not transformed. Nine correlation coefficients ρ_i ($i = 1, 2, \dots, 9$) were determined ($t_i = 1250, 1300, \dots, 1650$; cluster-inventory sizes N_i ranging from 18 to 44). The trajectory of correlation coefficients together with the corresponding 95% confidence intervals (computed with the `RVAideMemoire` package in R) is shown in Figure 4a.

An inspection of Figure 4a clearly shows that, as expected from the results of the previous analysis, the correlation between φ_{mpt} and λ increases until the end of the 15th century, not significantly at first, but then reaching significance and approximating a strong correlation of $\rho \cong 0.5$ (Cohen 1992). However, in the 1500 period, the correlation drops close to zero and becomes non-significant again. Apart from the 1600 period, which exhibits a significant relationship again, this trend stays the same from then on.

Looking at the confidence intervals alone, which, crucially, do overlap when looking at the 1450 and 1500 periods, we cannot confidently claim that the data show the existence of two substantially different periods (before and after 1500). Hence, a clustering technique, variability-based neighbor-clustering, was employed which allows for the identification of stages in sequential data.

Variability-based neighbor clustering (VNC, Gries & Hilpert 2008) is a hierarchical clustering method which has the advantage of keeping a fixed ordering of the leaves of the hierarchy tree, because the clustering proceeds in such a way that only the direct neighbors in a previously defined sequence—here, successive time periods—are eligible for clustering on the various hierarchical levels. Hence VNC provides an excellent method to detect sets of similarly behaving periods in diachronic developments.

Figure 4b shows the dendrogram⁵ which results from the application of VNC to the trajectory of correlation coefficients ρ_i . It clearly divides the observation period into two stages, one corresponding to the periods before 1500 and one to the periods after 1500. Hence, it can be concluded that from 1250 to 1500, the cluster

⁵ VNC computations were done in R (version 3.0.2) with a script written by Stefan Gries and Martin Hilpert (see http://global.oup.com/us/companion.websites/fdscontent/uscompanion/us/static/companion.websites/nevalainen/Gries-Hilpert_web_final/vnc.individual.html; accessed 16.02.2016).

inventory evolved as hypothesized. After the break in 1500, the development does not show a clear pattern.⁶

3.3.3 Interpretation of the break: the transition from LME to EModE

Both analyses, the GAM as well as the VNC analysis of the correlation coefficients, suggest that at the end of the ME period, the dynamics of the cluster inventory showed a substantial change. Figure 5 shows the cluster inventories from these two periods overlaid in the same plot. Morphonotactic frequency was normalized with respect to the maximum score in each respective period. In order to identify the clusters that behave differently in the two periods, the cluster locations are labeled by the respective phonological representation (light gray indicates 1450 data, and dark gray represents the 1500 data). For the sake of illustration, linear regression lines were included, although due to the distributional properties of the data, they should be treated with caution. The regression lines are nevertheless helpful for identifying the cluster types which are responsible for the change in the correlation coefficient.

Two sets of cluster types seem to be particularly involved in the changing behavior: /Cs/ clusters, which, contrary to our predictions, become more frequent and less lexical, and /Cn/ clusters, which become slightly less frequent and more lexical (locations indicated by circles).

In ME, instances of the /Cs/ group occur as verbal present tense inflections (Northern dialects, Horobin & Smith 2002: 117), as well as nominal plural and genitive forms (all dialects). At the end of the 15th century, the inflectional competitors of the *-(e)s* ending (*-eth* and contracted *-t*) were ousted, partially due to migration and, in the end, standardization processes which took place at this time (cf. Horobin & Smith 2002), so that *-(e)s* became the default choice for expressing 3rd person singular in verbs and plural in nouns. Thus, morphonotactic final /Cs/ clusters became more frequent.

⁶ One might wonder, at this point, why a configuration of morphonotactic and lexical clusters which is so dispreferred that it needs to be repaired by processes operating during first language acquisition, as we intended to demonstrate by the development in the pre-1500 era, came about at all. We propose the following quite straightforward answer: only when schwa-loss had produced a reasonably large number of consonant clusters was it possible for analogy to function in language acquisition. We assume that this must have been at around 1200.

In contrast, word-final /Cn/ clusters went through a completely different development. The ME *-e(n)* suffix played a substantial role in the inflectional morphology of nouns (as a plural suffix) and verbs (as a subjunctive and partially indicative plural suffix, and as an infinitival suffix). In the transition from the ME to the EModE period, this suffix began to become less productive, and eventually dropped out of inflectional morphology. As a consequence, morphotactic word-final /Cn/ clusters became less frequent and more lexical.

In summary, the above observations show that both sets of clusters exhibit a development that is not in line with the proposed monotonously increasing relationship between morphotactic frequency and lexicality.

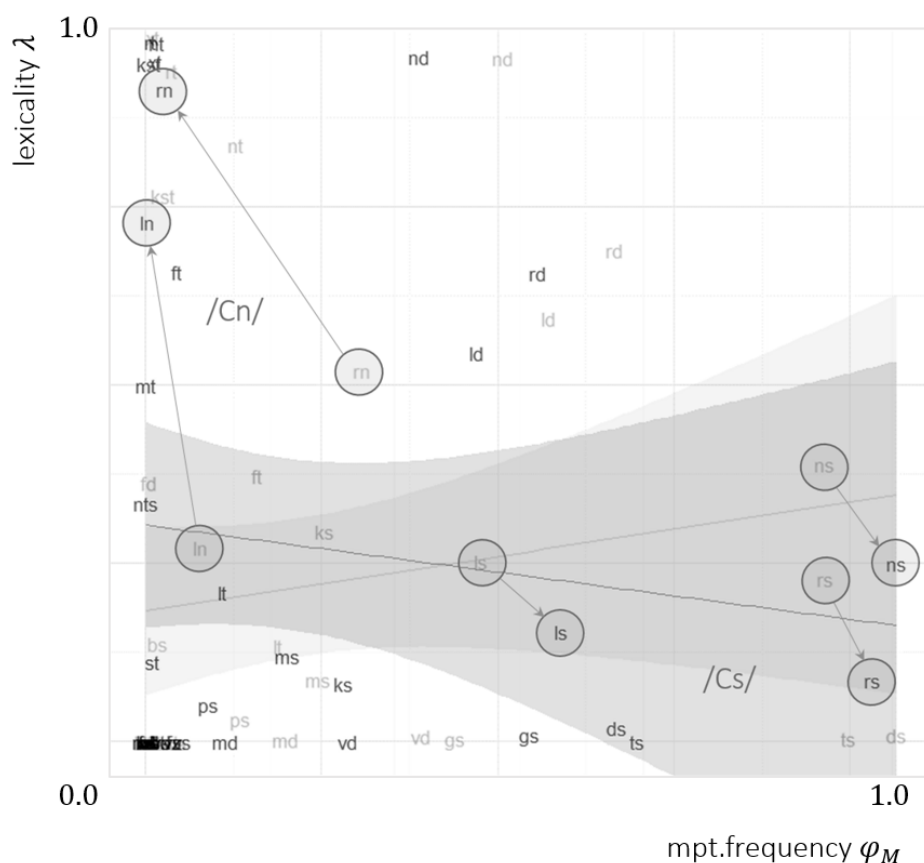


Figure 5. Plot of the superimposed morphotactic spaces of the cluster inventories in the successive periods 1450–1500 (light gray labels) and 1500–1550 (dark gray labels), respectively. The horizontal axis was scaled in such a way that the respective maximal and minimal scores in both periods coincide. Linear regression models (solid lines) together with 95% confidence areas (gray) were added and illustrate a positive correlation in the 1450–1500 period in contrast to a (not significant) negative correlation in the 1500–1550 period. Circles indicate morphosyntactically relevant dynamics corresponding to a set of /Cn/ and /Cs/ clusters.

The crucial point is that these developments are driven by morphosyntactic and sociolinguistic factors, and hence are not phonologically or phonotactically conditioned. We conclude that the period before 1500 provides corroborating evidence for the hypothesis about analogy effects. However, the period after 1500 does not yield a clear picture. Indeed, more recent data (i.e. ModE after 1700) is needed in order to satisfactorily address the question at hand.

4 Conclusion

At the outset of this paper we put forth the question of whether morphonotactic consonant clusters provide supporting effects for lexical instances of the same cluster type via analogy. This hypothesis was motivated by observations and conjectures found in the (mor)phonotactic literature about the inhibiting and supporting effects among the two cluster categories. We hypothesized that frequency effects in the first stages of language acquisition could give rise to these supporting effects, and that as a diachronic reflex of these effects, characteristic diachronic patterns were proposed to be observable. More specifically, we expected morphonotactically token-frequent clusters to become more lexically present (see hypothesis 2).

The latter claim was formalized in the following fashion: diachronically, a positive monotone relationship between the morphonotactic token frequency and the lexicality of the consonant clusters in the inventory of the language was expected to establish itself. Data from the ME period were used to test this hypothesis quantitatively by means of two different modeling approaches. Using ME data for addressing the research question at hand suggested itself, since through schwa-loss a completely new set of consonant clusters was created, so that the diachronic reflexes of the hypothesized analogy effects should be clearly observable.

In the first modeling approach, a generalized additive model (GAM) was fit to the data. It showed that indeed in the first part of the ME period, the relationship between morphonotactic frequency and lexicality evolved as expected, but that later on and contrary to expectation, intermediately frequent clusters became more lexical than highly frequent clusters. The second modeling approach provided a detailed look at when the change in the behavior of the cluster inventory took place. It was shown that before 1500 the inventory behaved as expected under the assumption of frequency-driven analogy in language acquisition, while after 1500 no particular pattern could be observed. Looking at the clusters involved in this change, we showed that the shifts in frequency or lexicality in a number of cluster types can be attributed to morphosyntactic or sociolinguistic, and thus

phonology-external, changes. Although an investigation of the development after 1500 would naturally be interesting, a systematic survey of this period exceeds the scope of our data. Hence, we can conclude that at least before the onset of inflectional reduction and standardization, the diachronic dynamics of the ME coda-cluster inventory suggest a supporting relationship between morphotactic and lexical clusters (see hypothesis 1).

This has interesting implications for morphotactic theory. One of the major claims about morphotactic consonant clusters is that their functionality in terms of signaling morpheme boundaries is diminished by the presence of structurally similar lexical clusters, so that there is an inhibitory relationship between the two cluster categories, as outlined in Section 2. A corollary of this is that cluster types should disambiguate so that they become either purely morphotactic or purely lexical. The findings from the present study add two novel aspects to the expected diachronic dynamics of consonant clusters.

First, it can be specified which clusters should become more lexical and which ones should become more morphotactic. According to our findings it should be—somewhat counterintuitively—the morphotactically *highly* token-frequent clusters that evolve towards the lexical boundary, while their low-frequency counterparts are expected to evolve into the less lexical, i.e. morphotactic direction.

Second, the findings in this study bolster the evidence for the supporting rather than inhibiting effects between morphotactic and lexical clusters, so that a diachronic development towards more ambiguous configurations would be expected. In this sense, two opposing forces are at work in the diachronic dynamics of consonant clusters, one which favors unambiguous clusters and a second one which favors ambiguous ones. The—doubtless language-specific—nature of the interaction between these two forces, however, still remains to be explored in further studies.

Comments invited

PiHP relies on post-publication review of the papers that it publishes. If you have any comments on this piece, please add them to its comments site. You are encouraged to consult this site after reading the paper, as there may be comments from other readers there, and replies from the author. This paper's site is here:

<http://dx.doi.org/10.2218/pihph.1.2016.1693>

Acknowledgements

We would like to thank the participants of the 3rd International Workshop on Phonotactics and Phonotactic Modeling (Vienna, November 2015), in particular Basilio Calderone, for many helpful comments on the present project, Kamil Kaźmierski, Ines Nitsche, Paula Orzechowska, and Paulina Zydrowicz for helping us with analyzing the historical data, as well as Elnora ten Wolde and Klaus Hofmann for their support. This research was supported by the Austrian Science Fund (FWF, grant No. P27592-G18) and the Österreichische Austauschdienst (OeAD, grant No. PL8/2014).

First Author contact details

Andreas Baumann
 Department of English and American Studies
 University of Vienna
 Spitalgasse 2-4, Hof 8.3
 A-1090 Vienna, Austria
andreas.baumann@univie.ac.at

References

- Baayen, Harald. 2013. Multivariate Statistics. In Robert Podesva & Devyani Sharma (eds.), *Research Methods in Linguistics*, 337–372. Cambridge: Cambridge University Press.
- Berent, Iris, Donca Steriade, Tracy Lennertz & Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104(3). 591–630.
- Blythe, Richard A. & William Croft. 2012. S-curves and the mechanism of propagation in language change. *Language* 88(2). 269–304.
- Box, George & David Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society Series B* 26(2). 211–252.
- Brinton, Laurel J. & Elizabeth C. Traugott. 2005. *Lexicalization and language change* (Research surveys in linguistics). Cambridge, UK, New York: Cambridge University Press.
- Brown, Roger. 1973. *A first language: The early stages*. Harvard: Harvard University Press.
- Brunner, Karl. 1984. *Die englische Sprache: Ihre geschichtliche Entwicklung* (Sammlung kurzer Grammatiken germanischer Dialekte. B, Ergänzungsreihe Nr. 6), 2nd edn. Tübingen: Niemeyer.

- Calderone, Basilio, Chiara. Celata, Katharina Korecky-Kröll & Wolfgang U. Dressler. 2014. A computational approach to morphotactics: evidence from German. *Language Sciences* 46. 59–70.
- Cohen, Jacob. 1992. A power primer. *Psychological Bulletin* 112. 155–159.
- Denison, David. 2003. Log(ist)ic and simplistic S-curves. In Raymond Hickey (ed.), *Motives for Language Change*, 54–70. Cambridge: Cambridge University Press.
- Dobson, Eric J. 1957. *English Pronunciation 1500-1700: Volume II: Phonology*. Oxford: Clarendon Press.
- Dressler, Wolfgang U. & Katarzyna Dziubalska-Kołodziejczyk. 2006. Proposing Morphotactics. *Wiener Linguistische Gazette* 73. 69–87.
- Dressler, Wolfgang U. Katarzyna Dziubalska-Kołodziejczyk & Lina Pestal. 2010. Change and variation in morphotactics. *Folia Linguistica Historica* 31. 51–68.
- Dziubalska-Kołodziejczyk, Katarzyna & Paulina Zygorowicz. 2014. The production of high-frequency clusters by native and non-native users of Polish. *Concordia Working Papers in Applied Linguistics* 5. 130–144.
- Fehér, Olga, Elizabeth Wonnacott & Kenny Smith. 2016. Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language*.
- Ferreira, Victor S. & Kathryn Bock. 2006. The functions of structural priming. *Language and Cognitive Processes* 21(7-8). 1011–1029.
- Fisiak, Jacek. 1968. *A Short Grammar of Middle English: Part One: Graphemics, Phonemics and Morphemics*. Warsaw: Państwowe Wydawnictwo Naukowe.
- Fruehwald, Josef. 2015. *Generations, lifespans, and the zeitgeist*. The Second Edinburgh Symposium on Historical Phonology. Edinburgh.
- Giegerich, Heinz. 1999. *Lexical strata in English: Morphological causes, phonological effects*. Cambridge: Cambridge University Press.
- Gries, Stefan T. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34(4). 365–399.
- Gries, Stefan T. & Martin Hilpert. 2008. The identification of stages in diachronic data: variability-based neighbour clustering. *Corpora* 3(1). 59–81.
- Hogg, Richard M. & Christopher B. McCully. 1987. *Metrical Phonology: A Course Book*. Cambridge: Cambridge University Press.
- Horobin, Simon & Jeremy J. Smith. 2002. *An Introduction to Middle English*. Oxford: Oxford University Press.

- Jäger, Gerhard & Anette Rosenbach. 2008. Priming and unidirectional language change. *Theoretical Linguistics* 34(2). 85–113.
- Juszyk, Peter W. Paul Smolensky & Theresa Allocco. 2002. How English-Learning Infants Respond to Markedness and Faithfulness Constraints. *Language Acquisition* 10. 31–73.
- Kiparsky, Paul. 1982. Lexical phonology and morphology. In Ik-Hwan Lee (ed.), *Linguistics in the morning calm*, 1st edn. 3–91. Seoul: Hanshin Publ.
- Kroch, Anthony. 1989. Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change* 1. 199–244.
- Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. *Penn-Helsinki Parsed Corpus of Early Modern English*. <http://www.ling.upenn.edu/hist-corpora/>.
- Kroch, Anthony & Ann Taylor. 2000. *Penn-Helsinki Parsed Corpus of Middle English*. <http://www.ling.upenn.edu/hist-corpora/>.
- Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1. 85–97.
- Martin, Andrew. 2007. *The evolving lexicon*. Los Angeles: UCLA PhD dissertation.
- McMahon, April M. S. 2002. *An Introduction to English Phonology* : Edinburgh University Press.
- McQueen, James. 1998. Segmentation of Continuous Speech Using Phonotactics. *Journal of Memory and Language* 39. 21–46.
- Minkova, Donka. 1991. *The history of final vowels in English: The sound of muting* (Topics in English linguistics 4). Berlin, New York: M. de Gruyter.
- Mossé, Fernand. 1991. *A handbook of Middle English*, 10th edn. Baltimore, Md.: Johns Hopkins University Press.
- Pickering, Martin J. & Victor S. Ferreira. 2008. Structural priming: a critical review. *Psychological Bulletin* 134(3). 427–459.
- Plag, Ingo, Julia Homann & Gero Kunter. 2015. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*. 1–36.
- Post, Brechtje, William D. Marslen-Wilson, Billi Randall & Lorraine K. Tyler. 2008. The processing of English regular inflections: Phonological cues to morphological structure. *Cognition* 109(1). 1–17.
- R Development Core Team. 2013. *R: A Language and Environment for Statistical Computing: A language and environment for statistical*. Vienna, Austria: R Foundation for Statistical Computing.

- Rumelhart, David E. & James L. McClelland. 1986. On learning the past tenses of English verbs. In PDP Research Group (ed.), *Parallel distributed processing: explorations in the microstructures of cognition*, 216–271. Cambridge, MA: MIT Press.
- Schlüter, Julia. 2005. *Rhythmic grammar: The influence of rhythm on grammatical variation and change in English* (Topics in English linguistics 46). Berlin, New York: Mouton de Gruyter.
- Shields, Lynne W. & David A. Balota. 1991. Repetition and associative context effects in speech production. *Language and Speech* 34(1). Shockey.
- Wang, William & James Minett. 2005. The invasion of language: emergence, change and death. *Trends in Ecology and Evolution* 20. 263–269.
- Wieling, Martijn, John Nerbonne & Harald Baayen. 2011. Quantitative Social Dialectology: Explaining Linguistic Variation Geographically and Socially. *PloS one*.
- Wood, Simon. 2006a. *Generalized Additive Models: an introduction with R*. Boca Raton, Florida: Chapman & Hall/CRC.
- Wood, Simon. 2006b. Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models. *Biometrics* 62. 1025–1036.