Cognitive influences in language evolution:

Psycholinguistic predictors of loan word borrowing

Padraic Monaghan[1,2] & Seán Roberts[3]

[1]Department of Psychology, Lancaster University, UK

[2]Max Planck Institute for Psycholinguistics, Netherlands

[3]excd.lab, Department of Anthropology and Archaeology, University of Bristol, UK

Correspondence to:

Padraic Monaghan

Department of Psychology

Lancaster University

Lancaster LA1 4YF

UK

Tel: +44 1524 593813

Fax: +44 1524 593744

Email: p.monaghan@lancaster.ac.uk

Abstract

Languages change due to social, cultural, and cognitive influences. In this paper, we provide an assessment of these cognitive influences on diachronic change in the vocabulary. Previously, tests of stability and change of vocabulary items have been conducted on small sets of words where diachronic change is imputed from cladistics studies. Here, we show for a substantially larger set of words that stability and change in terms of documented borrowings of words into English and into Dutch can be predicted by psycholinguistic properties of words that reflect their representational fidelity. We found that grammatical category, word length, age of acquisition, and frequency predict borrowing rates, but frequency has a non-linear relationship. Frequency correlates negatively with probability of borrowing for high-frequency words, but positively for low-frequency words. This borrowing evidence documents recent, observable diachronic change in the vocabulary enabling us to distinguish between change associated with transmission during language acquisition and change due to innovations by proficient speakers.

Key-words: language evolution; vocabulary change; loan-words; frequency; age of acquisition; word length; language acquisition; language incrementation

Cognitive Influences in Language Evolution: Psycholinguistic Predictors of Loan Word

Borrowing

An adequate model of diachronic language change must establish not only the drivers of change, but also uncover the properties of the communicative system that result in stability of language structure (Croft, 2000). Forces driving language change are diverse, originating in social, cultural and economic influences (Labov, 2001, 2007), but there has been a recent increase in interest in the role of cognitive processing as a factor affecting how and whether language structures change or remain stable (e.g., Corballis, 2017; Steels, 2017). Investigations of language change have been energised by cladistics studies of languages, which propose groupings of languages, and words within languages, according to their common ancestry. The logic of this approach is that languages that differ to a greater degree are more distinct in time (Gray & Atkinson, 2004) and in space (Atkinson et al., 2011). Such cladistic studies have, however, typically investigated the similarities and distinctions that occur between small sets of vocabulary items, such as the Swadesh lists of 200 words (e.g., Atkinson et al., 2008, 2011; Blasi et al., 2016; Gray & Atkinson, 2004) which can then be used to determine which vocabulary items are more or less prone to change by investigating the variability of word forms expressing a similar concept across sets of languages. These descriptions of stability and change in small sets of vocabulary items have begun to render suggestions as to the social and cognitive contributors to changes in the vocabulary (Lieberman, Michel, Jackson, Tang, & Nowak, 2007; Monaghan, 2014; Pagel et al., 2007; Vejdemo & Hörberg, 2016; Winter et al., 2014).

For instance, studies of cognates across languages have demonstrated that particular psycholinguistic properties of a word can predict the stability or volatility of its form, suggesting how cognitive processes involved in language representation and language acquisition may result in language change (Monaghan, 2014). From a cladistics study based on Swadesh word lists, Pagel et al. (2007) estimated the rate of change of individual words within a language. This was done by first determining the number of distinct word forms expressing a similar concept across multiple languages (so two words that are similar but overlapping, such as *aqua* and *agua,* would not count as distinct forms), and then determining the rate of change required to generate the observed variety of forms across languages within the Indo-European language family. Pagel et al. (2007) then discovered that the grammatical category of the word was highly related to rate of lexical change of a word – words that were nouns had a higher rate of change than words from other grammatical categories, and function words had the lowest rate of change.

In addition, Pagel et al. found that the contemporary frequency of a word predicted its rate of lexical change – higher frequency words had a lower rate of change. This was interpreted as indicating that high frequency usage of a word provides protection for that word against forces of change: If one speaker changed the form of a high frequency word, that would result in a greater number of misunderstandings than if a lower frequency word was used with alternative forms (Boyd & Richerson, 1985; Bybee, 2007; see also Pagel et al., 2013). Pagel et al.'s (2007) approach thus aligns the parallel traditions of social and cognitive approaches to language change: language usage models can thus be related to the psycholinguistic analysis of the effect of frequency. In another recent study on a subset of the Swadesh word list, Vejdemo and Hörberg (2016)

confirmed that higher frequency words and a distinction between nouns and verbs were predictors of greater stability of word forms, and in addition high imageability and a small number of synonymous meanings for the word form were also protective factors against change.

Monaghan (2014) further extended these analyses to demonstrate that speakers' history of learning also related to rate of lexical change. In addition to grammatical category and frequency (Bybee, 2007), later age of acquisition predicted greater lexical change, meaning that the words that children acquire earlier are more resistant to change. We know that earlier acquired words are accessed more quickly and accurately (Brysbaert & Ghyselinck, 2006; Juhasz, 2005), remain easier to access throughout the lifespan (Hodgson & Ellis, 1998), and are more likely to be sustained in the vocabulary of people with dementia (Holmes, Fitch, & Ellis, 2006). Hence, Monaghan (2014) suggested that rate of lexical change was not just a consequence of its usage, as discovered by Pagel et al. (2007), but also due to how robustly a vocabulary item is represented by the speaker in their vocabulary, as determined by a word's age of acquisition.

Monaghan's (2014) analysis showed that word forms acquired early in language development are less prone to change, and suggested that innovations in the forms of words, driving language evolution of the vocabulary, are more likely to occur in terms of communication between proficient speakers of the language (Bybee & Slobin, 1982; Joseph, 1992; Slobin, 2005) rather than occurring as a consequence of the pressures of acquisition, counter to views that language acquisition is the primary driver of innovation and change (Bickerton, 1990; Hudson Kam & Newport 2005; Kempe, Gauvrit & Forsyth,

2015; Newport, 1990). These results are consistent with Labov's (2007) distinction between (vertical) transmission and (horizontal) diffusion as drivers of language change: During early language acquisition, the child must first replicate the care-giver's language structure, which is consistent with earlier acquired words being more stably represented in the language. In later stages of acquisition, the child then begins the process of "incrementation" of the language, where innovations and adaptations can occur − supported by observations of a greater rate of lexical change for later acquired vocabulary items. In other words, transmission of the language involves replicating the structures within the language community, whereas diffusion documents the role of influence and change in language structure across language groups.

In these previous analyses of rate of change of words, the measure of vocabulary change is estimated by comparing word forms across many languages for small sets of concepts, and determining the diversity of these word forms. Concepts expressed in very different forms across languages indicate a higher rate of language change. However, the set of words used in these studies are not only small but likely to be unrepresentative of the vocabulary more broadly. Previous studies based on Swadesh word lists provide the word forms for approximately 200 concepts that are almost universally shared across cultures, including body parts, simple actions and objects, and family relations (Swadesh, 1955). Furthermore, these lists comprise "core vocabulary" which are specifically chosen to be unlikely to undergo replacement, according to Swadesh's (1955) original rationale for constructing his word lists. Therefore, analyses derived from words in Swadesh lists are unlikely to reflect language change representative of the whole vocabulary. This becomes especially important for determining effects of frequency on language change,

where vocabulary change may be driven by multiple forces associated in different ways with frequency. For instance, higher-frequency words may be stably present in speakers' vocabulary because of their greater fidelity of representation and the importance of avoiding frequent confusion between speakers (Pagel et al., 2007), but for a word form to be replaced at all, it has to be used. For instance, change as a consequence of contact between languages, or as a process of incrementation to define a social group (Labov, 2007), requires exposure to those language structures in the first place between interlocutors. A very low frequency structure may be less likely to be replaced, or undergo innovation, than a slightly higher frequency word. Hence, effects of frequency may vary according to whether a word is higher or lower, or very low in frequency, and thus a wider range of vocabulary items is needed to explore change.

A more direct approach to identifying vocabulary change than cladistics studies is to investigate the inclusion of loan words into a language, as a process of borrowing between languages (Bloomfield, 1933; Haspelmath & Tadmor, 2009; Thomason & Kaufman, 1988). There are three key effects a borrowed word may exert. First, it may replace an existing word in the language, an "intimate" borrowing in Bloomfield's (1933) terminology. Alternatively, the borrowed word may coexist with a word in the language with a similar meaning, or it may be an insertion, i.e., referring to a concept that was not previously expressed in the language, both forms of "cultural" borrowing (Bloomfield, 1933). Compatible with the outcome from the cladistics studies of language change, if a concept is expressed with a word form that is stably represented in the speaker's vocabulary then we propose that it is less likely to be *replaced* and less likely to admit of *coexistence* with a word form from another language, thus it will be more resistant to

borrowing. In contrast, if the word form is represented less robustly then it is more likely to be replaced or will be more compatible in terms of coexistence with a word borrowed from another language.

It might be anticipated that the cognitive mechanisms influencing *insertion* of a borrowed word may well be different, as these insertions do not result in changes to the previously existing vocabulary. Hence, the suggestion of representational robustness does not make predictions about these insertions. We distinguish these different effects of borrowing in the analyses in Study 1 (see supplementary materials S4). We found that all borrowing effects demonstrated a similar relation to the psycholinguistic variables, with the exception that insertions were distinct than other borrowing effects with respect to concreteness: a word is more likely to be an insertion borrowing if it is high in concreteness.

The advantage of analyzing such a data set is multiple. First, it can provide additional tests of the psycholinguistic properties of words, reflecting the acquisition and maintenance of representations of word forms in the speaker's vocabulary, that relate to whether words are resistant or prone to change. Second, it can enable an analysis of a substantially larger word set than is possible from using Swadesh lists. This larger set not only provides greater power to the analyses, but it also means that analyses of vocabulary stability and change are more likely to be representative of the vocabulary as a whole, potentially revealing different effects of the influences observable from smaller, specially chosen word lists. Third, it provides insight into language change in plain sight – borrowings across languages have occurred within recent history, and are occurring right now as vocabularies change and innovate to accommodate new ideas, concepts, as well

as cross-linguistic, social and cultural influences (Grant, 2009; Thomason & Kaufman, 1988). We anticipate that the cognitive drivers of stability and change in the vocabulary that have been observed previously (e.g., Monaghan, 2014; Pagel et al., 2007; Vejdemo & Hörberg, 2016) will also be observable in these larger sets of loanwords.

There have been some previous studies that relate loanwords to psycholinguistic properties of words' semantic and syntactic features. Grammatical accounts of borrowing have suggested (but not tested) the hypothesis that words that do not require grammatical adaptation are more likely to be borrowed than others (Myers-Scotton, 1993). According to Myers-Scotton (1993) a noun from a donor language can be incorporated into a sentence in the recipient language without substantial reconfiguration of the morphology of the word. A verb, however, is likely to require more complex affixation before inclusion into the recipient language and so is less likely to be borrowed into the language due to the additional cognitive cost in terms of complexity of this integration. There may be substantial cross-linguistic variation, however, in terms of the complexity of the morphological system applying to each grammatical category (see, Thomas, 1983, for examples). We return to this issue in the General Discussion.

Cassidy (1999) investigated the relation between grammatical category and borrowings in English, and found that nouns, then adjectives, then verbs were more likely to originate in non-Germanic languages (i.e., are borrowed), and function words were more likely to be Germanic in origin. Cassidy (1999) also described the relation between frequency and borrowing in English, using the Brown corpus, and found that 98% of the most frequent 100 words in English were of Germanic origin, and this dropped to 88.5% of the most frequent 200 words. For the 5001-6000th most frequent words in English,

only 34.2% were of Germanic origin. Thus, the relation between frequency and vocabulary change (Pagel et al., 2007) appears also to be observable in vocabulary borrowing. However, the relative contribution of grammatical category, frequency, and a range of other psycholinguistic variables that may also predict borrowing have not been explicitly tested. It could be, for instance, that frequency effects are due to the confound of frequency with different grammatical categories, or even because of the confound between frequency and length (Zipf, 1949), or with age of acquisition (Monaghan, 2014). It is thus important to test these variables simultaneously for their contribution to stability and change of the vocabulary.

In the current study, we provide a psycholinguistic analysis of predictors of whether a word is borrowed from another language. We assessed whether predictors in rate of lexical change analyses were verified within analyses of changes within a vocabulary as measured by loan word borrowing. In particular, we wanted to test whether grammatical category, frequency, word length, and age of acquisition were predictors of borrowing in the same way as they applied to rate of lexical change analyses. Furthermore, we wanted to determine whether innovations in the vocabulary were observable in a larger vocabulary than was possible in previous analyses of lexical change, which were previously restricted to a 200 item vocabulary derived from the Swadesh lists (Dyen et al., 1992; Swadesh, 1952; see also Newberry, Ahern, Clark, and Plotkin, 2017, for analysis of language change in morphology from a slightly larger word set). The World loanword database (WOLD, Haspelmath & Tadmor, 2009) provides lists of 41 languages with between 1000 and 2000 vocabulary items, where words are derived from the Intercontinental Dictionary Series list of synonyms provided for a large range of

Indo-European languages (Key & Comrie, 2015). These lists were in turn derived from Buck (1949), and were intended to provide a set of core concepts expressed across many languages, similar to the motivation of Swadesh (1952; 1955), but for a substantially larger set of words. Indeed, a comparison between the Swadesh list of English and the loanword database for English by Grant (2009) documents a far broader range of etymological origins of words in the loanword list than the Swadesh list. The database also contains an indication of whether the word is borrowed or if there is no evidence of borrowing within each language, and, if so, the historical point that it was first borrowed. For English, these diachronic data are derived from etymological dictionaries and from historical texts corpora (Grant, 2009). Though the Intercontinental Dictionary Series lists include a far larger set of concepts than Swadesh lists, covering a much broader range of word frequencies than the Swadesh lists, see Study 3 for analysis and Borin, Saxena, and Comrie (2013) for further description, it is important to note that they still include only a fraction of the whole vocabulary.

Pagel et al.'s (2007) analyses on the Swadesh lists ensured that the vocabulary items that were analysed were fundamental to communication and general across cultures. Thus, Pagel et al.'s measure of rate of lexical change concerns replacement of vocabulary items that were likely already existing in the language. Indeed, it has been claimed that vocabulary change in these lists are not due to effects of borrowing (Atkinson et al., 2011), even though the Swadesh list of English words contains words that are attested from at least 7 other languages (Grant, 2009, p.371). However, it is certainly the case that the WOLD also contains words from a richer spectrum of usage, and as such documents loan words for a substantially wider range of concepts, some of

which may have replaced existing terms, but others that may be due to insertions resulting from cultural influence, including concepts for which there may not have been a word in the language prior to borrowing, such as for technical developments, newly discovered animal species, incorporation of new schools of thought (e.g., religious terms), or even names for foods or other cultural products that are introduced to the language community (Bloomfield, 1933; Croft, 2000; Haugen, 1950; Matras & Sakel, 2007; Myers-Scotton, 1993). Thus, whereas the WOLD provides a more extensive test of the psycholinguistic predictors of lexical change it is also important to account for additional contributors to loan word exchange, such as concreteness (Brysbaert, Warriner, & Kuperman, 2014; Vejdemo & Hörberg, 2016), with cultural or technological concepts introduced into a culture, and hence requiring a word form that did not previously exist, more likely to be abstract (Thomason & Kaufman, 2000), but terms for animals or foods more likely to be concrete. It is also feasible that borrowing may be facilitated for concrete words compared to abstract words, since they are recognised and retrieved faster and more accurately (Strain, Patterson, & Seidenberg, 1995; Walker & Hulme, 1999), they may be more similar in meaning between languages (Van Hell & De Groot, 1998), and they may be borrowed alongside their physical referents (e.g., during trade).

In Study 1 we provide an analysis of loan words in English, determining the probability that a word is borrowed from another language as an indication that it is less likely to be stable as a word form, and determining whether previously observed psycholinguistic predictors of change are also observed in this larger, more representative sample of the vocabulary. In all the studies in this paper, in line with previous studies of

psycholinguistic predictors of language change, we assume that word frequency and age of acquisition of words remain relatively stable over time (Monaghan, 2014; Pagel et al., 2007; Vejdemo & Hörberg, 2016). Diachronic corpus analyses make similar assumptions – indeed the rare instances of change in relative frequencies of words permits key changes in word usage to be detected in historical corpora (Baron, Rayson, & Archer, 2009; Kilgariff, 1997; Sinclair, 1991).

As English is somewhat unique in terms of the extent of the social and cultural influences from various language families over the last two thousand years (Leslie et al., 2015), we also determined whether a similar set of psycholinguistic predictors of stability and change applied to another language, or whether the observed pattern is specific to English. In Study 2, we selected Dutch for analysis because it is a language for which the set of psycholinguistic measures are currently available. Although English and Dutch are closely related Germanic languages they have undergone rather different processes of borrowing – with English borrowing substantially from Latin and French, whereas Dutch borrowing rates are lower overall, and drawn from more varied sources (Haspelmath & Tandor, 2009).

In Study 3 we related the loan word status of words to Pagel et al.'s (2007) rate of lexical change measure to determine the relation between the Swadesh words and the larger set of words analysed here, to determine whether the loanword analyses confirm or contrast with conclusions derived from this smaller set of words. Finally, in Study 4, we tested whether word length from Old English, rather than a contemporary measure of word length as used in Studies 1 to 3, predicted whether a word would be borrowed or not subsequent to the Old English period.

Study 1: Psycholinguistic influences on loan words in English

Method

Materials

The set of words were taken from the WOLD database (Haspelmath & Tadmor, 2009), which provides the word form, whether the word is borrowed from another language, and the point at which it is first documented as borrowed (Grant, 2009).

The English list was compiled by Grant (2009) and comprised 1515 word meanings. Word forms with distinct meanings are entered separately in the database, such as "sink" as noun and as verb. If the borrowed status of the word form was the same for each of the word's multiple usages then the entries were combined (so for "sink" both the noun and the verb usage were classified as borrowings). There were two word forms where usages differed in their borrowing status ("calf", where the young cow meaning had no evidence for borrowing, and the meaning referring to part of the leg was "probably borrowed" from Old Norse; and "boil", where the noun meaning provided no evidence of borrowing, but the verb meaning was "clearly borrowed" from Latin via Anglo-Norman French).   Grammatical category information was taken from Brysbaert et al. (2014), which provided the grammatical category associated with the most frequent usage for each word form (e.g., for English word "sink", the noun usage is more frequent).

Phonology for the words in the list was taken from CELEX (Baayen, Pipenbrock & Gulikers, 1995). Words that did not appear in CELEX were hand-coded. Entries from WOLD that were multi-word phrases (e.g., low tide, driving licence) were omitted to

ensure that all entries were individual words. Length in phonemes was then calculated, with diphthongs and affricates counting as single phonemes in the analyses.

Frequency was taken from the Zipf SUBTLEX-UK frequency (van Heuven, Mandera, Keuleers, & Brysbaert, 2014) which is an established measure of frequency that closely approximates spoken language usage.

Age of acquisition for each word was taken from Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012). This measure was generated by participants making subjective ratings about the age at which they first learned each word, but these subjective measures have been shown to be highly valid when compared to actual acquisition age (Gilhooly & Gilhooly, 1980; Morrison, Chapell, & Ellis, 1997). A recent study has also provided objective test-based measures of age of acquisition for English (Brysbaert & Biemiller, 2017), and these ratings were also analysed to confirm the effect of the subjective age of acquisition measure.

We also included a measure of concreteness (Brysbaert, Warriner, & Kuperman, 2014) in order to test for cultural and technical innovations as a potential confound in analyses of loan word borrowing likelihood.

For English, there were 1390 distinct word forms with values for all psycholinguistic variables. Of these, 1317 were classified either as "clearly borrowed" or "no evidence of borrowing". All data and analysis scripts are available in an online repository (https://github.com/seannyD/BorrowingFreqAoA) and in supplementary materials 1, 2, 3, 4 and 5.

Of the words that were classified as clearly borrowed, 167 had the effect of replacement, 162 were coexistences, and 197 were insertions. The supplementary

materials (S4) show that the key psycholinguistic variables of interest in our analyses –

length, age of acquisition, and frequency – apply similarly across all these borrowing

types. The results also showed that higher concreteness significantly related to greater

probability of insertions, indicating that novel innovations described in the language

tended to be high concrete terms. This was due to many of the borrowings with highest

concreteness being insertions, with technological and social concepts being rare, but

labels for novel animals, domestic effects (such as *pillow*, or *toilet*), foods and clothes

being frequent, each of which has high concreteness.

Analysis

Whether a word was classified as "clearly borrowed" or "no evidence of

borrowing" was taken as the dependent variable, with borrowing coded as 1, and no

borrowing as 0. The set of psycholinguistic variables were taken to be predictors (see

supplementary materials S1).

Preliminary analyses revealed substantial non-linear relationships in the data.

Linear models can accommodate this by adding extra fixed effects to represent the

quadratic or cubic terms of a variable to represent an exponential, parabolic relationship

(or higher terms to model an s-curve, and so on). However, this considerably increases

the number of possible terms in the model, especially when considering multiple

variables and interactions, making it difficult to identify the best model. To avoid this

issue, we used general additive models (GAM) which fit non-linear curves to the data

(using the R package *mgcv*, Wood, 2011). The analysis attempts to fit the data while

avoiding overfitting by penalising the use of more complex curves. For each independent

term in the model, two basic statistics are provided: how well the term predicts the dependent variable (likelihood of borrowing in this case), represented by a χ^2 statistic and a *p*-value (lower *p*-values indicate a relationship unlikely to be due to chance); and how non-linear the relationship is, represented by the estimated degrees of freedom (edf, a value of 1 indicates a linear relationship, with higher values for more complex curves). We included fixed effects for length, age of acquisition, frequency and concreteness (variables were scaled and centred).

We included grammatical category as a random effect, which modelled the possibility that some parts of speech are more likely to be borrowed. Grammatical category was selected to be a random rather than a fixed effect primarily because the number of words in each category varied, which could have biased the model in favour of smaller categories, and also because there are additional parts of speech that occurred in Brysbaert et al. (2014) that did not occur in the current analysis. As an additional check, we ran additional analyses with grammatical category as a fixed effect and the results were very similar, so we do not report these analyses further here.

To test the possibility that the psycholinguistic variables have a greater effect for particular parts of speech, we included random slopes for each fixed effect with grammatical category. We also conducted model comparisons to determine whether there were significant interactions between fixed effects, but none of these were significant. See the supplementary materials for the full data and analysis scripts.

Results and discussion

The model accounted for 16.2% of the deviance ($R^2 = 0.19$). The probability of a word being borrowed was significantly predicted by grammatical category (edf = 5.88, χ^2 = 39.7, $p < 0.0001$), length (edf = 1.62, $\chi^2 = 32.3$, $p < 0.0001$), frequency (edf = 3.41, $\chi^2 = 32.6$, $p < 0.0001$), age of acquisition (edf = 1.00, $\chi^2 = 35.6$, $p < 0.0001$), with a marginal effect for concreteness (edf = 2.68, $\chi^2 = 7.64$, $p = 0.073$). The random slopes did not account for a significant proportion of the variation, meaning that the observed psycholinguistic effects were not distinct across the various grammatical categories.

There was little difference in the model fit when using the objective measure of age of acquisition instead of the subjective ratings (overall deviance explained 15.7%, length edf = 2.12, $\chi^2 = 34.8$, p < .001, frequency edf = 3.44, $\chi^2 = 25.3$, p < .001, objective age of acquisition edf = 1.76, $\chi^2 = 28.3$, $p < 0.0001$), except that concreteness now exhibited a significant effect (edf = 2.22, $\chi^2 = 7.03$, p = .043): more concrete words were less likely to be borrowed than more abstract words.

For grammatical category, the percentage of borrowings by category for the set of words in English is shown in Table 1. Consistent with Cassidy (1999), borrowing is more likely for nouns than adjectives and verbs, which are more likely to be borrowed than the other categories (although the percentage borrowing of some categories such as names is high, the number of observations is much lower, providing a lower confidence bound on the estimated probability of borrowing).

Grammatical category	Percentage Borrowed	Total Number of words
Noun	49.5% [46.1,52.9]	812
Verb	33.3% [28.1,39.0]	288
Adjective	35.2% [27.4,43.9]	125
Name	50% [9.5, 90.5]	2

Conjunction	16.7% [3.0, 56.4]	6
Pronoun	11.1% [2.0, 43.5]	9
Determiner	7.7% [1.4, 33.3]	13
Number	5.6% [1.0, 25.8]	18
Adverb	3.3% [0.6, 16.7]	30
Interjection	0% [0.0, 65.8]	2
Not	0% [0.0, 79.3]	1
Preposition	0% [0.0, 25.9]	11

Table 1. The percentage of words borrowed in each grammatical category in English. Numbers in brackets give the confidence intervals for the percentage estimate using Wilson's (1927) binomial method. Categories are sorted by the estimated lower bound.
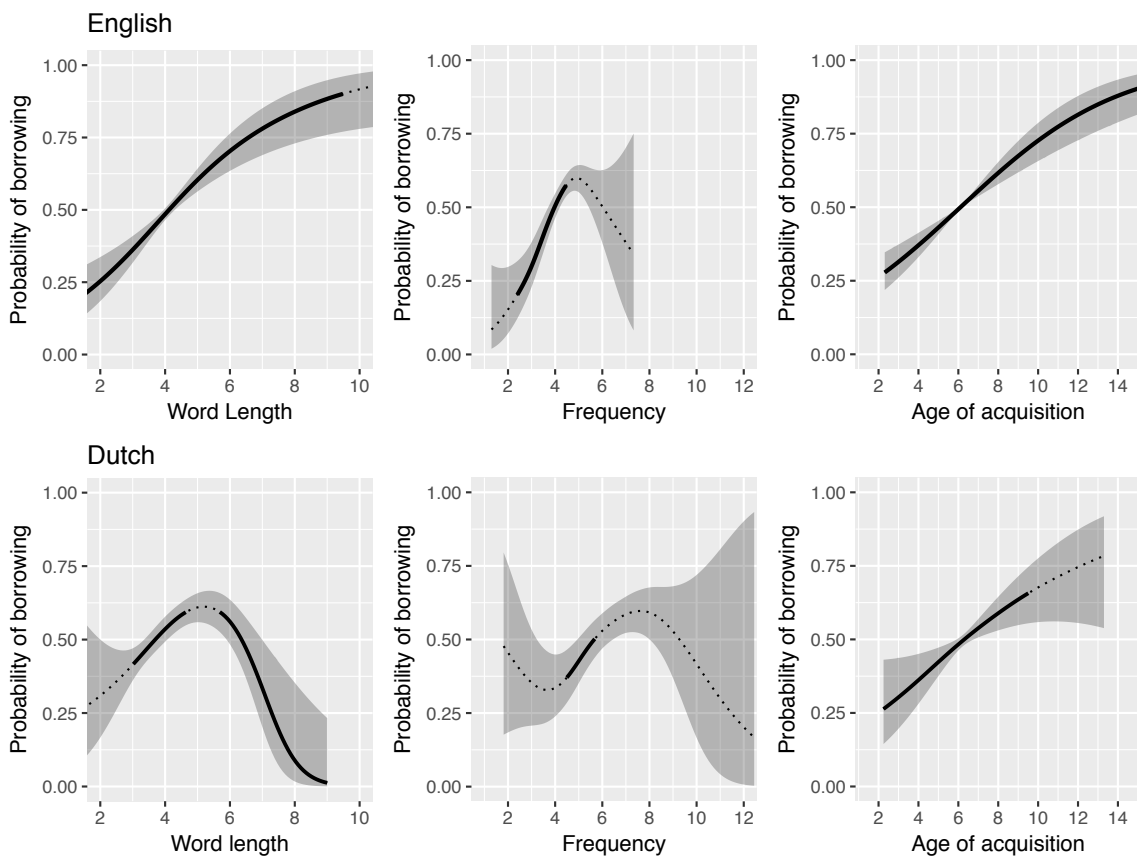

Figure 1. Relationship between probability of borrowing with length, frequency, and age of acquisition for English (top) and Dutch (bottom). Lines are model estimates (dotted

lines are estimates and solid lines indicate where the estimates are significantly increasing

or decreasing). Shading indicates 95% confidence intervals around the estimates.

Figure 1 (upper panels) shows the model estimates for the psycholinguistic

variables that were significantly related to probability of borrowing in English. For

length, as the word's form gets longer, the probability of it being borrowed increases in a

monotonic relationship. Similarly, the relationship between age of acquisition and

probability of borrowing is monotonic, where words that are acquired later tending to be

more likely to be borrowed than those words acquired earlier in language development.

These results are compatible with the observations of Monaghan (2014) for stability and

volatility of a word's form relating to its length and age of acquisition: longer, later-

acquired words are more vulnerable to change.

However, the relationship between borrowing probability and frequency is not

straightforward. In Figure 1, the unit of frequency is $\log_{10}$ of words per billion, which is

van Heuven et al.'s (2014) Zipf frequency measure. For mid- to high-frequency words

(above approximately 5, corresponding to words with frequency greater than one

occurrence every ten thousand words in a corpus), the relationship is negative and

monotonic, with increasing frequency relating to reduced likelihood of borrowing,

consistent with Pagel et al. (2007). However, for mid- to low-frequency words (with

frequencies less than one per ten thousand) the relationship is positive and monotonic,

with increasing frequency resulting in *greater* likelihood of borrowing, contrasting with

previous observations of the link between vocabulary change and frequency in smaller

sets of words.

In order to determine for each psycholinguistic variable for which region of the slope was it significantly increasing or decreasing, we determined the derivatives and standard errors for the derivatives along the gradient of the slope. Where the confidence interval for the derivatives do not overlap zero, there is a significant increase or decrease in the slope. Figure 1 shows the regions of the slopes that indicate significant change. For age of acquisition and length the results are straightforward, the monotonic increase in borrowing probability is significant at all ages, and for all lengths, except for the very longest words. For frequency, the derivatives of the slope analysis demonstrated that the only significant effect of frequency was for the low to mid-frequency words, where increasing frequency resulted in a greater likelihood of borrowing. For the mid- to high-frequency range, the negative slope did not have a significant decreasing gradient, and so for this larger set of words, there was not clear evidence of the negative effect of frequency on vocabulary change that was observed in studies of the Swadesh lists (e.g., Pagel et al., 2007).

It is possible that collinearity between the psycholinguistic variables may have affected the model fit, and so we also fit GAM models to probability of borrowing with each psycholinguistic predictor entered individually. The values and shapes of the GAM curves were very similar to the analysis when predictors were entered simultaneously, (length, edf = 2.09, χ^2 = 30.67, p < .001, frequency, edf = 2.94, χ^2 = 14.03, p = .006, age of acquisition, edf = 1.80, χ^2 = 34.87, p < .001, though concreteness was now significant, edf = 1.10, χ^2 = 23.85, p < .001). Thus, the effect of frequency, length, and age of acquisition, was qualitatively similar regardless of whether the variables were considered simultaneously or individually.

The results replicate previous studies of psycholinguistic factors predicting vocabulary change. Previous studies showing grammatical category affecting probability of borrowing were confirmed by these analyses: nouns were more likely to be borrowed than other open class word categories, which were in turn more likely to be borrowed than function words. Furthermore, age of acquisition and word length, which relate to the robustness of the representation of a word in the speaker's vocabulary, were both found to relate significantly to the probability of a word being borrowed. Frequency, however, showed a different relation in the larger vocabulary than that previously observed in smaller-scale cladistics studies of rate of change in the vocabulary: a low-frequency word was less likely to be borrowed than a medium-frequency word, possibly because of the lower opportunity for exposure to words from another language. For mid- to high-frequency words, there was a trend in the same direction as found in cladistics studies of lexical change, where increasing frequency related to lower probability of borrowing.

Study 2: Psycholinguistic influences on loan words in Dutch

The overlap between the 41 languages available in WOLD and databases for which large-scale psycholinguistic measures are available is very small. However, broad coverage in terms of psycholinguistic variables were also available for Dutch, and so we determined whether observations of loan words in English in Study 1 showed a similar pattern in another language.

Method

      Materials

The set of Dutch words were again taken from the WOLD database (Haspelmath & Tadmor, 2009), providing the word form, borrowing status, and the point at which it is first documented as borrowed. The Dutch list, compiled by van der Sijs (2009), was composed of 1588 words. Homonyms were combined in the same way as for the English database, i.e., if the borrowing status for homonyms was the same then they were combined, if borrowing status was distinct then the homonyms were maintained as separate entries in the database.

Grammatical category information was taken from Keuleers, Brysbaert, and New (2010) for Dutch. Phonology for the words in the list was again taken from CELEX (Baayen et al., 1995), and frequency for Dutch was the Zipf SUBTLEX-NL frequency (Keuleers et al., 2010). Age of acquisition was taken from Brysbaert, Stevens, De Deyne, Voorspoels, and Storms (2014), which was collected in an equivalent way to the English database (Kuperman et al., 2012). Concreteness was also taken from Brysbaert et al.'s (2014) database.

For Dutch, there were 1028 distinct word forms for which all the psycholinguistic variables were available and which were classified as either clearly borrowed or no evidence of borrowing.

Analysis

As for the English analysis, whether a word was classified as "clearly borrowed" or "no evidence of borrowing" was taken as the dependent variable, with borrowing coded as 1, and no borrowing as 0. As for the English analyses, we included fixed effects for length, age of acquisition, frequency and concreteness, with variables scaled and

centred. Grammatical category was again a random effect, as there were additional parts of speech in Keuleers et al. (2010) that were not present in the current corpus. Once again additional analysis checks with grammatical category as a fixed effect resulted in very similar results.

As for English, model comparison indicated that there were no significant interactions between fixed effects, so we report the same GAM analyses as for English, with models with fixed effects, and random slopes for each grammatical category.

Results and discussion

For Dutch, the model explained 13.7% of the deviance ($R^2 = 0.11$), and the results were similar in terms of significant effects (see supplementary materials S2): there were significant effects of grammatical category (edf = 3.72, $\chi^2 = 38.6$, $p < 0.0001$), length (edf = 2.64, $\chi^2 = 16.9$, p = 0.0004), frequency (edf = 3.67, $\chi^2 = 12.0$, $p = 0.02$), age of acquisition (edf = 1.37, $\chi^2 = 11.3$, $p = 0.007$), but not concreteness (edf = 1.65, $\chi^2 = 3.03$, $p = 0.23$). As in English, random slopes did not account for significant variation.

Grammatical category	Percentage Borrowed	Total number of words
Noun	27.6% [24.3,31.2]	630
Adjective	14.4% [8.9, 22.4]	104
Verb	4.6% [2.4, 8.5]	196
Name	7.1% [1.3, 31.5]	14
Preposition	6.3% [1.1, 28.3]	16
Number	5.6% [1.0, 25.8]	18
Adverb	4.3% [0.8, 21.0]	23
Conjunction	0% [0.0, 43.4]	5
Interjection	0% [0.0, 65.8]	2
Article	0% [0.0, 65.8]	2
Pronoun	0% [0.0, 17.6]	18

Table 2. The percentage of words borrowed in each category in Dutch. Numbers in brackets give the confidence intervals for the percentage estimate using Wilson's (1927) binomial method. Categories are sorted by the estimated lower bound.

Table 2 shows the percentage of borrowings for each grammatical category in Dutch, and Figure 1 (lower panels) shows the model estimates for the significant psycholinguistic variables. Table 2 confirms that overall borrowing rates are lower in Dutch than English (Haspelmath & Tandor, 2009), and reflects Cassidy's (1999) observations of greater borrowings for nouns and adjectives than other parts of speech. However, verbs were not borrowed at a higher rate than other parts of speech, distinct from the results for English.

As shown in Figure 1, for length, the relationship is monotonic and positive for words up to length 5 phonemes (as in English) but then is negative for longer words, so longer words are less likely to be borrowed than medium length words. To formally test the differences between languages, we ran a single GAM with the combined data from

both languages and a random effect for language (see supplementary materials S5).

Adding a term which allowed the effect of length to vary by language significantly

improved the model (χ^2 = 18.5, p < 0.001), suggesting that there are significant

differences between the languages regarding the relationship between the probability of

borrowing and word length.

      This could reflect the role of compounding in Dutch, which tends to result in a

greater proportion of longer, polymorphemic words in the vocabulary than in languages

which have less of a tendency to compound words, such as English. Multimorphemic

words require greater grammatical adaptation for the lexical root, as it requires

accommodation to patterns of derivational and inflectional morphology. To explore this

possibility, we ran a post hoc analysis on the borrowing dataset above, adding

the morphological classification of each word from CELEX (85% of the words were

monomorphemic). We fit a new GAM which allowed the effect of word length to vary by

whether the word was monomorphemic or not. This new model significantly improved

the fit compared to the first model above (χ^2 = 9.53, p = 0.002) and the effects of

grammatical category, frequency, age of acquisition and concreteness did not change

qualitatively (see supplementary materials S2). In this analysis, the relationship for length

is now split into two components, illustrated in Figure 2. The relationship for

monomorphemic words is monotonic and positive (edf = 1.08, χ^2 = 14.8, p = 0.0003),

much like the English data, while multimorphemic words have a weaker, negative

relationship (edf = 2.02, 9.20, p = 0.06) with longer multimorphemic words less likely to

be borrowed than medium length multimorphemic words. For completeness, we also ran

this analysis for English (using morphological classifications from CELEX, 80% of

words were monomorphemic). Allowing separate curves for monomorphemic and multimorphemic words significantly improved the model ($\chi^2 = 4.75$, p $= 0.03$, see supplementary materials S1). Both types of word had a significant, positive, linear relationship with probability of borrowing (monomorphemic: edf $= 1.0$, $\chi^2 = 35.9$, p < 0.0001; multimorphemic: edf $= 1.0$, $\chi^2 = 11.7$, p $= 0.0006$, see SI), and the relationship of the other variables did not differ qualitatively. Though the analysis was conducted post hoc, it was consistent with broader theories of grammaticalisation affecting probability of borrowings.
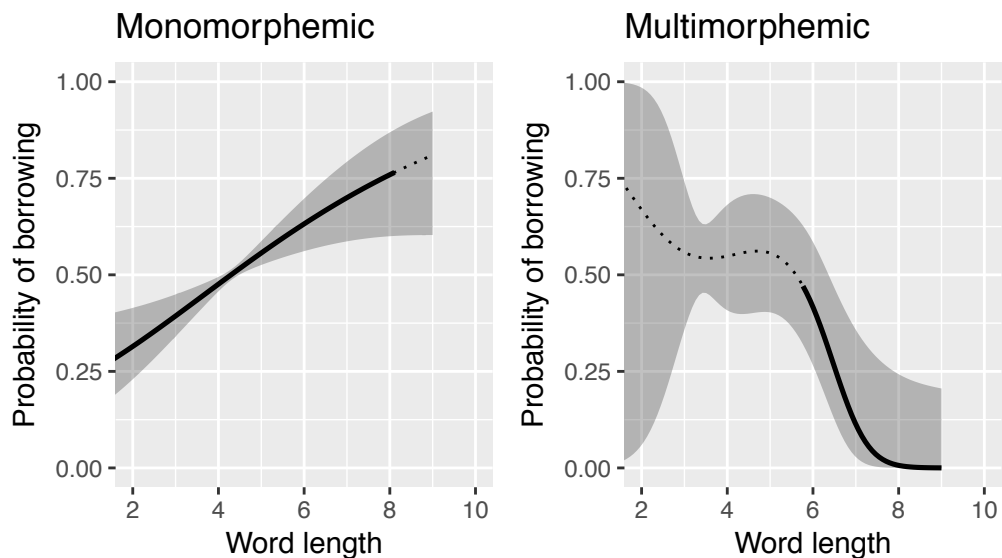

Figure 2: Relationship between probability of borrowing with length for Dutch words that are monomorphemic (left) and multimorphemic (right). Lines are model estimates (dotted lines are estimates and solid lines indicate where the estimates are significantly increasing or decreasing). Shading indicates 95% confidence intervals around the estimates.

Returning to the other psycholinguistic predictors in the model, as illustrated in Figure 1, age of acquisition in Dutch has the same effect as in English: earlier acquired words are less likely to be borrowed than later acquired words. A formal comparison, as above, showed no significant difference between languages ($\chi^2 = 0.8$, $p = 0.64$).

The relationship for frequency is similar to English, with a monotonic and negative relationship between frequency and probability of borrowing for mid- to high-frequency words, and a positive relationship between frequency and probability of borrowing for low- to mid-frequency words. A formal comparison revealed that the frequency curves significantly differed between languages ($\chi^2 = 9.2$, $p < 0.001$), but this was due to differences for very low frequency words, where the confidence intervals are so large that the differences should be considered unreliable.

The disparity in the frequency effect found in our analyses of probability of borrowing and that found in previous cladistics studies of vocabulary change are possibly due to increased power in the analyses resulting from a larger data set, or due to a different range of distributions of frequency in the data. However, it remains possible that borrowing is a very different process to that measured by comparing similarities and differences of cognate forms across different languages as in cladistics studies. Another possibility is that the non-linear statistical modelling we applied may have resulted in very different statistical results compared to a model that applies only a linear fit to the data, as has been used in previous analyses of psycholinguistic contributors to language change (Monaghan, 2014; Pagel et al., 2007; Vejdemo & Hörberg, 2016). In Study 3, we compared the distribution of psycholinguistic variables in the 200 word Swadesh vocabulary and the larger set of words in the loanword database. We also measured the

relationship between borrowing and rate of lexical change in Pagel et al.'s (2007) set of

vocabulary items, and reanalyzed rate of lexical change in the smaller set of words using

non-linear modelling techniques to determine whether non-linear effects of frequency, or

length and age of acquisition, were also observable in the English Swadesh word list.

Study 3: Relations between lexical rate of change of words and loan word status

Method

Data

The data were the English Swadesh list, with rate of lexical change taken from

Pagel et al. (2007), and grammatical category, length, age of acquisition, concreteness,

and frequency measures exactly those used in Study 1 for the English loanword data set.

The overlap between the loanword word set and the Swadesh list comprised 190 words.

Analysis

We first compared the distribution of psycholinguistic variables in the loan word

data set in English and the English Swadesh list. We then determined the relation

between Pagel et al.'s (2007) rate of lexical change of English words and probability of

borrowing with a linear regression model. We next analysed rate of lexical change using

a preliminary GAM analysis with the psycholinguistic predictors of length, frequency,

age of acquisition, and concreteness. However, these analyses demonstrated that all

relations were linear (see supplementary materials S3). We thus used a linear regression

model to predict Pagel et al.'s (2007) rate of lexical change of English words, using the

same set of predictors as employed in Study 1. Model comparison was used to test

whether or not to include each of the psycholinguistic variables as main effects. Finally,

we repeated the GAM analyses of Study 1 with borrowing as the dependent variable for this subset of words. Together these analyses enable a test of whether the sampling of the Swadesh list was representative of the larger corpus.

Results and discussion

Figure 3 compares the distribution of the key psycholinguistic variables in the Swadesh word list and the larger loanword database. The results indicate that the inclusion of words from the wider vocabulary has extended the range of all variables compared to the Swadesh list.
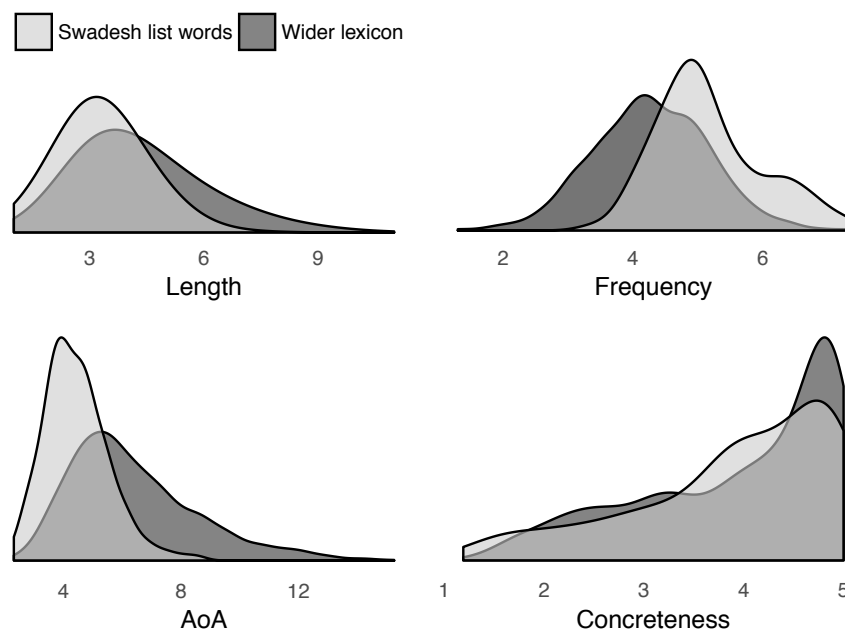

Figure 3. The distribution of words in the English data for length, frequency, age of acquisition, and concreteness for words according to whether they appear in the Swadesh word list and the wider lexicon in the loanword database.

Rates of lexical change were greater for borrowed words (M = 4.08, SD = 2.10, N = 162) than for non-borrowed words (M = 2.89, SD = 1.73, N = 23, $t(188)$ = 3.25, p = .001, d = .613), indicating that borrowings provide a record consistent with diversity calculated from cladistics studies.

In a linear model of the multiple psycholinguistic predictors of rate of lexical change, with grammatical category included as a random effect, we replicated the effect observed by Pagel et al. (2007) of a significant main effect of frequency on rate of lexical change (β = -.312, SE = .086, p < .001). When frequency, age of acquisition and length were entered, we found that word length was also significantly related to rate of change (β = .189, SE = .068, p = .006), as was age of acquisition (β = .174, SE = .071, p = .015), similar to Monaghan (2014). There was no significant quadratic (non-linear) effect for frequency, length nor age of acquisition (see supplementary materials S3), indicating that a linear, rather than a non-linear model fit was effective for predicting rate of change for the Swadesh list of words. Rates of change were higher for less frequent words, longer words, and words acquired later in language development (Figure 4). The effect of frequency for rates of lexical change corresponds to the result found for the relation between mid- to high-frequency words in the loanword database analysis, where increasing frequency related to lower levels of borrowing.
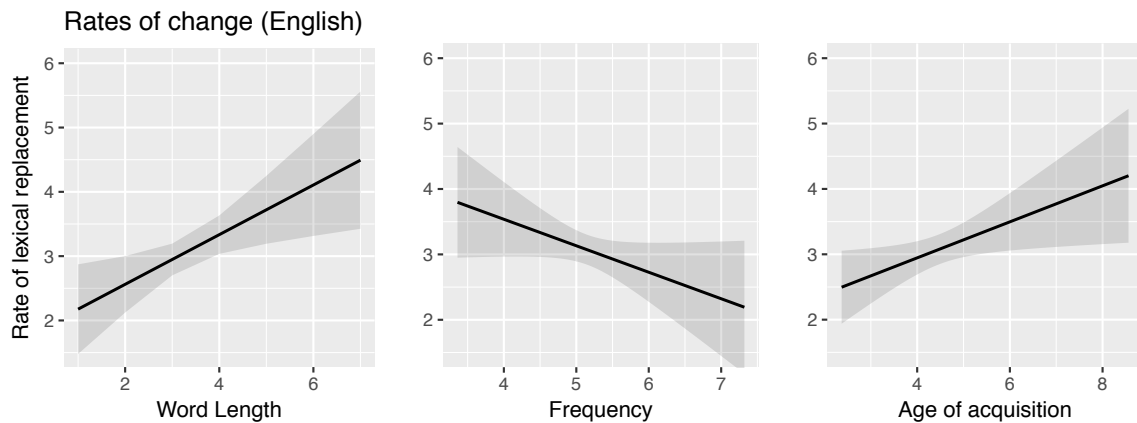
Figure 4. Model estimates for the relationship between rate of lexical change (vertical axis) and word length, frequency and age of acquisition.

The GAM analysis of probability of borrowing for the Swadesh word list resulted in only a significant effect of length (edf = 2.17, χ^2 = 9.42, p = 0.0272 (see Supplementary Materials S3 for full details), suggesting that for a binary variable the dataset was too limited to replicate effects shown in the lexical replacement data. Frequency was monotonically decreasing, demonstrating a similar but non-significant pattern to the lexical replacement analysis (edf = 1.00, χ^2 = 1.32, p = 0.252), age of acquisition showed no monotonic change (edf = 1.80, χ^2 = .006, p = 0.9360), and grammatical category was also not significant (edf < .001, χ^2 = 0.00, p = 0.6892).

Thus, the results of analyses of the smaller Swadesh set of words are less likely to represent the mechanisms of vocabulary change that are evident for low- to mid-frequency words, where increasing frequency related to higher levels of borrowing.

Study 4: Changes from Old English to Modern English

We have thus far presented analyses of word length, frequency, and age of acquisition as predictors of vocabulary change, which are derived from principles of how robustly individual word forms are represented in the speaker's vocabulary. For frequency and age of acquisition, it is a reasonable assumption that these factors remain relatively stable over time – certain concepts are acquired earlier, and used more frequently, independent of transient cultural influences (e.g., Baron et al., 2009; Li, Engelthaler, Siew, & Hills, in press; Sinclair, 1991).

However, we have also used length of contemporary forms of words as a predictor of stability, due to longer words being less accurately represented than shorter words in models of memory and lexical access (e.g., Davies, Arnell, Birchenough, Grimmond, & Houlson, 2017). This measure is somewhat problematic, though, because if longer words are found to be more likely to be borrowed, this may instead be a relic of the language from which the word is borrowed rather than being a driver of change in the vocabulary. For instance, if donor languages tend to contain longer words than the language to which the word is donated, then a borrowed word is more likely to be longer than original forms. To address this, we conducted an additional analysis on borrowing in English with length taken from Old English word forms. If length of the word form influences stability of that form, then longer words in Old English are more likely to be subsequently replaced by another word than shorter words in Old English.

Method

Data

The list of words was as for Study 1, with Old English word forms taken from the Intercontinental Dictionary Series (Key & Comrie, 2015). Frequency and age of

acquisition were the same as Study 1, measured from contemporary sources. Length was taken from the Old English form of the word. Words which were recorded as borrowed since 900CE were classified as borrowed, words which had not been borrowed since 900CE were classified as not borrowed. There were 1139 words in the database with all psycholinguistic predictors available, of which 403 had been borrowed since 900CE.

Analysis

We conducted GAM analysis on borrowing status since 900CE using the same procedure as for Study 1, with random effect of grammatical category, fixed effects for Old English word length, frequency, age of acquisition, and concreteness, and random slopes for the psycholinguistic variables.

Results and discussion

As for the analyses of English using contemporary word length, there was a significant effect of grammatical category (edf $= 5.27$, $\chi^2 = 32.6$, $p < 0.001$), frequency (edf $= 3.49$, $\chi^2 = 25.4$, $p < 0.0001$), age of acquisition (edf $= 1.00$, $\chi^2 = 33.3$, $p < 0.001$), and the effect of concreteness was also now significant (edf $= 2.56$, $\chi^2 = 18.3$, $p < .001$). There was no significant main effect of Old English length (edf $= 1.00$, $\chi^2 = .851$, $p = .356$), but there was a significant slope effect for Old English length by grammatical category (edf $= 2.31$, $\chi^2 = 41.0$, $p = .002$), indicating that length had a significant effect on borrowing status which varied according to different parts of speech. No other slope effects were significant (all $\chi^2 < 0.001$, $p > .456$).

Analyses of borrowings for each grammatical category separately showed that Old English length was significant and monotonically increasing for nouns (edf $= 1.00$, χ^2

= 13.01, $p < .001$), and adjectives (edf = 2.16, χ^2 = 8.31, $p = .025$), but was not significant for verbs (edf = 2.48, χ^2 = 4.08, $p = .238$). Other parts of speech could not be analysed because the number of items and range of lengths was too small. Finally, a GAM without random slopes for grammatical category showed that overall there was a significant effect of Old English word length (edf = 1.00, χ^2 = 11.27, $p < .001$, see supplementary materials S4). Thus, longer words in Old English are more likely to have been subsequently replaced with a borrowed form from another language over the last 1100 years of English language history. This is consistent with our hypothesis that shorter word length forms are more robustly represented in the speaker's vocabulary which results in the greater stability of the form in the speaker's language.

General Discussion

The analyses of a substantially larger dataset than that used in previous cladistics studies enabled us to test the extent to which observations of psycholinguistic variables relating to vocabulary change were confirmed in a larger and more representative vocabulary. The analyses in English largely confirmed previous psycholinguistic predictors of language change, with one important exception. Whereas effects of grammatical category, length (longer words are more likely to change, in terms of being borrowed into the language), and age of acquisition (later acquired words are more likely to be borrowed) showed similar effects to previous analyses (e.g., Monaghan, 2014), the effect of frequency was shown to be more nuanced than in the original analyses of rate of lexical change for 200 words (Pagel et al., 2007; Vejdemo & Hörberg, 2016). For higher frequency words, the results were consistent with these previous analyses: increasing

frequency related to reduced probability of borrowing. But for the majority of words in the language, which were lower to mid-range frequency, the effect of frequency was reversed: increasing frequency related to increased probability of borrowing.

The results in Dutch confirmed these analyses: grammatical category, length, age of acquisition and frequency were all significant predictors, with frequency demonstrating a similar non-linearity in relation to borrowing as that found in English.

Taken together, these results indicate that numerous factors relate to the likelihood of a word being borrowed. The observed effects are, first of all, consistent with accounts that suggest that grammaticalisation influences likelihood of borrowing. Open class words (such as nouns, verbs, and adjectives) are more likely to be borrowed than closed class words, such as pronouns, numerals, conjunctions, prepositions, or determiners. Interestingly, adverbs were also found to be less likely to be borrowed than other open word classes, a result that appears to be cross-linguistically valid (Tadmor, 2009). Such effects of grammatical category are consistent with greater difficulty of borrowing a word when it requires adaptation to the syntax and morphology of the recipient language (Myers-Scotton, 1993). For example, in Lithuanian, a language with a complex inflectional system, the incorporation of loan words into the language has been shown to be a gradual process, with forms varying in terms of whether they are fully, partially, or not at all integrated into the inflectional system (Rimkutė & Raižytė, 2014). The complexity of the grammaticalisation of a loan word may also vary within grammatical categories. For instance, an analysis of Czech, Russian, and Serbo-Croatian shows that loanwords ending in consonants can be entered into the morphological system for nouns within an established declension class (Thomas, 1983). Loanwords ending in

vowels, however, must be adjusted because no noun stems can end in a vowel, requiring adaptation of their form. A prediction arising from this observation, if grammaticalisation complexity is a driver of borrowing, would be that Slavic languages are less likely to borrow words ending in a vowel than those ending in a consonant.

The discrepancy between English and Dutch in terms of borrowings of verbs – they are more likely than other grammatical categories in English, but not in Dutch – may be due to the influence of more complex conjugation in Dutch than English. Dutch has distinct forms of conjugation for person (first, second, and third person), tense and mood, whereas English generally has only an inflected form for the third person singular, past tense, past participle, and a gerund or present participle form. Thus, there is larger system of grammaticalisation of verbs in Dutch than English that would have to be applied to a newly introduced verb form. However, without a broader cross-linguistic analysis, such comparisons between pairs of languages remain speculative.

The finding that several psycholinguistic predictors relate to vocabulary change suggest that cognitive processes of memory and learning exert a profound effect on diachronic language change. We contend that this is because words that are more salient and robustly represented in the language user's vocabulary are less likely to be prone to external forces of change – social, cultural, economic – on language structure. However, our analyses are not able to establish causality between these cognitive influences – and the psycholinguistic predictors that relate to the cognitive processes – and language change.

Nevertheless, it remains the case that earlier acquired words are more robustly represented in the vocabulary in terms of accuracy and speed of access (Juhasz, 2005).

This representational fidelity is a consequence of learning (Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2010) but continues to exert its effect throughout the lifespan (Hodgson & Ellis, 1998). As in Monaghan's (2014) analyses, the effect of age of acquisition applies not only to representation of a word in the individual's vocabulary, but also has implications for the robustness of the word's representation in the community's language (Baumann & Ritt, in press). These observations of acquisition influences on language stability are also consistent with Labov's (2007) suggestion that transmission requires replication of the care-giver's language, thus early acquired forms must be reproduced accurately in the learner's language, whereas incrementation occurs once the language structure has been acquired to a degree of accuracy. Thus, the pressure points for change in the language are more likely to be for later acquired vocabulary items.

The length of the word in English was also found to be a predictor of whether a word was borrowed, consistent with previous studies of lexical change based on cladistics studies (Monaghan, 2014; Pagel, 2007; Pagel et al., 2013). Again, this effect is consistent with models of memory and representational fidelity – it is easier to remember a short sequence than a long sequence (Baddeley, Thomson, & Buchanan, 1975), which then is more likely to be robustly established in the vocabulary of the learner. However, the effect of length may also be partially due to English words borrowed from Latin or French being longer than those that were original Germanic word forms (Reilly, Westbury, Kean, & Peelle, 2012). The analyses in Study 4 of Old English word length predicting whether a word is subsequently borrowed are not compatible with this explanation. Furthermore, the role of the source language is partly offset by the inclusion

of concreteness as a measure, which can partially account for the effect of loan words from French or Latin which tend to refer to cultural and technological terms which are generally more abstract (Lewis & Frank, 2016; Reilly & Kean, 2007). Though, note that concreteness did not reliably predict likelihood of whether a word was borrowed or not in English and so effects of whether a word was a cultural or technological innovation are likely to explain only limited deviance in probability of borrowing.

      The analyses of Dutch did not reveal a monotonic effect of word length, but found that mid-length words were more likely to be borrowed than shorter and longer words. The robustness of representation might then be a combination of length in phonemes and morphological complexity. In Dutch compared to English, longer words are more likely to be multimorphemic, hence composed of morphemes that individually can be represented with fidelity, but also exerting a grammaticalisation effect on the lexical root which has to be adapted by the morphological rules of the language. Thus, longer, multimorphemic words were less likely to be borrowed than a medium length word, either because of the increased grammaticalisation required for attachment of suffixes, or because the individual morphemes were short and more easily memorable than a medium length word. Further explorations are required to decide between these accounts. Note, however, that these analyses of word length were based on contemporary word forms in Dutch. Unlike in English, no corpus was available to determine the length of words historically, but we assume that the consistency between analyses in English, Old English, and Dutch (for monomorphemic words) indicates that length is operating as predicted across these languages in terms of affecting representational robustness and stability of the word form.

The effect of frequency in the analyses was largely consistent across English and Dutch, but distinct from previous smaller-scale analyses of the relation between frequency and vocabulary change. Previously, higher-frequency words were found to be more resilient to change in cladistic studies of vocabulary change (Pagel et al., 2007). For the substantially larger set of words analysed here, only the comparison between mid-frequency and high-frequency words replicated the negative linear relationship between frequency and change, with higher frequency resulting in lower rates of change. However, when the range of lower-frequency words was represented more effectively as it was in the loanword database analysis, the effect of frequency between low- and mid-frequency words was in the reverse direction. Lower frequency words were less likely to be borrowed than mid-frequency words. Note that this discrepancy in the results was not due to borrowing data being a very different measure of change than rate of lexical change in Pagel et al.'s (2007) analyses, as there was a significant relation between these measures: borrowed words tended to have greater rates of lexical change, as shown in Study 3.

Taken together, these psycholinguistic predictors of borrowing suggest that more robustly represented vocabulary items, as reflected in psycholinguistic predictors of speed and accuracy of vocabulary access (e.g., Brysbaert & Ghyselinck, 2006), highlight cognitive principles explaining processes in diachronic change of the vocabulary. But these effects are, in some cases, modulated by other factors influencing change. The analyses of the larger data set provide ways by which these multiple forces may be distinguished. For instance, a word is less likely to be replaced by a word from another language if it is very high frequency. In these cases, there is an overlap between

explanations based on representational fidelity of the word and social-usage models of

language change, whereby misunderstandings will be more frequent if speakers use

distinct words for the same concept (Boyd & Richerson, 1985; Bybee, 2007; Pagel et al.,

2013). However, if the word is not high frequency, then borrowings become more likely

if the speaker is using the concept more than occasionally in speech. Incrementation of

the language (Labov, 2007) appears to be licensed for words that are not the most

common in the language, but that are still in regular use, either by increasing the chances

of exposure to an alternative form from a speaker of another language or through

innovation within a group of speakers.

How general are the results to other languages? There are suggestions that some

languages may demonstrate more borrowings with basic words than more complex

words. In several Southeast Asian languages, there are observations of numerals and

pronoun borrowing (Cooke, 1968; Thomason & Everett, 2001), which are parts of speech

less likely to be borrowed in English. However, these languages tend to have very large

sets of pronouns, and it may be that the less frequent pronouns are the ones that come to

be borrowed (Tadmor, 2007), thus the results may be consistent with the principles of the

influence of psycholinguistic effects presented for English and Dutch. One particular

example is Ceq Wong (Kruse, 2009), which demonstrates high levels of borrowing of

basic vocabulary from Malay. Tadmor (2007) suggests that this could be due to

assimilation of the words that Ceq Wong speakers knew in Malay, with these comprising

the basic vocabulary items, though this is curious because of the lack of cultural influence

from Malay on this society (Kruse, 2009). Such an explanation is consistent with

innovations in vocabulary resulting from exposure to novel items – the basic vocabulary

items cannot be borrowed unless they are heard. However, Kruse's (2009) analysis of Ceq Wong still demonstrates that a higher proportion of nouns than function words are borrowed, but an equal number of verbs and function words, which is indeed a similar pattern to that observed for Dutch (see Table 2).

Thus, it is too early to say whether languages exist where the psycholinguistic effects established for English and Dutch, based on cognitive principles of learning and representation, may be consistent or inconsistent with other languages. What is needed is the construction of reliable psycholinguistic variables in each language. For Ceq Wong, there is a current absence of both frequency and age of acquisition measures, and these measures are necessary in order to determine the generality of the results for English and Dutch, and the extent to which societal influences may over-rule the proposed cognitive principles driving lexical stability and change. But the observed frequency effects in English and Dutch for whether a word is borrowed or not (i.e., mid frequency words are most likely to be borrowed) may indeed be consistent with views of vocabulary assimilation that are at least in part driven by usage – those words frequently encountered from other dialects may be the ones that are more likely to be incorporated into the language.

These results require a reconsideration of accounts of language evolution that claim that change is driven by the process of acquisition (Bickerton, 1990). The age of acquisition results demonstrate that acquisition is implicated in evolution of the vocabulary, but actually as a preserving influence rather than a driver of change. The process of acquisition, then, is much more likely to establish stability in the vocabulary. The alternative theory that change in language results from innovations by proficient

speakers of the language (Bybee & Slobin, 1982; Joseph, 1992; Labov, 2007; Slobin, 2005; Smith & Durham, 2009) is more consistent with the current psycholinguistic results, but with the clarification that innovations are constrained by cognitive processes determining representations of the vocabulary.

Nevertheless, the extent to which acquisition prevents rather than drives change may be constrained to evolution of only certain language levels, such as vocabulary (Monaghan, 2014) or sublexical structure (Baumann & Ritt, 2018). The role of acquisition may be very different for evolution of morphology or syntax, where the process of acquisition appears to exert a profound influence on syntactic structural change (Rische & Komorova, 2016; Senghas & Coppola, 2001). Indeed, it has been shown that syntactic restructuring can emerge quite independently of changes to individual word forms (Coppola & Newport, 2005), indicating that the processes influencing vocabulary change may be very different than those applying to syntax or discourse (Haselow, 2018). Such a view is consistent with separations in speech production models between vocabulary access and syntactic constructions (e.g., Levelt, 1989) and this evolutionary approach can thereby provide an additional set of data that can highlight distinctions in the mechanisms of implicit memory associated with syntax and explicit memory associated with vocabulary learning (Ullman, 2004). Clearly, a full account of language evolution will require a more complex account of how change driven by acquisition or by proficient speaker innovations operate at different levels of language structure.

Acknowledgements

References

Atkinson, Q. D. (2011). Phonemic diversity supports serial founder effect model of language expansion from Africa. *Science, 332*, 346-349.

Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. J., & Pagel, M. (2008). Languages evolve in punctuational bursts. *Science*, *319*(5863), 588-588.

Baayen, R., Pipenbrock, R., and Gulikers, L. (1995). The CELEX lexical database (Version 2) [CD-ROM]. University of Pensylvania, Linguistic Data Consortium, Philadelphia.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575-589.

Baron, A., Rayson, P., & Archer, D. (2009). Word frequency and key word statistics in historical corpus linguistics. *Anglistik*, *20*(1), 41-67.

Baumann, A., & Ritt, N. (2018). The basic reproductive ratio as a link between acquisition and change in phonotactics. *Cognition*, 176, 174-183.

Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Borin, L., Comrie, B., & Saxena, A. (2013). The Intercontinental Dictionary Series - a rich and principled database for language comparison. In L. Borin, & A. Saxena (Eds.), *Approaches to Measuring Linguistic Differences* (pp. 285-302). Berlin: Mouton de Gruyter.

Boyd, R. & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition, 13*(7-8), 992-1011.

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica, 150*, 80-84.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*, 904–911.

Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods, 49*, 1520-1523.

Bybee, J. L. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, J. L., & Slobin, D. I. (1982). Why small children cannot change language on their own: Suggestions from the English past tense. In A. Ahlqvist (Ed.), Papers from the 5th International Conference on Historical Linguistics (pp. 29–37). Amsterdam/Philadelphia: John Benjamins.

Buck, C. D. (1949). *A dictionary of selected synonyms in the principal Indo-European languages*. Chicago, IL: University of Chicago Press.

Cassidy, F. G. (1999). English: A Germanic language? In Carr, G. F., Harbert, W., & Zhang, L. (Eds.), *Interdigitations: Essays for Irmengard Rauch* (pp. 75-79). New York: Peter Lang.

Coppola, M., & Newport, E. L. (2005). Grammatical subjects in home sign: Abstract

linguistic structure in adult primary gesture systems without linguistic input.

Proceedings of the National Academy of Sciences of the United States of America,

102(52), 19249-19253.

Corballis, M. C. (2017). Language evolution: a changing perspective. *Trends in Cognitive*

Sciences, *21*, 229-236.

Cooke, J. R. (1968). *Pronominal reference in Thai, Burmese, and Vietnamese*. Berkeley:

University of California Press.

Croft, W. (2000). Explaining language change: An evolutionary approach. Harlow:

Pearson Education.

Davies, R. A., Arnell, R., Birchenough, J. M., Grimmond, D., & Houlson, S. (2017).

Reading through the life span: Individual differences in psycholinguistic effects.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *43*(8),

1298-1338.

Dyen, I., Kruskal, J. B., & Black, P. (1992). An Indo-European classification, a

lexicostatistical experiment. *Transaction of the American Philosophical Society, 82*,

1–132.

Embleton, S. (1986). *Statistics in historical linguistics.* Bochum: Brockmeyer.

Gilhooly, K. J., & Gilhooly, M. L. M. (1980). The validity of age-of-acquisition ratings.

British Journal of Psychology, *71*(1), 105-110.

Grant, A. (2009). Loanwords in British English. In Haspelmath, M. & Tadmor, U. (Eds.),

Loanwords in the world's languages: A comparative handbook (pp.360-383). Berlin:

De Gruyter Mouton.

Gray, R. D., & Atkinson, Q. D. (2003). Language-tree divergence times support the

Anatolian theory of Indo-European origin. *Nature*, *426*(6965), 435-439.

Haselow, A. (2018). Language change from a psycholinguistic perspective: The long-

term effects of frequency on language processing. *Language Sciences, 68,* 56-77.

Haspelmath, M. & Tadmor, U. (2009). *World loanword database*. Leipzig: Max Planck

Institute for Evolutionary Anthropology. http://wold.clld.org/vocabulary/13,

Accessed on 2016-09-16.

Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In Haspelmath, M., &

Tadmor, U. (Eds.)*, Loanwords in the world's languages: A comparative handbook*

(pp.35-54)*.* Berlin: De Gruyter Mouton.

Haspelmath, M., & Tadmor, U. (2009). The loanword typology project and the world

loanword database. In Haspelmath, M., & Tadmor, U. (Eds.), *Loanwords in the*

world's languages: A comparative handbook (pp.1-34). Berlin: De Gruyter Mouton.

Haugen, E. (1950). The analysis of linguistic borrowing. *Language, 27*, 210-231.

Hodgson, C. & Ellis, A.W. (1998). Last in, first to go: Age of acquisition and naming in

the elderly. *Brain and Language, 65*, 146-163.

Holmes, S. J., Fitch, F. J., and Ellis, A. W. (2006). Age of acquisition affects object

recognition and naming in patients with Alzheimer's disease. *Journal of Clinical and*

Experimental Psychology, 28, 1010-1022.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The

roles of adult and child learners in language formation and change. *Language*

Learning and Development, *1*(2), 151-195.

Joseph, B. D. (1992). Diachronic explanation: Putting speakers back into the picture. In

    G. Davis & G. Iverson (Eds.), Explanation in historical linguistics (Current Issues in

    Linguistic Theory, 84) (pp. 123–144). Amsterdam: John Benjamins.

Juhasz, B.J. (2005). Age-of-acquisition effects in word and picture identification.

    Psychological Bulletin, 131, 684-712.

Kempe, V., Gauvrit, N., & Forsyth, D. (2015). Structure emerges faster during cultural

    transmission in children than in adults. *Cognition*, *136*, 247-254.

Keuleers, E., Brysbaert, M. & New, B. (2010). SUBTLEX-NL: A new frequency

    measure for Dutch words based on film subtitles. *Behavior Research Methods, 42*(3),

    643-650.

Key, M. R., & Comrie, B. (2015). *The intercontinental dictionary series.* Leipzig: Max

    Planck Institute for Evolutionary Anthropology. http://ids.clld.org, Accessed on

    2016-09-16.

Kilgarriff, A. (1997). Using word frequency lists to measure corpus homogeneity and

    similarity between corpora. *Proceedings of the 5th ACL workshop on very large*

    corpora, 231-245.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition

    ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978-990.

Labov, W. (2001). *Principles of linguistic change, vol. 2: Social factors*. Oxford:

    Blackwell.

Labov, W. (2007). Transmission and diffusion. *Language*, *83*(2), 344-387.

Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., ... & Falush, D. (2015). The fine-scale genetic structure of the British population. *Nature*, *519*, 309-314.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Lewis, M. & Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition, 153*, 182-195.

Li, Y., Engelthaler, T., Siew, C. S. Q., & Hills, T. T. (in press). The Macroscope: A tool to examine the historical structure of language. *Behavior Research Methods, in press*.

Lieberman, E., Michel, J. B., Jackson, J., Tang, T., & Nowak, M. A. (2007). Quantifying the evolutionary dynamics of language. *Nature, 449*(7163), 713-716.

Matras, Y., & Sakel, J. (2007). *Grammatical borrowing in cross-linguistic perspective.* Berlin: Mouton de Gruyter.

Monaghan, P. (2014). Age of acquisition predicts rate of lexical evolution. *Cognition, 133,* 530-534.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology: Section A*, *50*(3), 528-559.

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching.* Oxford: Clarendon.

Newberry, M. G., Ahern, C. A., Clark, R., & Plotkin, J. B. (2017). Detecting evolutionary forces in language change. *Nature, 551*, 223-226.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11-28.

Pagel, M., Atkinson, Q. D., Calude, A. S., & Meade, A. (2013). Ultra conserved words point to deep language ancestry across Eurasia. *Proceedings of the National Academy of Sciences, 110*(21), 8471–8476.

Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature, 449*(7163), 717–720.

Reilly J., & Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: Analyses and theoretical implications. *Cognitive Science, 31*, 1–12.

Reilly, J., Westbury, C., Kean, J., & Peelle, J. E. (2012). Arbitrary symbolism in natural language revisited: when word forms carry meaning. *PLoS One*, *7*(8), e42286.

Rische, J. L., & Komarova, N. L. (2016). Regularization of languages by adults and children: A mathematical framework. *Cognitive Psychology*, *84*, 1-30.

Rimkutė, E., & Raižytė, J. (2010). Morphological adapting of loanwords in Lithuanian. *Lietuvių kalba*, (4). Retrieved from http://www.lietuviukalba.lt/index.php/lietuviu-kalba/issue/view/109, August 11th 2018.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, *12*(4), 323-328.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Slobin, D. I. (2005). From ontogenesis to phylogenesis: What can child language tell us about language evolution? In J. Langer, S. T. Parker, & C. Milbrath (Eds.), *Biology and knowledge revisited: From neurogenesis to psychogenesis* (pp. 255–285). Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, J., & Durham, M. (2011) A tipping point in dialect obsolescence? Change across

    the generations in Lerwick, Shetland. *Journal of Sociolinguistics, 15*, 197-225.

Steels, L. (2017). Human language is a culturally evolving system. *Psychonomic Bulletin*

    and Review, *24*(1), 190-193.

Strain, E., Patterson, K. E., & Seidenberg, M. S. (1995). Semantic influences on word

    recognition. *Journal of Experimental Psychology: Learning, Memory, and*

    Cognition, 21, 1140–1154.

Swadesh, M. (1952). Lexico-statistic dating of prehistoric ethnic contacts: with special

    reference to North American Indians and Eskimos. *Proceedings of the American*

    Philosophical Society, 96, 452–463.

Swadesh, M. (1955). Towards greater accuracy in lexicostatistic dating. *Linguistics, 21*,

    121-137.

Tadmor, U. (2007). Is borrowability borrowable. *Paper presented at the Symposium for*

    Language Contact and the Dynamics of Language. Max Plack Institute for

    Evolutionary Anthropology, Leipzig, May 2007.

Tadmor, U. (2009). Loanwords in the world's languages: Findings and results. In

    Haspelmath, M., & Tadmor, U. (Eds.), *Loanwords in the world's languages: A*

    comparative handbook (pp.55-75). Berlin: De Gruyter Mouton.

Thomas, G. (1983). A comparison of the morphological adaptation of loanwords ending

    in a vowel in contemporary Czech, Russian, and Serbo-Croatian. *Canadian Slavonic*

    Papers, *25*(1), 180-205.

Thomason, S. G., & Everett, D.L. (2001). Pronoun borrowing. *Berkeley Linguistic*

    Society, 27, 301-315.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231-270.

van der Sijs, N. (2009). Loanwords in Dutch. In Haspelmath, M., & Tadmor, U. (Eds.), *Loanwords in the world's languages: A comparative handbook* (pp.338-359)*.* Berlin: De Gruyter Mouton.

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representations in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193–211.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*, 1176–1190.

Vejdemo, S., & Hörberg, T. (2016). Semantic factors predict the rate of lexical replacement of content words. *PLoS ONE, 11(1)*, e0147924.

Walker, I. & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1256-1271.

Wilson, E. B. (1927). "Probable inference, the law of succession, and statistical inference". Journal of the American Statistical Association. 22: 209–212.

Winter, B., Thompson, G., & Urban, M. (2014). Cognitive factors motivating the evolution of word meanings: Evidence from corpora, behavioral data and

encyclopedic network structure. In *Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)* (pp. 353-360).

Wood, S.N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society B, 73(1)*, 3-36.