Decoding Surprisal and Iconicity in American English

Alexander J. Kilpatrick[1], *Rikke L. Bundgaard-Nielsen*[2]

[1]Nagoya University of Commerce and Business, [2]University of Melbourne
`alexander_kilpatrick@nucba.ac.jp, rikkieb@unimelb.edu.au`

Abstract

This meta-study investigates how phonemic bigram surprisal and iconicity affect word processing in American English. It shows that high surprisal words are harder to process than words with lower levels of surprisal, and that iconic words are easier to process than arbitrary words. The results also show that both high surprisal and iconic words are associated with improved memory recall. Additionally, the study shows that longer words generally convey less information than shorter words, especially at word boundaries, and that iconic words are more likely to maintain their high surprisal irrespective of length. These findings suggest that language evolves to the cognitive processing limitations of speakers with bigram surprisal and iconicity playing important roles in this process.

Index Terms: speech processing; surprisal; iconicity; memory recall; age of acquisition

1. Introduction

Most psycholinguists assume that language processing efficiency reflects a balance between and individual's cognitive capacity and the complexity and quantity of the information that is received [e.g., 1]. This assumption has been extensively explored and theorized in research focusing on the word level [2, 3, 4, 5] and highlights the differential effects of *predictability* on language processing. *Cognitive load theory* [6], for instance, posits that cognitive capacity is finite and that unpredictable words are processed less efficiently than predictable ones [7, 8] because of requiring increased cognitive effort. Other research indicates that decreased predictability can facilitate lexical recall. For instance, the *lossy-context surprisal model* [9] suggests that while unpredictable words pose processing challenges due to misalignment with prior expectations, the additional cognitive effort enhances the memorability of these words. This suggests that recall and comprehension are affected differentially by predictability in the input.

Here, we test the effect of two variables that influence processing: iconicity—or form-to-meaning mapping—and surprisal—a measure of the quantity of information presented. Iconicity refers to the resemblance between a word's form and its meaning and has been demonstrated to play a crucial role in language development, cognition, and processing efficiency [10, 11, 12]. For instance, longitudinal studies indicate that both children and parents tend to use more iconic words during early childhood, with a transition towards using more arbitrary language occurring as cognitive development progresses [10]. Iconic words are also processed faster and more effortlessly than arbitrary words and recalled more reliably [13]. Moreover, iconicity facilitates word learning by establishing a connection between form and meaning based on resemblance, simplifying the learning process [14]. This facilitatory role of iconicity underscores its significance in language processing and acquisition.

Research has also explored the effect of *phonological iconicity* across various sensory domains in human language [15] including size [16], shape [17], and colour [18]. Iconic words often exhibit marked phonological traits such as the use of rare or foreign speech sounds (e.g., blech [blɛx]), phonotactic violations (e.g., vroom [vɹuːm]), expressive gemination (e.g., KAP-POW! [kəˈpːaʊ]), vowel lengthening (e.g., WHAAT? [wæːt]), and expressive metathesis (e.g., aks [æks] from ask) [19]. Here, the meta-data analysed is cross-referenced with a dictionary, so these examples do not feature in the dataset; however, we explore the idea that increased average surprisal is another way that iconic words exhibit markedness.

Despite considerable variation in the average *speaking rate* (phonemes over time) across the languages of the world, speakers of different languages transmit very similar volumes of *information* per unit of time. This optimization for efficient communication supports the cognitive economy principle discussed in the current study, where we posit that languages evolve to balance communicative efficiency with processing limitations. For example, a comparison between English and Japanese indicates that English speakers typically produce 6.19 syllables/second while Japanese speakers manage 7.84 syllables/second [20]. Conversely, English has a larger phonemic inventory than Japanese and allows greater phonotactic variation, which has the consequence that the average possibility of any two phonemes co-occurring being much lower in English than in Japanese, and that English expresses more information *per phoneme* than Japanese does. The slower rate of speech in English, however, results in *a similar overall rate of information transmission* between the two languages. The similar rate of information transmission between Japanese and English is not unique to this pairing. Indeed, [21] examined the relationship between speech rate and information expression in 17 languages and found that they trend towards encoding similar information rates (39 bits/s), and [21] proposes that this is evidence that information transmission rate is modulated by universal processing limitations. Similarly, [22] showed that less-probable words tend to have segments that provide more information early, facilitating quicker and more accurate identification in a cross-linguistic study.

The present study moves away from a focus on words in some prior research and examines how predictability at the *phoneme level* affects language comprehension in American English by examining their transitional probability. To do so we calculate average bigram surprisal (hereafter: average surprisal) which is Shannon's surprisal [23] calculated on the predictability of each phoneme given its prior. This returns values in bits of information where high information reflects low predictability. Average surprisal is calculated as the sum of information divided by the number of bigrams. This calculation is included in a series of models designed to measure the

◀◀ ◀ Table of Contents for this manuscript ▶ ▶▶

influence of average surprisal on the results of pre-existing psycholinguistic tests. The results are consistent with psycholinguistic theories that suggest a trade-off between the cognitive load imposed by unexpected input and the facilitative effect of iconicity on language processing [1, 24]. In addition, we observe that while increased word length correlates with decreased surprisal at word boundaries, iconic words defy this trend and maintain high surprisal irrespective of length.

2. Method

2.1. The Master Dataset

The Master Dataset in the present study is comprised of the SUBLEX-US corpus [25] which was cross-referenced with the Carnegie Mellon University Pronouncing Dictionary [26] to convert English orthography to phonemic transcriptions based upon Standard American English pronunciation. Word length is the number of phonemes in each word. After surprisal was calculated, the master dataset was cross-referenced with additional datasets to obtain parts of speech [27], morpheme counts [28], and iconicity ratings [29] from a study in which 1400 American English speakers rated how similar each word "sounds like" its meaning on a 7-point Likert scale. In the master dataset, any word that did not find a match was discarded, resulting in a dataset of 39,136,598 instances of 13,336 unique words.

2.2. Psycholinguistic Datasets

We include datasets from five different published psycholinguistic studies: The first dataset consists of an auditory lexical decision test (MALD: [30]) completed by 231 American English-speaking participants. This study provides a dataset of reaction times and accuracy of recognizing real words, with 10,340 samples matching the master dataset, discussed above. The second dataset comes from a speeded reading experiment [31], with 816 native English-speaking participants recruited from six universities in the United States. We also include two datasets examining the age-of-acquisition of words: the first [32] involved 1960 American English participants indicating the age at which they believed they would have understood a given word, yielding 12,465 words that match the Master Dataset, while the second dataset [33] comprised responses from 829 participants at the University of Glasgow, rating words based on when they learned them, with 4101 matching words in the master dataset. Finally, we include data from a word recognition accuracy study [34] completed by 120 undergraduate students trained on a list of words in one experimental session and a subsequently tested on the accuracy of their recall in a second session within the same week.

2.3. Predictions

Hypothesis 1: We expect that word length is associated with decreased average surprisal. **Hypothesis 2:** We also expect to find that iconic words carry more information and that the effect of word length on average surprisal to be dampened in iconic words due to the inherent effortlessness of their processing. **Hypothesis 3:** We also expect to find some influence of word position on surprisal whereby the correlation between length and surprisal is dependent on the position of the bigram. **Hypothesis 4:** In reference to the psycholinguistic battery, we expect that iconic words will be processed more accurately and faster and learned at a younger age while high average surprisal words will be processed less accurately, more slowly, and learned at an older age. **Hypothesis 5:** Finally, we predict that both increased iconicity and average surprisal will be associated with increased memory recall.

3. Results

All models were constructed in R [35]: https://osf.io/mduc9/?view_only=e20f9f4d92d54044b0ab941122b3d24a. A linear regression model was constructed to examine the relationship between length and average surprisal (**H1**). The regression equation was significant, $F(1, 13704) = 501.5$, $p < .001$, with an R-squared value of .035, indicating that approximately 3.5% of the variance in average surprisal can be explained by length. The coefficient for average surprisal was significant, $b = -0.42$, $t(13704) = -22.39$, $p < .001$, indicating that average surprisal decreases as a function of length. A multiple regression analysis was conducted to predict average surprisal based on length, iconicity, number of morphemes, and parts of speech. The results are presented in Table 1.

Table 1. *Multiple linear regression model results. Asterisks denote statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).*

Variable	Result
Intercept	33.456***
F statistic	66
Degrees Freedom	15:13288
Adjusted R2	0.069
Phonemic Length	-0.085
Iconicity	9.600***
Phonemic Length:Iconicity	-6.567***
Morphemic Length	9.527***
PoS_Adverb	-3.625***
PoS_Article	-3.943***
PoS_Conjunction	-1.679
PoS_Determiner	-3.035**
PoS_Interjection	5.136***
PoS_Name	-0.042
PoS_Noun	5.818***
PoS_Number	-2.858**
PoS_Prepositon	-3.675***
PoS_Pronoun	-5.542***
PoS_Verb	-3.203**

The regression equation was significant, $F(15, 13288) = 66.01$, $p < .001$, with an R-squared value of .069, indicating that approximately 6.9% of the variance in average surprisal can be explained by the predictors. Among the predictors, iconicity ($b = 0.26$, $p < .001$), number of morphemes ($b = 0.15$, $p < .001$), and the interaction between length and iconicity were significant predictors of average surprisal (**H2**). However, length alone was not a significant predictor of average surprisal ($b = -0.00$, $p = .932$). This suggests that iconic words carry more information than arbitrary words, additional morphemes are associated with increased average surprisal when phoneme count is controlled, and the effect of length in the previous model is explained entirely by the interaction between length and iconicity whereby average surprisal does not appear to be affected by increased length in iconic words.

To test **H3**, we generated heatmaps using bigram surprisal, phonemic length, and word position to visualize how information is expressed across words. Figure 1 reveals that increased length appears to only influence surprisal at word boundaries. We considered that this may be the result of affixes which are highly predictable sequences of sounds that occur at word boundaries in English. To test this possibility, we produced a second heatmap (Figure 2) using only monomorphemic words to control for affixation. The word-boundary pattern is consistent across heatmaps. We explored this further by running a series of simple linear regression analyses to assess the relationship between word length and bigram surprisal at word start, middle, and end positions. Longer words have lower surprisal at their onset ($\beta = -0.139$, $p < .001$, $R^2 = 0.033$), and similarly at their end ($\beta = -0.171$, $p < .001$, $R^2 = 0.043$), highlighting the significance of word boundaries in information expression. The Word Middle model showed a much smaller effect size ($\beta = -0.017$, $p = .027$, $R^2 < 0.001$), suggesting a weaker influence of word length on surprisal in the middle of words. These findings are consistent in monomorphemic words only.

To test **H4** and **H5**, a series of multiple linear regression models were constructed using the accuracy and reaction times of the auditory lexical decision task, the accuracy and reaction times of the speeded reading experiment, the age of acquisition (AoA) scores of the two AoA experiments, and the accuracy scores of the memory recall experiment (Table 2). The analyses are consistent with H4, demonstrating both that iconic words are processed more accurately and rapidly, and acquired at a younger age compared to non-iconic words, and that high surprisal words, are processed with lower accuracy, slower response times, and tend to be learned at an older age. Finally, **H5** is also supported, as both increased iconicity and average surprisal are associated with enhanced recall, highlighting the mnemonic advantage conferred by these linguistic features.

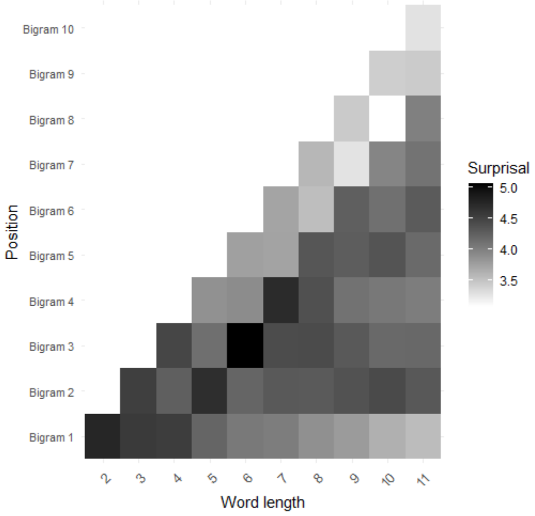

Figure 1: *Heatmap of surprisal across words according to length.*

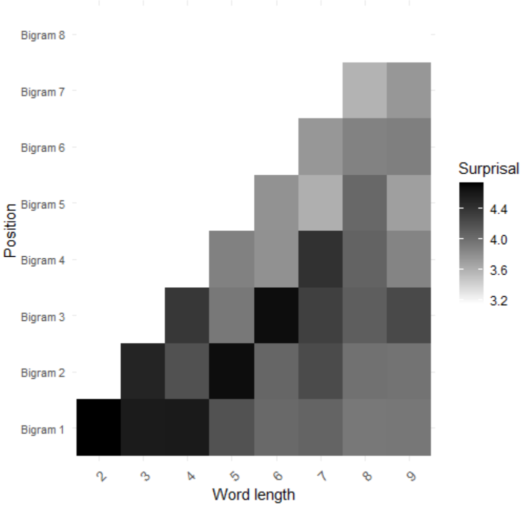

Figure 2: *Heatmap of the distribution of information according to length in monomorphemic words.*

Table 2. *Multiple linear regression model results constructed to test how various variables influence Average Surprisal. Asterisks denote statistical significance (* p < 0.05, ** p < 0.01, *** p < 0.001).*

Test	**Lexical Decision**		**Reading Task**		**Age of Acquisition**		**Memory**
Variable	**Accuracy**	**RT**	**Accuracy**	**RT**	**AOA:[32]**	**AOA:[33]**	**Accuracy**
Intercept	88.98***	67.69***	244.53***	122.09***	44.76***	20.84***	30.25***
F statistic	11.05	28.4	44.79	417.6	261.8	70.83	58.71
DF	15:10324	15:10324	15:13093	15:13093	15:12499	14:4086	11:4561
Adjusted R2	0.016	0.04	0.049	0.3228	0.239	0.195	0.122
Average Surprisal	**-3.02****	**3.05****	**-6.10*****	**9.37*****	**8.79*****	**6.62*****	**30.25*****
Iconicity	**3.1****	**-2.17***	**16.07*****	**-11.77*****	**-20.26*****	**-6.89*****	**8.09*****
Phonemic Length	6.46***	13.25***	-13.40***	56.65***	33.40***	21.38***	3.32***
Morphemic Length	1.97*	2.169*	5.81***	-4.44***	1.42	0.17	-2.36*
PoS_Adverb	1.33	-0.6	3.30***	-5.64***	-11.47***	-3.60***	-3.87***
PoS_Article	-4.64***	1.33	0.92	0.10	-2.57*	-1.54	
PoS_Conjunction	-1.39	1.02	0.27	-1.06	-5.09***		
PoS_Determiner	-0.36	0.42	1.98*	-1.56	-5.95***	-2.35*	
PoS_Interjection	-0.35	2.80**	-1.17	0.33	-3.81***	-2.54*	0.28
PoS_Name	0.68	2.49*	1.08	-0.32	-0.80	-1.01	2.57*
PoS_Noun	2.63**	0.73	5.42***	-3.53***	-8.40***	-2.24*	5.47***
PoS_Number	0.48	0.82	2.03*	-1.56	-8.48***	-2.44*	0.49
PoS_Prepositon	-1.6	0.77	1.92	-1.56	-6.53***	-2.71**	-1.06
PoS_Pronoun	1.22	0.016	2.56*	-3.93***	-10.06***	-1.97*	
PoS_Verb	2.13*	2.94**	4.84***	-0.60	1.16	0.37	-10.39***

4. Discussion

In the present investigation, we demonstrated that word length significantly influences surprisal within words, with longer words typically conveying less information per phoneme (**H1**). Interestingly, the effect of length on surprisal is largely explained by its interaction with iconicity (**H2**). This supports previous studies that suggest a unique processing advantage for iconic words [24]. Additionally, bigram position in words modulates information distribution (**H3**), with length's impact on surprisal particularly pronounced at word boundaries, possibly to facilitate early identification [22]; however, this does not explain why the effect was observed at the end of words and morphemes. Our analysis of the data from the psycholinguistic tests presented in [30, 31, 32, 33, 34] also demonstrate that highly iconic words are associated with increased processing efficiency and decreased age of acquisition, while high surprisal words exhibit the opposite pattern (**H4**). Both increased surprisal and iconicity correlate with improved recognition memory (**H5**), highlighting a complex relationship between linguistic properties and cognitive performance. These findings provide evidence for a complex relationship between cognitive cost and language efficiency, and indicate that cognitive constraints might be an important factor in language evolution. They also suggest that existing models, such as *cognitive load theory* [6] and the *lossy-context surprisal model* [9] should be extended beyond the word-level to include the transitional probabilities of phonemes. Within these existing frameworks, words consisting of predictable phoneme sequences are assumed to be easier to process while those made up of unpredictable sequences are more difficult. However, the additional investment of cognitive resources in these more difficult words enhances long-term recall. Iconic words are an important piece of this puzzle because they are processed with inherent effortlessness. Therefore, they can be used as a benchmark to measure the influence of increased cognitive costs by observing how little these effects impact iconic words, compared to their non-iconic counterparts. This includes the effect of word length on average surprisal which is entirely accounted for by its interaction with iconicity according to the model presented in Table 1.

The relationship between word length, iconicity, and average surprisal sheds light on how language is adapted to speakers' cognitive demands. The fact that longer words typically convey less information per phoneme can be argued to reflect a principle of linguistic efficiency. Despite this, iconic words, which have higher surprisal, are processed with greater accuracy and speed, suggesting a cognitive advantage for iconicity that is strong enough to counter the increased processing cost associated with higher levels of information. This preference for iconicity not only enhances cognitive processing and recall, but potentially also, influences how English has evolved. *The iconic treadmill hypothesis* [36], for example, posits that iconic words tend to evolve towards arbitrariness. Perhaps that process of evolution towards arbitrariness also creates more predictable sequences of sounds to offset the increased cognitive demands of processing non-iconic words. Consider the evolution of the English word *laugh*, which goes back to Old English *hlehhan* [37] where the iconic associations are much more evident. In old English, the hl-onset was very likely a low frequency/high surprisal sequence because it is no longer phonotactically legal. As the word evolved, it lost most of its iconic associations and became a comparatively predictable CVC sequence. This example speaks to how the transition from iconicity to arbitrariness in language may involve a reduction in information, making words easier to process.

Cognitive load is closely tied to predictability in language processing and *cognitive load theory* [6] suggests that the brain expends more cognitive resources to integrate unpredictable words into a given context than it does predictable words, resulting in longer reading times and slower response times. However, despite the increased processing effort, unpredictable words may be more memorable, as suggested by the *lossy-context surprisal model* [9]. This model proposes that the processing difficulty of a word is linked to its surprisal value within a memory representation of the context. High-surprisal words, although more challenging to process initially, may lead to deeper encoding and better memory recall due to the additional cognitive effort required for integration. Indeed, the experiments in our study support this idea, showing a positive correlation between memory recall and surprisal.

The investigation into how the position of bigrams within words influences surprisal reveals that the impact of word length on information is primarily observed at word boundaries. This would suggest that word boundaries may play a particularly important role in directing the allocation of cognitive resources by listeners for upcoming signals, if a word can be (correctly) retrieved from the lexicon/memory based on its initial segments only. However, this interpretation is preliminary and warrants further exploration to understand the significance of word boundaries as cues for cognitive resource allocation. This allows us to explore the interaction between linguistic structure and cognitive mechanisms. Additionally, it would be valuable to investigate potential differences between complex multi- and monomorphemic words. [38] highlights that complex words have multiple points at which the probability of a target word shifts. These points influence response latencies due to the Surprisal carried by phonemes suggesting that the cognitive cost of updating probability distribution differs between simple and complex words.

We recognize that our study, which focuses exclusively on English words, is of course limited. The results raise the question of whether similar patterns exist in other languages, with different phonemic inventories, different phonotactics, and differently shaped lexicons, prompting the need for further exploration. Future studies could broaden their scope to encompass languages from various linguistic families and cultural contexts. Such comparative analyses would not only deepen our comprehension of universal versus language-specific phenomena but also illuminate broader principles governing human language processing and evolution.

Our findings suggest that (at least) English has evolved in response to cognitive demands by minimizing cognitive load. Longer words tend to convey less information, particularly at word boundaries, but their impact diminishes with easier processing. Despite carrying more information, iconic words are processed more efficiently, illustrating a cognitive advantage. These insights, though limited to American English, prompt further investigation across diverse languages to better understand the universal versus language-specific effects on language processing and evolution.

5. Acknowledgements

We wish to thank Aleksandra Ćwiek, Brett Baker, Sharon Peperkamp, Shigeto Kawahara, David Sidhu, Maria Flaksman, and Kimi Akita for their insightful comments and suggestions for this project. We also thank the researchers who made their data publicly available so we could conduct this meta-analysis. This project was funded by the Japan Society for the Promotion of Science (# 20K13055).

6. References

[1] Bybee, J., “From usage to grammar: The mind's response to repetition,” Language, 82(4):711-733, 2006.

[2] Ehrlich, S. F., & Rayner, K., “Contextual effects on word perception and eye movements during reading,” Journal of Verbal Learning and Verbal Behavior, 20(6):641-655, 1981.

[3] Altmann, G. T., & Kamide, Y., “Incremental interpretation at verbs: Restricting the domain of subsequent reference,” Cognition, 73(3):247-264, 1999.

[4] Trueswell, J. C., Tanenhaus, M. K., & Kello, C., “Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths,” Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(4):800-819, 1994.

[5] Elman, J. L., Hare, M., & McRae, K., “Learning and the structure of syntactic categories,” in K. Johnson & E. N. Gibson [Eds.], Perceptual and cognitive aspects of language and speech, pp. 103-128, Springer, 2005.

[6] Sweller, J., “Cognitive load during problem solving: Effects on learning,” Cognitive Science, 12(2):257-285, 1988.

[7] Federmeier, K. D., & Kutas, M., “A rose by any other name: Long-term memory structure and sentence processing,” Journal of Memory and Language, 41(4):469-495, 1999.

[8] Kutas, M., & Federmeier, K. D., “Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP),” Annual Review of Psychology, 62:621-647, 2011.

[9] Futrell, R., Gibson, E., & Levy, R. P., “Lossy‑context surprisal: An information‑theoretic model of memory effects in sentence processing,” Cognitive Science, 44(3), 2020.

[10] [Perlman, M., Fusaroli, R., Fein, D., & Naigles, L., “The use of iconic words in early child-parent interactions,” in the 39th Annual Conference of the Cognitive Science Society (CogSci 2017), pp. 913-918, Cognitive Science Society, 2017.

[11] Edmiston, P., Perlman, M., & Lupyan, G., “Repeated imitation makes human vocalizations more word-like,” Proceedings of the Royal Society B: Biological Sciences, 285(1874):20172709, 2018.

[12] Nölle, J., Fusaroli, R., & Tylén, K., “Iconicity in sign grounding: Representation or disambiguation,” in The Evolution of Language: Proc. of the 13th Int. Conf. Evolution of Language (EVOLANG XIII), pp. 318-320, 2020.

[13] Sidhu, D. M., Khachatoorian, N., & Vigliocco, G., “Effects of Iconicity in Recognition Memory,” Cognitive Science, 47(11), 2023.

[14] Nielsen, A. K., & Dingemanse, M., “Iconicity in word learning and beyond: A critical review,” Language and Speech, 64(1):52-72, 2021.

[15] Perniss, P., & Vigliocco, G., “Iconicity: A review,” in The Oxford Handbook of Synesthesia, pp. 947-973, Oxford University Press, 2014.

[16] Sapir, E., “A study in phonetic symbolism,” Journal of Experimental Psychology, 12(3):225-239, 1929.

[17] Ćwiek, A., et al., “The bouba/kiki effect is robust across cultures and writing systems,” Philosophical Transactions of the Royal Society B, 377(1841):20200390, 2022.

[18] Erben Johansson, N., The building blocks of sound symbolism, Doctoral dissertation, Lund University, 2020.

[19] Voeltz, F. K. E., & Kilian-Hatz, C. [Eds.], Ideophones: Typological studies in language, Vol. 44, John Benjamins, 2001.

[20] Pellegrino, F., Coupé, C., & Marsico, E., “Across-language perspective on speech information rate,” Language, 87(3):539-558, 2011.

[21] Coupé, C., et al., “Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche,” Science Advances, 5(9), 2019.

[22] King, A., & Wedel, A., “Greater early disambiguating information for less-probable words: The lexicon is shaped by incremental processing,” Open Mind, 4:1-12, 2020.

[23] Shannon, C. E., “A mathematical theory of communication,” Bell System Technical Journal, 27(3):379-423, 1948.

[24] Sidhu, D. M., Vigliocco, G., & Pexman, P. M., “Iconicity in language processing and learning,” Annual Review of Psychology, 71:633-663, 2020.

[25] Brysbaert, M., & New, B., “Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English,” Behavior Research Methods, 41(4):977-990, 2009.

[26] Weide, R., “The Carnegie Mellon pronouncing dictionary,” Release 0.6, 1998.

[27] Brysbaert, M., New, B., & Keuleers, E., “Adding part-of-speech information to the SUBTLEX-US word frequencies,” Behavior Research Methods, 44:991-997, 2012.

[28] Sánchez-Gutiérrez, C. H., et al., “MorphoLex: A derivational morphological database for 70,000 English words,” Behavior Research Methods, 50:1568-1580, 2018.

[29] Winter, B., et al., “Iconicity ratings for 14,000+ English words,” Behavior Research Methods, 1-16, 2023.

[30] Tucker, B. V., et al., “The massive auditory lexical decision (MALD) database,” Behavior Research Methods, 51:1187-1204, 2019.

[31] Balota, D. A., et al., “The English lexicon project,” Behavior Research Methods, 39:445-459, 2007.

[32] Kuperman, V., et al., “Age-of-acquisition ratings for 30,000 English words,” Behavior Research Methods, 44:978-990, 2012.

[33] Scott, G. G., et al., “The Glasgow Norms: Ratings of 5,500 words on nine scales,” Behavior Research Methods, 51:1258-1270, 2019.

[34] Cortese, M. J., et al., “Recognition memory for 2,578 monosyllabic words,” Memory, 18(6):595-609, 2010.

[35] R Core Team, “R: A language and environment for statistical computing,” Build 548, 2024.

[36] Flaksman, M., “Iconic treadmill hypothesis: The reasons behind continuous onomatopoeic coinage,” in A. Zirker, et al. [Eds.], Dimensions of Iconicity, pp. 15-38, John Benjamins, 2017.

[37] Wedgwood, H., & Atkinson, J.C., A dictionary of English etymology, Trübner & Company, 1872.

[38] Balling, L. W., & Baayen, R. H., “Probability and surprisal in auditory comprehension of morphologically complex words,” Cognition, 125(1):80-106, 2012.