Iconicity and Diachronic Language Change

Padraic Monaghan,[a,b] Seán G. Roberts[c]

[a]*Department of Psychology, Lancaster University, Lancaster, UK*
[b]*Amsterdam Center for Language and Communication, University of Amsterdam, Amsterdam, The Netherlands*
[c]*School of English, Communication and Philosophy, Cardiff University, Cardiff, UK*

Abstract

Iconicity, the resemblance between the form of a word and its meaning, has effects on behavior in both communicative symbol development and language learning experiments. These results have invited speculation about iconicity being a key feature of the origins of language, yet the presence of iconicity in natural languages seems limited. In a diachronic study of language change, we investigated the extent to which iconicity is a stable property of vocabulary, alongside previously investigated psycholinguistic predictors of change. Analyzing 784 English words with data on their historical forms, we found that stable words are higher in iconicity, longer in length, and earlier acquired during development, but that the role of frequency and grammatical category may be less important than previously suggested. Iconicity is revealed as a feature of ultra-conserved words and potentially also as a property of vocabulary early in the history of language origins.

Keywords: Age of acquisition; Frequency; Grammatical category; Iconicity; Language evolution; Psycholinguistics; sSound symbolism

1. Introduction

When asked to produce a novel sign to refer to an object or action, participants tend to use a form for the sign that resembles one or more aspects of its meaning (Perniss, Thompson, & Vigliocco, 2010). This resemblance, or "iconicity" (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015), between communicative symbol and meaning has been found

---

in a range of behavioral tasks. Iconicity has been found in the forms produced by participants for linguistic constructions (Meir, Padden, Arono, & Sandler, 2013), oral productions that are restricted to be nonlinguistic, that is, not containing phonemes (Perlman, Dale, & Lupyan, 2015), and even for auditory signs produced by a whistle (Verhoef, Kirby, & de Boer, 2016). Iconicity is also evident in spoken natural language, as demonstrated in investigations of sound symbolism (Dingemanse, Schuerman, Reinisch, Tufvesson, & Mitterer, 2016; Lockwood & Dingemanse, 2015; Ohala, 1994; Westbury, Hollis, Sidhu, & Pexman, 2018). A prime example is onomatopoeia, where words for animal calls, for instance, reflect the auditory properties of the animal's sound. Yet iconicity extends well beyond the boundaries of onomatopoeia, and it has been proposed to widely suffuse the vocabulary (Dingemanse, 2012; Dingemanse et al., 2015; Perry, Perlman, & Lupyan, 2015; Winter, Perlman, Perry, & Lupyan, 2017).

Experiments investigating production of iconicity in communicative studies have been taken as providing insight into the origins of communication (e.g., Ramachandran & Hubbard, 2001). A proposal emerging from this literature is that systems of communication— such as the origins of language—begin with sets of iconic representations, only gradually becoming eroded as communication develops. In this paper, we investigate whether there is direct evidence for iconicity in natural language evolution, by analyzing a diachronic corpus of vocabulary forms, determining the extent to which iconicity relates to language stability and change.

Studies of iconicity in natural language have investigated the extent to which speakers judge resemblance between a word's sound and its meaning. Perry et al. (2015) and Winter et al. (2017) asked participants to make decisions about the extent to which a word's sound fitted its meaning, which was their implementation of iconicity. Perry et al. (2015) found variation in participants' judgments of the iconicity of words, with words high in iconicity more likely to be those that were earlier acquired during language development, and low iconicity rated words more likely to be later acquired. Thus, iconicity may be especially useful to assist children in acquiring their first words.

Iconicity ratings for words potentially draw on sensory correspondences between the sound of the word and its meaning (Perry et al., 2015, Perry, Perlman, Winter, Massaro, & Lupyan, 2018; Winter et al., 2017). However, this does not mean that iconicity can be reduced to concepts of concreteness or imageability of the word. Winter et al. (2017) found that iconicity, though related to concreteness, was divergent, with some abstract concepts in the language having high iconicity, and concrete concepts expressed by words with low iconicity.

These studies show that iconicity can certainly be found in subsets of natural language vocabulary (Perry et al., 2015), yet the evidence for iconicity in the broader vocabulary seems to be limited. Monaghan, Shillcock, Christiansen, and Kirby (2014) investigated the extent to which similarities between sounds of words related to similarities between meanings of words in a representative sample of the vocabulary of English. This statistical correspondence between sound and meaning is referred to as systematicity (Dingemanse et al., 2015), and though not identical to iconicity, it ought to be related. If words which sound similar also have similar meaning and words which sound distinct have unrelated meanings, then there is higher systematicity in the vocabulary. Whereas, if the extent to which words are similar

or distinct in sound does not relate to whether those words are similar or distinct in meaning, then the vocabulary is not systematic. If iconicity was found to be prevalent throughout the vocabulary, then this ought to be reflected in correspondences between the sound space and the meaning space of language: If there are widespread resemblances between sounds and meanings, then similar sounding words ought to relate to similar meaning concepts. Monaghan et al. (2014) measured this systematicity in the vocabulary of English and found that, though the systematicity in the vocabulary was greater than expected by chance, the sounds of words explained only a very small amount of variance in the meanings of words. Dautriche, Mahowald, Gibson, and Piantadosi (2017) confirmed this effect for English and found a similar level of systematicity across 99 other languages: Natural language vocabularies are systematic, but only just.

Taken together, these studies raise the question: If iconicity is instantiated in early communicative signs, and is advantageous for learning the communicative system, why is iconicity not more prevalent in natural language? There are several possible explanations for how sound symbolism—whether iconic or systematic—may decline as language systems change. Ahlner and Zlatev (2010) suggested that communicative systems may begin with iconicity in the signification of meaning but that processes of conventionalization reduce this iconicity, introducing greater compositionality (Kirby, Tamariz, Cornish, & Smith, 2015) and abstraction of the signs (Fay, Ellison, & Garrod, 2014; Senghas & Coppola, 2001). For example, Nölle, Staib, Fusaroli, and Tylén (2018) showed that participants who were required to communicate via silent gestures developed systems that involved systematicity and compositionality rather than iconicity in the signs that they used, though iconicity was more likely when there was a smaller, more predictable set of referents for the signs.

Another driver to reduce iconicity between form and meaning in the vocabulary is the need for efficient communication between speaker and hearer (Gibson et al., 2019). For the speaker, iconicity may well be an effective means by which to generate a sign for an intended referent, but decoding the sign is subject to different constraints. In order to ascertain the speaker's intended meaning, the hearer must determine which of the possible set of referents is being referred to by the speaker. As referents with related meanings tend to co-occur in language and in the environment (Landauer & Dumais, 1997), this means that if there is a similarity in the sound of words with similar meanings, then the sound of the word provides less information to distinguish the possible referent (Monaghan, Christiansen, & Fitneva, 2011). Monaghan et al. (2011) showed that, for English and French, there is greater systematicity toward the end of words and more distinctive information toward the beginning of the word that can support efficient word identification. The growing vocabulary exerts greater pressure on the forms of the language to produce distinctive forms for concepts that are similar in meaning (Brand, Monaghan, & Walker, 2018), resulting in a reduction in iconicity as the vocabulary of the language community expands. Furthermore, reducing iconicity permits greater expressive freedom in the language for expanding to abstract terms which cannot be bound to sensation in the same way as more concrete concepts (Lupyan & Winter, 2018).

Monaghan et al. (2014) predicted that words that are learned earlier in life are likely to be more systematic than those learned later in life, because when the vocabulary is small the need to distinguish forms is less because the meaning space is less densely populated

(Gasser, Sethuraman, & Hockema, 2010). This finding from natural language was corroborated in an artificial language learning task (Brand et al., 2018): Iconicity in word forms was advantageous in the early stages of learning a small vocabulary, but the advantage diminished as the vocabulary grew.

Thus, iconicity provides advantages for processing and acquisition of language, yet there are substantial pressures on the vocabulary to ensure efficient communication that push against iconicity. Nevertheless, despite these pressures against iconicity, all else being equal, if iconicity is a crucible of early communicative systems, then it ought to be observed to some degree in contemporary vocabulary structure. Furthermore, if iconicity supports acquisition, then it ought to be observed as a property of the language that is resistant to change. Whereas systematicity drives against communicative efficiency (Monaghan et al., 2011), it is possible for iconicity to be present in the vocabulary without necessarily resulting in similarity of forms. Thus, words with similar meanings can have iconic forms with the iconicity carried in different aspects of the signal. For instance, for size, front vowels and unvoiced consonants, and frication have all been shown to relate to smaller size referents (Klink, 2000; Knoeferle, Li, Maggioni, & Spence, 2017; Lockwood & Dingemanse, 2015; Monaghan & Fletcher, 2019; Nichols, 1971; Ohala, 1994; Ultan, 1978), and so for words relating to meanings associated with small, iconicity could be in the vowel quality for one word, and the consonant manner for another word, resulting in little systematicity (and confusability) but maintaining iconicity.

Perry et al. (2015) included a measure of systematicity alongside iconicity in their study of a large set of words in English and showed that iconicity was strongly related to age of acquisition independent from the systematicity of the form–meaning relationships in the vocabulary. Similarly, Perry et al. (2015) found evidence of iconicity relating to children's expressive vocabularies in the first few years of their language acquisition, again independent of systematicity of forms. These studies open the door to the possibility that iconicity may be a property of early communicative systems, supportive for language transmission and resistant to pressures of language change.

Studies of diachronic lexical change have uncovered the features of the vocabulary that result in stability of change in forms. Pagel, Atkinson, and Meade (2007) examined the list of 200 basic vocabulary items from the Swadesh word lists (Swadesh, 1952) and estimated the rate of lexical change for the forms referring to each of these meanings by comparing the extant forms across a range of languages in the Indo-European language family. The idea of this approach is that words that change more rapidly are those where a greater diversity of forms is found across these languages. Pagel et al. (2007) found that higher-frequency words were less likely to change than lower-frequency words. In follow-up analyses, Monaghan (2014) found that earlier acquired and shorter words were also less likely to change, and Vejdemo and Hörberg (2016) discovered that words with fewer different senses, with more synonyms, and with lower imageability were also more likely to change.

Recently, Monaghan and Roberts (2019) extended these small-scale studies to investigate the contributors to lexical change in terms of words that are borrowed into a language. There are three contexts in which a word can be borrowed: as a replacement for a pre-existing form (such as the Old French derived *autumn* replacing Old English *hærfest* from Proto-Germanic

*harbitas); as a form that coexists with an existing form (such as *baby* recorded in English in the late 14th century to exist alongside *child* from the Proto-Germanic *kiltham); or as an insertion (such as *citrus*, from Latin, which was first recorded in English in the 19th century) which conveys a novel meaning that did not previously exist in the language. Whether a word is borrowed or not thus highlights the extent to which it is stable within the language: Overall, if a word is classified as not borrowed, then this indicates that it has undergone less change within the language. Approximately 1500 words from the World Loan-Word Database (WOLD, Haspelmath & Tadmor, 2009) were analyzed in order to investigate which psycholinguistic properties of words related to the probability of a word being borrowed into the language. Monaghan and Roberts (2019) confirmed the results from rate of lexical change studies showing that shorter length and earlier age of acquisition both related to lower probability of the word being borrowed, and thus greater resistance to change. For frequency, mid-range frequency was least resistant to borrowing, with higher frequency and lower frequency words less likely to be borrowed, providing a more nuanced indication of effects of frequency than that reported in Pagel et al. (2007) study.

In the current study, we extend the investigation of loan words to determine whether iconicity is also a property of the language that results in resistance to change. If so, then this provides converging evidence that iconicity is a stable property of signs used in communication and, despite processes of conventionalization, proves resistant to alteration. If iconicity is found to relate to words that are unchanged, then this increases the likelihood that words with high iconicity were present in earlier stages of language evolution. We also provide novel analyses of multiple psycholinguistic predictors of lexical stability that control collinearity to ascertain the relative strength of these predictors in determining stability or change of words' forms.

2. Method

2.1. Corpus preparation

Our key aim was to test the effect of iconicity in predicting borrowing of words in the WOLD (Haspelmath & Tadmor, 2009). We focused on the English word set which comprised a list of 1,515 words compiled by Grant (2009). These words originated from the Intercontinental Dictionary Series (Key & Comrie, 2015), which were selected to provide a set of core concepts that are verbalized across most languages. The WOLD database indicates which words are borrowed and which have no evidence of borrowing. Only entries that were single words in English (e.g., omitting "lightning bolt" and "fishing line") were included.

As in Monaghan and Roberts (2019), we gathered information on a set of psycholinguistic properties for each word in order to test and control for properties of words in addition to their iconicity. Frequency was taken from the Zipf SUBTLEX-UK database (van Heuven, Mandera, Keuleers, & Brysbaert, 2014), which is an effective measure of spoken word frequency. We also gathered information on grammatical category (noun, verb, adjective, adverb, determiner, or number) from Brysbaert, Warriner, and Kuperman (2014), and derived concreteness from the same database. Phonological length of the word was derived from the

CELEX database (Baayen, Pipenbrock, & Gulikers, 1995) with words that did not appear in CELEX hand-coded for phonological form. Diphthongs and affricates were encoded as single phonemes in the length measure. Age of acquisition was taken from Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012).

Iconicity was taken from Perry et al. (2018) and was a measure of participants' judgments of the extent to which the word sounds like its meaning. Judgments were made on an 11 point scale from -5 ("words that sound like the opposite of what they mean") to $+5$ ("words that sound like what they mean").

As the dependent variable, we encoded whether the word was classified as "clearly borrowed" or "no evidence of borrowing." Intermediate judgments from the WOLD (such as "possibly borrowed") were omitted from the analysis as we wanted to focus only on those words where there was clear evidence of borrowing or not.

There were a total of 784 words with values for all the psycholinguistic variables (frequency, grammatical category, phonological length, age of acquisition, concreteness, and iconicity) of which 296 were loanwords and 488 were classified as not borrowed into the language.

2.2. Analysis

The relation between each of the psycholinguistic variables and the probability of borrowing was determined using general additive models (GAMs). GAMs were employed because the relations between several of the psycholinguistic variables (frequency, age of acquisition, and length) and probability of borrowing had previously been shown to be nonlinear (Monaghan & Roberts, 2019). Key interactions were tested, but not included as they did not contribute significantly to model fit (see Supplementary Analyses accompanying the data archive). GAMs are nonlinear models which minimize the nonlinearity required to best fit the data (Wood, 2011). The way each predictor relates to borrowing can be assessed by two sets of measures: a set of test statistics that indicate how well the variable predicts the probability of borrowing (estimated degrees of freedom, a χ^2 statistic comparing the smooth term coefficients to zero, and associated *p*-value, see Marra & Wood, 2012; Wood, 2013), and a measure of the nonlinearity of the fit of the independent variable to the probability of borrowing (EDF). An EDF value close to 1 indicates a linear relation between the psycholinguistic variable and likelihood of borrowing, a value exceeding 1 indicates nonlinearity. For the GAM analyses, we used the R package *mgcv* (Wood, 2011) to fit a binomial GAM predicting whether a word was borrowed or not. The psycholinguistic measures (frequency, AoA, length, concreteness, and iconicity) were scaled and centered and entered as smooth thin plate regression spline predictors. Following Monaghan and Roberts (2019), the model included random slopes for parts of speech (penalized by a ridge penalty) and interactions between part of speech and each psycholinguistic variable.

We further conducted an analysis of the derivatives and standard errors for the derivatives along the nonlinear slopes for each psycholinguistic variable to determine at which point of the model fit the slope is significant. The gradient of the slope is significant if the confidence interval for the derivates does not overlap with zero.

Finally, to address potential collinearity between some of the predictors, we used a decision tree and random forests analysis. Decision trees are machine learning tools that find optimal ways of dividing the data in order to predict the target variable, similar to a game of 20-questions (Strobl, Malley, & Tutz, 2009; for applications in linguistics, see Bürki, Alario, & Frauenfelder, 2011; Roberts, Torreira, & Levinson, 2015; Tagliamonte & Baayen, 2012). They can help identify which factors are more decisive in determining the probability of borrowing or find exceptions to general rules. Branches in the tree are recursively added as long as the split produces a significant difference in the target variable. Therefore, if there are no significant effects, there will be no branches in the tree.

Random forests is a method of producing many decision trees based on subsets of the data and predictors in order to assess the robustness of the decision tree (Breiman, 2001). "Importance values" are calculated to assess how relatively decisive each predictor is among the "forest" of other predictors. Importantly, these measures are immune to collinearity between predictors, providing a robust test of the independence of the predictor effects. The R package *party* (Hothorn, Bühlmann, Dudoit, Molinaro, & Van Der Laan, 2006; Hothorn, Hornik & Zeileis, 2006; Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008; Strobl, Boulesteix, Zeileis, & Hothorn, 2007) was used, predicting the probability of borrowing from length, age of acquisition, frequency, concreteness, iconicity, and grammatical category.

The data, R script of the analyses, and analysis results are available at https://osf.io/8mr6d

3. Results

We first computed correlations between the psycholinguistic variables for the set of 784 words in the current analyses. The results are shown in Table 1.

In previous studies with larger sets of words, iconicity has been shown to relate to concreteness, with more iconic words being more concrete (Winter et al., 2017), and also to age of acquisition, with more iconic words tending to be acquired earlier (Perry et al., 2015). However, these direct relations were not found in the current set of words, nor for the subset of nouns in the current data set (Table 2), where iconicity related only to frequency. Supplementary Analyses demonstrated that this discrepancy was due to the larger set of grammatical categories included in our analyses compared to those of Perry et al. (2015) and

Table 1

Correlations between psycholinguistic variables for the 784 words included in the analysis

	AoA	Length	Concreteness	Iconicity
Frequency	$-.482$***	$-.335$***	$-.497$***	$-.050$
AoA		$.243$***	$-.051$	$-.039$
Length			$.097$**	$-.021$
Concreteness				$-.038$

Note. ** $p < .01$; *** $p < .001$.

Table 2

Correlations between psycholinguistic variables for the 460 nouns included in the analysis

	AoA	Length	Concreteness	Iconicity
Frequency	−.501***	−.231***	−.208***	−.120**
AoA		.196***	−.230***	.033
Length			−.051	.009
Concreteness				−.037

Note. ** $p < .01$; *** $p < .001$.

Table 3

Results of the GAM for probability of borrowing with psycholinguistic variables as predictors (EDF, reference degrees of freedom for the $\chi 2$ test, $\chi 2$ value, and associated p-value)

	EDF	Ref.df	$\chi 2$	p	GAM Model % of Variance Explained (pseudo-R^2)	Random Forests Relative Importance (x1000)
Frequency	2.191	2.813	3.645	0.278	1.3	3.7
AoA	1.086	1.166	15.34	<.001	1.7	22.2
Length	2.409	3.074	40.022	<.001	7.4	46.1
Concreteness	1	1	0.174	0.676	0.2	7.0
Iconicity	1.761	2.245	17.385	<.001	0.4	17.3
Grammatical category	4.393	10	17.666	<.001		6.9

Note: The fifth column shows the percentage of variance explained by each variable, calculated using a pseudo-R^2 method from Wood (see Supplementary Materials). The sixth column shows the relative importance from the random forests analysis for reference.

Winter et al. (2017). We show in the Supplementary Materials that similar relations among the psycholinguistic variables are shown when a more restrictive set of grammatical categories are analyzed.

The results of the GAM model predicting probability of borrowing are shown in Table 3. The effect of frequency was not significant in the current study, though the trend was similar to that observed in Monaghan and Roberts (2019) for a larger set of words: mid-frequency words were more likely to be borrowed than low- or high-frequency words. Similar to previous studies, AoA and length were found to be positively and monotonically related to probability of borrowing: shorter, earlier-acquired words are less likely to be borrowed (see Fig. 1).

For the measure of interest—iconicity—the relation to probability of borrowing is illustrated in Fig. 1; as predicted, words which are judged to be more iconic are less likely to be borrowed. The GAM indicates that the relation is monotonic and reducing. The significant regions of the model fit are for words in the central range of the distribution of iconicity.
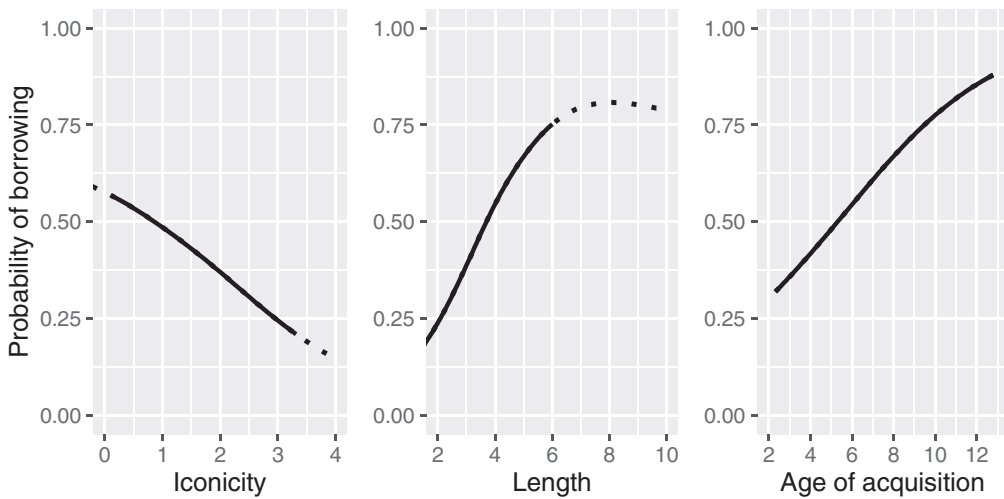
Fig. 1. Relation between iconicity, word length, and age of acquisition with the probability of borrowing. For iconicity, there is a negative relation with borrowing (more iconic words are less likely to be borrowed). Solid sections of the lines indicate regions of the model fit at which the gradient is significantly changing.

This is because there is some sparsity in the distribution of words with very high or very low iconicity.

As there are correlations between some of the psycholinguistic variables that may have altered or obscured the independent effect of iconicity, we repeated the GAM using only iconicity as a predictor in order to ascertain whether the effect of iconicity was due to the presence of the other predictors in the model. The results were very similar, EDF = 1.194, Ref.df = 1.367, $\chi^2 = 22.09$, $p < .001$. We also repeated the analyses using objective age of acquisition rather than subjective age of acquisition measures, from Brysbaert and Biemiller (2017). Again, the results were very similar (see Supplementary Materials).

For the three contexts in which a word can be borrowed—replacement of a pre-existing form, coexistence with an existing form, or as an insertion—we determined which of these borrowing effects were predicted by iconicity, and the other psycholinguistic properties. We repeated the GAM analyses examining each borrowing effect separately (see Supplementary Materials). Overall, the effects of AoA and length were similar for each type of borrowing effect: Late acquired and longer words are more likely to be borrowed into the language as replacements, coexisting forms, or insertions compared to words that are not borrowed. Iconic words are less likely to be borrowed into the language as coexisting forms or insertions, but the effect for borrowing as replacement of previously existing forms was not found to be significant.

Fig. 2 shows the decision tree. It suggests that the most decisive predictor is phonological length, with only around 20% of words with three or fewer phonemes being borrowed. For words with more than three phonemes, age of acquisition is decisive, with words learned later in life being generally more likely to be borrowed (especially for words longer than five
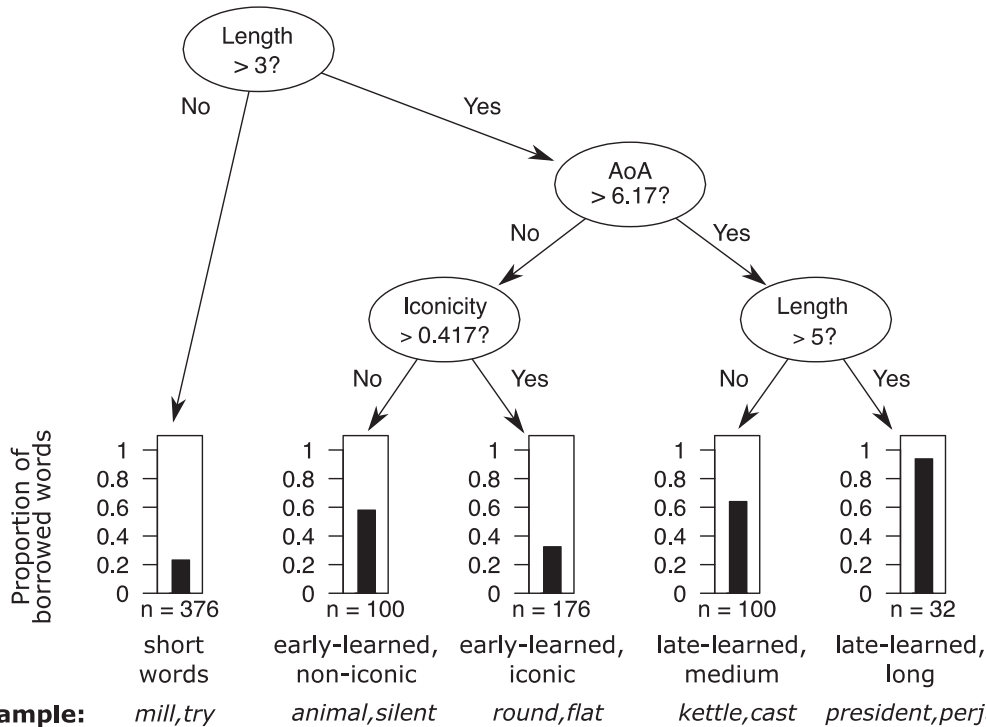
Fig. 2. A decision tree splitting the borrowing data into partitions. The first divide is by length, followed by age of acquisition (AoA). The numbers on the branches show the conditions for the split (e.g., the first split divides words with three or fewer phonemes from those with three or more phonemes). The *n* values indicate the number of observations in each partition. The bars at the bottom show the proportion of borrowed words in each partition, with examples of borrowed words below each.

phonemes). For early-learned words, iconicity is decisive, with more iconic words being less likely to be borrowed.

The importance measures from the random forests are shown in the final column of Table 2. The units are not meaningful, and only the relative sizes are informative. Length is the most decisive predictor, followed by age of acquisition, and then iconicity. The importance measures agree well with the pseudo-R^2 measures in the GAM. The random forests results thus suggest that previous analyses of lexical change may be overestimating the influence of grammatical category, and potentially underestimating the role of iconicity. Note that the results of the decision tree algorithm did not include grammatical category, frequency, and concreteness, and these measures have low importance values, suggesting that they are not effective independent predictors of borrowing. In summary, the effect of word length and age of acquisition are most important in predicting borrowing, as reflected by the pseudo-R^2 GAM values and relative importance values for the random forests analysis, with the effect of iconicity applying to words falling in the middle ground of these variables, as shown in the decision tree in Figure 2.

4. Discussion

Incidences of iconicity may be relatively sparse in natural language (Monaghan et al., 2014), yet the words that are iconic seem to have a privileged status in the vocabulary. Previous studies have shown that the extent to which the sound of a word is judged to reflect its meaning—the iconicity of a word—is related to language acquisition: Earlier acquired words are more likely to be iconic than those acquired later (Perry et al., 2015, 2018).

The current analyses demonstrate, in addition, that iconic words are also less likely to be borrowed into the vocabulary of the English language: Iconicity related negatively to probability of borrowing, particularly for the introduction of coexisting forms and insertions of words representing novel meanings into the language. The recorded borrowings of words provide observable insight into which words are stable and which are more prone to change in the vocabulary. Words that are not borrowed (i.e., according to the WOLD database, this indicates that they have preserved their form at least since proto-German) are those that are less likely to have been altered in the vocabulary.

Stability of forms of words has previously been linked to the frequency, age of acquisition, length, and number of senses of words (Monaghan, 2014; Monaghan & Roberts, 2019; Pagel et al., 2007; Vejdemo & Hörberg, 2016). Each of these properties of individual words relates not only to diachronic change but also to the representational fidelity of the word (Monaghan & Roberts, 2019). Those words which are more easily accessed, produced, and identified are those that are least prone to change in the vocabulary. By analogy to the other psycholinguistic properties, we can now add iconicity to this list of properties that highlight representational strength of the word.

However, the decision trees and random forests analyses show that certain of these psycholinguistic predictors may have a prominent role in preservation and change of lexical items. Frequency (e.g., Bybee, 2007; Pagel et al., 2007) and grammatical category (e.g., Myers-Scotton, 1993; Pagel et al., 2007) may be less important in predicting which words change and which do not than length, age of acquisition, as well as iconicity, when these variables are considered as independent contributors to determine lexical change.

Pagel, Atkinson, Calude, and Meade (2013) suggest that ultra-preserved words—those that are not prone to change—can provide insight into language ancestry, highlighting the word forms that existed in early communication. The results of the current study provide some indication that iconicity is one feature of these preserved word forms. If a word is more iconic, then it is less prone to change. It follows that iconic words are more likely to be those that existed in the vocabulary in our communicative history. It remains speculation to infer that iconicity is a property of the origins of human communication that is vestigially present in contemporary vocabulary. Nonetheless, the results indicate that regardless of how an iconic form is introduced into the language, it is more likely to remain in the language than an arbitrary form.

The current results provide another example of data where there is an intersection of principles of language acquisition, language processing, and language change. The same properties of words that highlight stability of the word form in the history of the language are those that relate to more efficient processing and ease of acquisition, indicating the multiple points

of convergence between mechanisms of language evolution and cognitive processing (Christiansen & Chater, 2008).

ACKNOWLEDGMENTS

References

Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Systems Studies*, *38*, 298–348.

Baayen, R. H., Pipenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Brand, J., Monaghan, P., & Walker, P. (2018). The changing role of sound symbolism for small versus large vocabularies. *Cognitive Science*, *42*, 578–590.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods*, *49*(4), 1520–1523.

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

Bürki, A., Alario, F. X., & Frauenfelder, U. H. (2011). Lexical representation of phonological variants: Evidence from pseudohomophone effects in different regiolects. *Journal of Memory and Language*, *64*, 424–442

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

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489–509.

Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform similarity increases with semantic similarity: An analysis of 100 languages. *Cognitive Science*, *41*(8), 2149–2169.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, *6*(10), 654–672.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*, 603–615.

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language*, *92*(2), e117–e133.

Fay, N., Ellison, M., & Garrod, S. (2014). Iconicity: From sign to system in human communication and language. *Pragmatics and Cognition*, *22*(2), 244–263.

Gasser, M., Sethuraman, N., & Hockema, S. (2010). Iconicity in expressives: An empirical investigation. In J. Newman & S. Rice (Eds.), *Empirical and experimental methods in cognitive/functional research* (pp. 163–180). Stanford, CA: CSLI Publications.

Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*, 389–407.

Grant, A. (2009). Loanwords in British English. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages: A comparative handbook* (pp. 360–383). Berlin: De Gruyter Mouton.

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

Hothorn, T., Bühlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. J. (2006). Survival ensembles. *Biostatistics*, *7*, 355–373.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, *15*, 651–674.

Key, M. R., & Comrie, B. (2015). *The intercontinental dictionary series*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Klink, R. R. (2000). Creating brand names with meaning: The use of sound symbolism. *Marketing Letters*, *11*(1), 5–20.

Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports*, *7*(1), 1–11.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*, 211–240.

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: A review of behavioural, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology*, *6*, 1246.

Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren't languages more iconic? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170137.

Marra, G., & Wood, S. N. (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, *39*(1), 53–74.

Meir, I., Padden, C., Arono, M., & Sandler, W. (2013). Competing iconicities in the structure of languages. *Cognitive Linguistics*, *24*, 309–343.

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

Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*, 325–347.

Monaghan, P., & Fletcher, M. (2019). Do sound symbolism effects for written words relate to individual phonemes or to phoneme features? *Language and Cognition*, *11*(2), 235–255.

Monaghan, P., & Roberts, S. G. (2019). Cognitive influences in language evolution: Psycholinguistic predictors of loan word borrowing. *Cognition*, *186*, 147–158.

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B*, *369*, 20130299.

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

Nichols, J. (1971). Diminutive consonant symbolism in Western North America. *Language*, *47*, 826–848.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93–104.

Ohala, J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, H. Nichols, & J. Ohala (Eds.), *Sound symbolism* (pp. 325–347). Cambridge: Cambridge University Press.

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 of the United States of America*, *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.

Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society Open Science*, *2*(8), 150152.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 1–15.

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLoS One*, *10*(9), e0137147.

Perry, L. K., Perlman, M., Winter, B., Massaro, D. W., & Lupyan, G. (2018). Iconicity in the speech of children and adults. *Developmental Science*, *21*(3), e12572.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies*, *8*, 3–34.

Roberts, S. G., Torreira, F., & Levinson, S. C. (2015). The effects of processing and sequence organization on the timing of turn taking: A corpus study. *Frontiers in Psychology*, *6*, 509.

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

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, *9*, 307.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures. *BMC Bioinformatics*, *8*, 1–10.

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*, 323–348.

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.

Tagliamonte, S. A., & Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, *24*, 135–178.

Ultan, R. (1978). Size-sound symbolism re-examined. *Psychological Bulletin*, *60*(2), 200–209.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *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.

Verhoef, T., Kirby, S., & de Boer, B. (2016). Iconicity and the emergence of combinatorial structure in language. *Cognitive Science*, *40*(8), 1969–1994.

Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*, *99*, 122–150.

Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic? *Interaction Studies*, *18*(3), 443–464.

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.

Wood, S. N. (2013). On *p*-values for smooth components of an extended generalized additive model. *Biometrika*, *100*, 221–228

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information