Lonely sensational icons: semantic neighbourhood density, sensory experience and iconicity

David M. Sidhu & Penny M. Pexman

To cite this article: David M. Sidhu & Penny M. Pexman (2018) Lonely sensational icons: semantic neighbourhood density, sensory experience and iconicity, Language, Cognition and Neuroscience, 33:1, 25-31, DOI: 10.1080/23273798.2017.1358379

To link to this article: https://doi.org/10.1080/23273798.2017.1358379

View supplementary material

Published online: 31 Jul 2017.

Submit your article to this journal

Article views: 128

View Crossmark data

Citing articles: 2 View citing articles
Lonely sensational icons: semantic neighbourhood density, sensory experience and iconicity

David M. Sidhu and Penny M. Pexman
Department of Psychology, University of Calgary, Calgary, AB, Canada

ABSTRACT
Perry, Perlman, and Lupyan (2015. Iconicity in English and Spanish and its relation to lexical category and age of acquisition. PLoS One, 10, e0137147. doi:10.1371/journal.pone.0137147) found that a sample of English words was rated as being slightly iconic, on average, with words varying in their iconicity. Thus, the relationship between word form and meaning does not seem to be categorically arbitrary. We investigated factors that might explain variation in iconicity: specifically, that concepts with sparser semantic neighbourhoods have more iconic word forms, and that concepts with more sensory information are more likely to have iconic word forms (as in Winter, Perlman, Perry, & Lupyan, in press. Which words are the most iconic? Iconicity in English sensory words. Interaction Studies. Retrieved from http://sapir.psych.wisc.edu/papers/winter_perlman_perry_lupyan_interaction-studies.pdf), even after accounting for age of acquisition (AoA; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012. Age-of-acquisition ratings for 30,000 English words. Behavior Research Methods, 44, 978–990. doi:10.3758/s13428-012-0210-4). We found support for both predictions: words with sparser semantic neighbourhoods (ARC; Shaoul & Westbury, 2010. Exploring lexical co-occurrence space using HiDEx. Behavior Research Methods, 42, 393–413. doi:10.3758/Brm.42.2.393), and greater associated sensory experience (SER; Juhasz & Yap, 2013. Sensory experience ratings for over 5,000 mono-and disyllabic words. Behavior Research Methods, 45, 160–168. doi:10.3758/s13428-012-0242-9), were more iconic, even after accounting for AoA. ARC was also found to moderate SER. These results further our appreciation of iconicity as a general property of the lexicon.

ARTICLE HISTORY
Received 9 March 2017
Accepted 3 July 2017

KEYWORDS
Sound symbolism; iconicity; arbitrariness; semantic neighbourhoods; sensory features

Introduction
The nature of the relationship between the form of a word and its meaning has been debated since at least the time that Plato’s Cratylus takes place (fifth century BC; Sedley, 2006). The debate centres on whether the form of a word (i.e. its articulation and phonology) is related to its meaning. One position is that form and meaning have an arbitrary relationship, without any special connection (e.g. Hockett, 1963), such that aspects of a word’s form cannot be used as clues to its meaning (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015). For instance, consider the seemingly arbitrary word apple and the challenge of deriving its meaning based solely on its form.

However, it is also possible for aspects of a word’s form to be non-arbitrarily related to its meaning. One example of this is iconicity, in which aspects of a word’s form map onto aspects of its meaning (Emmorey, 2014; Taub, 2001). For instance, in the word ding aspects of form (i.e. abrupt onset, fading onset) map onto aspects of meaning via resemblance (Taub, 2001). While spoken languages (this article’s focus) allow for direct iconic mapping of auditory features, it is also possible for other kinds of sensory features to map onto form indirectly via sound symbolic associations. For instance, there is a well-documented association between the phonemes /b/, /m/, /l/, /n/, /u/ and /o/, and roundness (i.e. the Maluma/Takete effect; Köhler, 1929). Thus a word like balloon, which denotes a round object, could be considered indirectly iconic.

Blasi, Wichmann, Hammarström, Stadler, and Christiansen (2015) speculated that iconicity might explain their finding of consistency in the phonemes occurring in several basic vocabulary words, across nearly two-thirds of the world’s languages.

Instead of viewing the distinction between arbitrary and non-arbitrary words categorically, it has been suggested that words throughout the lexicon can have both arbitrary and non-arbitrary elements (e.g. Dingemanse et al., 2015). The word hiccups, for instance, sounds like the meaning it conveys (i.e. an iconic property), but its meaning cannot be derived purely from its form (i.e. an arbitrary property). By this view, the distinction between arbitrariness and non-arbitrariness is...
continuous, with non-arbitrariness (e.g. iconicity) appearing in varying amounts, throughout the lexicon.

Perry et al. (2015) quantified this, by having participants rate the iconicity of 592 English words on a scale from −5 to 5: −5 indicated the word sounded opposite to its meaning, 0 indicated the word sounded nothing like its meaning, and 5 indicated the word sounded just like its meaning. Words were presented visually (Experiment 1) or auditorily (Experiment 2). In both cases, words’ average rated iconicity was significantly higher than 0 (M_{Experiment 1} = 0.75, SD_{Experiment 1} = 0.99; M_{Experiment 2} = 0.78, SD_{Experiment 2} = 0.98), suggesting modest iconicity in this sample of English words (this was replicated in a larger item set by Winter et al., in press). The authors found, further, that onomatopoetic words were the most iconic, followed by interjections, adjectives and verbs. On average, nouns and function words were not rated as iconic. Lastly, Perry et al. (2015) also discovered that more iconic words tend to be acquired earlier.

These results suggest that iconicity may be present throughout the lexicon, but that it is not predominant. This is somewhat puzzling, given that iconicity makes communication more direct and vivid (Lockwood & Dingemanse, 2015), and facilitates language learning (for a review see Imai & Kita, 2014). If language is viewed as being named mulu, molo and lomo. Because similar meanings would beg similar forms, an iconic language would be populated by sets of words with similar forms and meanings. This would lead to ambiguity, and deficiencies in processing and learnability (e.g. Gasser, 2004).

However, concepts are not equally distributed in semantic space. On the contrary, some concepts have dense semantic neighbourhoods, in which there are many concepts with similar meanings, while others have sparse neighbourhoods. For instance, many concepts are similar in meaning to apple, while fewer are similar to balloon. This leads to the prediction that concepts with sparser semantic neighbourhoods can afford to have more iconic word forms, and enjoy the benefits of iconicity, without risking confusion. Conversely, concepts with denser semantic neighbourhoods may need to have relatively more arbitrary word forms. This illustrates how iconicity and arbitrariness can play complimentary roles in language (see Dingemanse et al., 2015; Perniss & Vigliocco, 2014), with each taking on a more prominent role in different contexts.

The main goal of the present paper was to examine the possibility that words used to describe concepts with sparser semantic neighbourhoods will be relatively more iconic. In addition, we sought to further examine the relationship between sensory experience and iconicity described in Winter et al. (in press). Because earlier acquired concepts tend to be both richer in sensory experience (Juhasz, Yap, Dicke, Taylor, & Gullick, 2011) and more iconic (Perry et al., 2015), it is important to examine the alternate explanation that the relationship between SER and iconicity is attributable to age of acquisition. We also examined whether the effects of semantic neighbourhood density and SER are additive. To accomplish these goals, we used the ratings collected by Perry et al. (2015) and Winter et al. (in press) to quantify iconicity. We used Shaoul and Westbury’s (2010) average radius of co-occurrence (ARC) variable to measure semantic neighbourhood density. ARC uses lexical co-occurrence information to quantify semantic similarity between a word and its neighbours in semantic space. We also used Juhasz and Yap’s (2013) sensory experience ratings (SER), to quantify the amount of sensory information associated with a concept.

Method

Materials and procedure

In these analyses we used the combined iconicity ratings from Perry et al. (Experiment 1; 2015) and Winter et al.
Our main interest was quantifying the iconicity of each word on a spectrum from arbitrary to non-arbitrary (i.e. iconic), so we eliminated 72 words that an item analysis revealed had iconicity ratings significantly below zero ($\alpha = .1$). Words rated below zero in the iconicity ratings studies were judged to have forms that mapped onto the opposite of their meaning, and thus were not arbitrary. Our interest was in the positive ratings since these captured degrees of iconicity, with lower values indicating more arbitrariness. We used a liberal criterion here because the danger of a Type 1 error (i.e. excluding words whose ratings were not actually different than zero) was less problematic than including words that affected the validity of the scale. In addition, we removed six onomatopoeic words and six interjections, since our goal was to examine the factors determining iconicity in the general lexicon.

Our analysis consisted of a hierarchical multiple regression predicting these iconicity ratings. In Step 1, we included control variables that are often used to control for lexical factors (e.g. Yap, Pexman, Wellsby, Harrgrevens, & Huff, 2012): letter length, number of phonemes, word frequency (Brysbaert & New, 2009), and orthographic Levenshtein distance (Yarkoni, Balota, & Yap, 2008). We also included age of acquisition (AoA; Kuperman Stadthagen-Gonzalez, & Brysbaert 2012). Lastly, as iconicity has been shown to vary by word type, we included three dummy coded predictors for words’ status as adjective/adverbs, verbs, or nouns. In Step 2 we added SER (Juhasz & Yap, 2013), and in Step 3 we added ARC (Shaoul & Westbury, 2010). Larger SER values indicate greater associated sensory experience; larger ARC values indicate denser semantic neighbourhoods. Finally, in Step 4 we added an interaction between SER and ARC. Data on all dimensions were available for 1709 words; see Table 1 for correlations between dimensions.

**Results**

The results from Step 1 showed that less frequent, less orthographically distinct and earlier acquired words were rated as more iconic. Replicating previous findings, nouns were also found to be less iconic than other word types. More importantly, the results from Steps 2 and 3 showed that words associated with more sensory experience, and with sparser semantic neighbourhoods, tended to be more iconic. The $\Delta R^2$ values reveal that both of these semantic variables accounted for incremental variance: $sr_{\text{SER}}^2 = .03$ and $sr_{\text{ARC}}^2 = .04$; these values are the same if ARC is entered before SER. See Table 2 for a summary of the model. Zero-order correlations indicated that SER and ARC were each associated with iconicity, when analyzed separately for adjectives/adverbs, verbs and nouns (see Table 3 and Figure 1). Lastly, there was also a significant interaction between SER and ARC (see Figure 2). Investigating the regions of significance indicated that

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iconicity</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Length</td>
<td>-.05*</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Number of Phonemes</td>
<td>-.06**</td>
<td>.81***</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Frequency</td>
<td>-.21***</td>
<td>-.31***</td>
<td>-.32***</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>OLD</td>
<td>-.08**</td>
<td>.81***</td>
<td>.71***</td>
<td>-.26***</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>AoA</td>
<td>.02***</td>
<td>.29***</td>
<td>.32***</td>
<td>-.59***</td>
<td>.28***</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>SER</td>
<td>.18***</td>
<td>.30***</td>
<td>.25***</td>
<td>-.24***</td>
<td>.27***</td>
<td>-.16***</td>
<td>--</td>
</tr>
<tr>
<td>ARC</td>
<td>-.34***</td>
<td>-.16***</td>
<td>-.14***</td>
<td>.73***</td>
<td>-.17***</td>
<td>-.37***</td>
<td>-.21***</td>
</tr>
</tbody>
</table>

Note: OLD = Orthographic Levenshtein distance; AoA = Age of Acquisition; SER = Sensory Experience Rating; ARC = Semantic Neighbourhood Density. *p < .05; **p < .01; ***p < .001.

**Table 2. Results of hierarchical regression predicting iconicity.**

<table>
<thead>
<tr>
<th>Variable (Control variables)</th>
<th>B</th>
<th>SEB</th>
<th>$\beta$</th>
<th>$sr^2$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 (Control variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.00</td>
<td>0.14**</td>
<td>0.14***</td>
</tr>
<tr>
<td>Number of Phonemes</td>
<td>-0.06</td>
<td>0.04</td>
<td>-0.06</td>
<td>0.00</td>
<td>0.14**</td>
<td>0.14***</td>
</tr>
<tr>
<td>Frequency</td>
<td>-0.48</td>
<td>0.04</td>
<td>-0.42</td>
<td>0.09***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLD</td>
<td>-0.22</td>
<td>0.09</td>
<td>-0.10</td>
<td>0.003*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AoA</td>
<td>-0.08</td>
<td>0.01</td>
<td>-0.18</td>
<td>0.02***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjective/Adverb</td>
<td>-0.00</td>
<td>0.16</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.17**</td>
<td>0.03***</td>
</tr>
<tr>
<td>Verb</td>
<td>-0.02</td>
<td>0.15</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.17**</td>
<td>0.03***</td>
</tr>
<tr>
<td>Noun</td>
<td>-0.57</td>
<td>0.15</td>
<td>-0.27</td>
<td>0.01***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SER</td>
<td>0.22</td>
<td>0.03</td>
<td>0.22</td>
<td>0.03**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARC</td>
<td>-2.53</td>
<td>0.28</td>
<td>-0.30</td>
<td>0.04***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SER x ARC Interaction</td>
<td>-0.54</td>
<td>0.20</td>
<td>-0.06</td>
<td>0.003**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: OLD = Orthographic Levenshtein distance; AoA = Age of Acquisition; SER = Sensory Experience Rating; ARC = Semantic Neighbourhood Density. *p < .05; **p < .01; ***p < .001.

**Table 3. Correlations between iconicity and ARC and SER, for each word type.**

<table>
<thead>
<tr>
<th>Word type</th>
<th>n</th>
<th>Correlation with Iconicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>SER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjectives/Adverbs</td>
<td>219</td>
<td>.18**</td>
</tr>
<tr>
<td>Function words</td>
<td>57</td>
<td>.12</td>
</tr>
<tr>
<td>Verbs</td>
<td>362</td>
<td>.46***</td>
</tr>
<tr>
<td>Nouns</td>
<td>1071</td>
<td>.13**</td>
</tr>
</tbody>
</table>

Note: SER = Sensory Experience Rating; ARC = Semantic Neighbourhood Density. *p < .05; **p < .01; ***p < .001.
higher SER was positively related with iconicity, except for words with especially high ARC values ($z > 1.63$).

Replicating previous research, we found that nouns were significantly less likely to be iconic than other word types. This was confirmed by an independent samples t-test comparing the iconicity of nouns ($M = 0.73$, $SD = 0.97$) to that of adjectives/adverbs/verbs ($M = 1.17$, $SD = 1.10$), $t(1076.22) = 8.08$, Glass’s $\Delta$Adjectives/Adverbs/Verbs = 0.40, $p < .001$ (unequal variances). We conducted supplementary analyses to further explore this finding. Surprisingly, nouns had equivalent semantic neighbourhood densities ($M = 0.56$, $SD = 0.12$) as adjectives/adverbs/verbs ($M = 0.56$, $SD = 0.14$), $t(1017.70) = 0.02$, $p = .99$ (unequal variances). Thus, this factor cannot explain difference in iconicity between word types. Additionally, nouns were associated with more sensory experience ($M = 3.41$, $SD = 0.98$) than were adjectives/adverbs/verbs ($M = 3.03$, $SD = 0.94$), $t(1650) = 7.52$, Cohen’s $d = 0.39$, $p < .001$. However, recall that Winter et al. (in press) discovered that certain kinds of sensory experience (i.e. auditory and tactile) were more important to iconicity than others. As such, it is interesting to note that nouns were associated with less auditory sensory experience (Lynott & Connell, 2009; 2013; Winter, 2016; $M = 0.53$, $SD = 0.95$) than were adjectives/adverbs/verbs ($M = 1.08$, $SD = 1.33$), $t(420.13) = 5.27$, Glass’s $\Delta$Adjectives/Adverbs/Verbs = 0.42, $p < .001$ (unequal variances). Nouns were not, however, associated with less tactile sensory experience ($M = 0.82$, $SD = 0.92$) than adjectives/adverbs/verbs ($M = 0.85$, $SD = 1.12$), $t(454.25) = 0.28$, $p = .78$ (unequal variances). Nevertheless, this is evidence that while nouns may be associated with more sensory experience overall, adjectives/adverbs/verbs may be associated with the particular kind of experience that lends itself to an iconic mapping.

**Discussion**

In recent research, arbitrariness and iconicity have been shown to exist on a spectrum, with words varying in
their arbitrary and iconic elements (e.g. Perry et al., 2015). We examined the prediction that concepts with sparser semantic neighbourhoods could afford to have more iconic forms, and further investigated the finding that concepts evoking a greater amount of sensory information are more mappable and thus more iconic (Winter et al., in press).

We found evidence for both predictions. Even after accounting for a variety of important lexical-semantic variables, words’ sensory experience ratings and semantic neighbourhood density were related to their iconicity. This former finding replicates Winter et al. (in press) even after accounting here for effects of age of acquisition, demonstrating that the relationship between sensory experience and iconicity cannot be fully attributed to this third variable. We also found further support for the notion that certain types of sensory experience are especially important to iconic mappings. While the relatively less iconic syntactic class of nouns was associated with more sensory experience than adjectives/adverbs/verbs, this latter syntactic group was associated with a greater amount of auditory sensory experience, a type of sensory feature that lends itself to iconic mappings in spoken language (Dingemanse, 2012; Winter et al., in press).

The unique contribution of the present paper is demonstrating that semantic neighbourhood density is an important factor in the iconicity of a given concept’s word form. This factor explained more unique variance than SER and AOA, factors that have previously been shown to be important to iconicity. The effect of semantic neighbourhood density highlights the cooperative roles of iconicity and arbitrariness in language (see Dingemanse et al., 2015; Perniss & Vigliocco, 2014). While iconicity conveys benefits to language, and can play a large role in some concepts’ word forms, there are instances in which arbitrariness must play the larger role, to avoid ambiguity and aid in discriminability (e.g. for concepts with dense semantic neighbourhoods).

The relationship between semantic neighbourhood density and iconicity has relevance for our understanding of the acquisition and evolution of iconic words. Some have suggested that early-acquired language can afford to be more iconic because early language maps

Figure 2. Scatterplots show the relationships between SER/ARC and iconicity, separately for adjectives/adverbs, verbs and nouns. Gray lines represent LOWESS functions (created using lowess function in R with the default smoother value of 66.67%), black lines represent zero-order correlations.
sparser semantic space (Gasser, 2004). Others have theorised that the original forms of many words may have been iconic, but that forms became more arbitrary as more words were added and semantic neighbourhoods became denser (see Imai & Kita, 2014). Our results are consistent with both of these claims. We also found a significant interaction between SER and ARC, the nature of which is consistent with these theories. In general, concepts are more able to have iconic word forms if they have a greater amount of sensory information. However, once a given region of semantic space becomes cluttered, words for these concepts tend to become arbitrary, regardless of their associated sensory experience. Instead of being additive in every case, high semantic neighbourhood density seems to nullify the effects of sensory experience.

The present study provides new insight about factors that modulate the iconicity of word forms and thus advances the ancient debate about the relationship between a word’s form and its meaning. We found that semantic neighbourhood density is important to iconicity, and that it moderates the extent to which sensory experience predicts the iconicity of concepts’ word forms. Of course, each of these variables explained a small amount of overall variance in iconicity, suggesting that there are likely other factors that are as yet unidentified. Nevertheless, these results reinforce the notion that iconicity is not simply a linguistic oddity but rather a more general property of the lexicon, one that cooperates with arbitrariness in shaping language.

**Notes**

1. We do not wish to make the mistake of evaluating the iconicity of spoken language as a whole based on spoken Indo European languages. Indeed many non-Indo European spoken languages contain many iconic words. However, even in these languages, we might ask why iconicity isn’t more prevalent.
2. Winter et al. (in press) also found that more frequent, less imageable words were more iconic. Imageability was not considered here, as it was only available for 63.32% of the words in our final sample.

**Acknowledgements**

The authors would like to thank Bodo Winter and Lynn Perry for providing trial level data of their iconicity ratings.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through: a Postgraduate Scholarship to DMS [grant number CGS D3 476111 2015] and a Discovery Grant to PMP [grant number RGPIN 217309-2013]; and by Alberta Innovates - Health Solutions (AIHS) through a Graduate Scholarship to DMS [grant number 201500125-1 CA# 3874].

**References**


