Effects of Iconicity in Lexical Decision

David M. Sidhu[a]

Gabriella Vigliocco[b]

Penny M. Pexman[a]

[a]University of Calgary

[b]University College London

Abstract

In contrast to arbitrariness, a recent perspective is that words contain both arbitrary and iconic elements. We investigated iconicity in word recognition, and the possibility that iconic words have special links between phonological and semantic features that may facilitate their processing. In Experiment 1 participants completed a lexical decision task (is this letter string a word?) including words varying in their iconicity. Notably, we manipulated stimulus presentation conditions such that the items were visually degraded for half of the participants; this manipulation has been shown to increase reliance on phonology. Responses to words higher in iconicity were faster and more accurate, but this did not interact with condition. In Experiment 2 we explicitly directed participants' attention to phonology by using a phonological lexical decision task (does this letter string *sound* like a word?). Responses to words that were higher in iconicity were once again faster. These results demonstrate facilitatory effects of iconicity in lexical processing, thus showing that the benefits of iconic mappings extend beyond those reported for language learning and those argued for language evolution.

1. Introduction

Traditional views of language have held that the relationship between the form of a word and its meaning is arbitrary (e.g., Hockett, 1963). That is, there is no special link between a word's form (i.e., its orthography, phonology or articulation) and its meaning. However, arbitrariness may be only one possible kind of relationship between form and meaning (see Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Perniss & Vigliocco, 2014). Language can also be non-arbitrary via *iconicity*: a resemblance between form and meaning, in which aspects of a word's form map onto aspects of its meaning (e.g., Perniss, Thompson & Vigliocco, 2010). In this paper we focus on phonological iconicity, wherein a word's phonological form maps onto meaning.

Iconicity can occur in several ways. In onomatopoeia, the phonology of a word imitates the sound to which it refers. English examples include *quack*, *boom* and *sizzle*. It is also possible for a word's form to map onto its meaning without direct imitation, but instead via association (i.e., *non-onomatopoeic iconicity*). This is enabled by the phenomenon of *sound symbolism*, in which particular language sounds evoke associations with non-auditory properties (see Lockwood & Dingemanse, 2015; Sidhu & Pexman, 2018). For instance, individuals associate the vowels /ɪ/ and /ɑ/ with smallness and largeness, respectively (Newman, 1933; Sapir, 1929). Thus, a word like *shrimp* could be considered iconic because its form evokes associations (i.e., smallness) that map onto aspects of its meaning.

It has been claimed that arbitrariness and iconicity both provide benefits to language. Arbitrariness allows any form to refer to any meaning without requiring resemblance, which may not always be possible (e.g., for entirely abstract concepts;

Dingemanse et al., 2015). Iconicity facilitates language learning, in part by helping infants and toddlers associate speech sounds with their referents (see Imai & Kita, 2014, Laing, Viham, & Portnoy, 2017; Perniss, Lu, Morgan, & Vigliocco, 2018; Perry, Perlman, Winter, Massaro, & Lupyan, 2017). In addition, it has been argued that the imitative, performative nature of iconic words makes language more direct and vivid (Dingemanse et al., 2015). In the present experiments we examined another potential advantage of iconicity: that iconic links between sound and meaning make these words easier to process.

Importantly, words are not entirely arbitrary or iconic. Rather, they fall on a spectrum from extremely arbitrary to extremely iconic–containing aspects of arbitrariness and iconicity (see Dingemanse et al., 2015; Perniss & Vigliocco, 2014). For instance, while the word *quack* sounds like a duck quacking, there are many other ways to imitate this sound (e.g., *mac* in Romanian, *vak* in Turkish, or *prääks* in Estonian), and thus the choice of *quack* is to some extent an arbitrary one. Indeed, other sets of phonemes might better imitate the sound of a duck. Also, while /ɪ/ in the word *shrimp* maps onto an aspect of its meaning, its other phonemes are seemingly arbitrary.[1] A word that contained a greater proportion of small-associated phonemes might seem more iconic. Perry, Perlman, and Lupyan (2015; supplemented by Winter, Perlman, Perry, & Lupyan, 2017) quantified the subjective iconicity of a large set of English words. They had participants rate a variety of English words in terms of their iconicity on a continuous scale and discovered that words existed along the entire spectrum.

[1] We thank the editor for pointing out that there is a difference in the nature of the arbitrariness in *quack* and *shrimp*. There are only so many phoneme combinations that could imitate the sound of a duck, and so while there is some room for variation, the phonemes in *quack* are still *motivated* and not entirely arbitrary. Conversely, there is seemingly no restriction in the non-iconic phonemes in *shrimp* (i.e., /ʃ/).

Formatted: Font: Times New Roman

Deleted: Instead of conceptualizing arbitrariness and iconicity as mutually exclusive categories, a more accurate view might be that words

Formatted: Font: Times New Roman

Deleted: contain both

Formatted: Font: Times New Roman

Deleted: y

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Deleted: elements

Deleted:

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Deleted: Thus, words exist on a spectrum ranging from extremely arbitrary, to extremely iconic. This was quantified for a large sample of English words by

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: Font: (Default) Times New Roman, 10 pt

Formatted: English (CAN)

There is a great deal of work to be done to understand the effects of such variation in iconicity on language processing. Triangle models of word recognition include three components: a word's meaning, its orthography and its phonology. In these models, a word's meaning is accessed via two paths: directly from a word's orthography, and indirectly from its orthography via its phonology (Harm & Seidenberg, 2004). While other models of word recognition exist (e.g., dual route models; Coltheart et al., 2001), here we focus on triangle models because they specify a route by which meaning is retrieved via phonology. Note that links between each of the three components of triangle models are bidirectional. Additionally, the extent to which paths or components are emphasized varies depending on task context (Balota, Paul, & Spieler, 1999). Evidence for this is the fact that phonological variables (e.g., spelling-sound regularity; Hino & Lupker, 1996) play more or less of a role in word recognition when the orthographic-phonological path is prioritized to a greater or lesser extent.

Mappings between phonology and semantics are considered arbitrary in triangle models: they must be learned through experience.[2] However, iconicity presents the possibility that some connections between phonology and semantics are not solely arbitrary. Such connections may emerge or exist naturally, or be easier to learn, by virtue of the inherent resemblance between phonology and semantics. Here we evaluate the possibility that iconic words possess extra and/or more direct links between phonology and semantics which will facilitate their processing. Note that this may be a quantitative

[2] Though Harm and Seidenberg (2004) mention systematic non-arbitrary patterns in orthographic-semantic links (e.g., relations among words shared an affix) that could affect learning.

and/or a qualitative difference in the nature of links for iconic as compared to non-iconic words.

Meteyard et al. (2015) examined the possibility that special links between phonology and semantics might make iconic words more resistant to aphasia. They compared processing of onomatopoeia and arbitrary words in aphasic patients using a variety of tasks. They found a benefit for onomatopoeia on the tasks that prioritized the mapping of phonology onto semantics. The authors also speculated that effects of iconicity may only be observable in individuals with developing or damaged language systems.

Peeters (2016) conducted an EEG study on the processing of Dutch onomatopoeia using an auditory lexical decision task. He found that onomatopoeia elicited a smaller N400, interpreted as reflecting facilitated lexical access for words with iconic mappings between form and meaning. However, despite this effect, there were no behavioural differences in the processing of onomatopoeic and arbitrary words. Peeters speculated that onomatopoeia may be processed both as linguistic stimuli and as environmental sounds (e.g., *boom* could be interpreted both as a word and the sound to which it refers). Peeters suggested that this could interfere with lexical decision where participants must judge stimuli to be linguistic stimuli or not.

Other evidence of iconicity potentially interfering with performance on a word recognition tasks comes from an analysis by Lupyan and Winter (2018). They analyzed data from a semantic decision task (abstract/concrete) collected by Pexman, Heard, Lloyd and Yap (2017), and found that higher iconicity actually led to a lower accuracy for more abstract items. Their interpretation was that words that are more iconic activate more specific semantic representations, which nudges participants towards a "concrete" (i.e., incorrect) decision.

In the present experiments we explored two questions. First, we examined whether the recognition of English iconic words is facilitated in a population of typical adults on a visual lexical decision task. Second, as a first step towards exploring the locus of such an effect, we varied the extent to which phonology was prioritized by the task.

In Experiment 1 we presented words varying in their iconicity in a *visual* lexical decision task (LDT: is this letter string a word?). While some previous studies have used auditory lexical decision tasks (e.g., Meteyard et al. 2015), we elected for a visual lexical decision task because: 1) there is still evidence of phonological processing on such a task (e.g., Pexman, Lupker, & Jared, 2001), and 2) using a more conservative test, and one that is more similar to everyday reading processes, allowed us to ensure that any effects would be more broadly applicable. Stimuli in the LDT were presented clearly or visually degraded. Visual degradation prevents participants from making lexical decisions solely on the basis of orthographic information, and thus tends to increase the extent to which phonological information is recruited (Hino & Lupker, 1996). If the nature of the links between phonology and semantics for iconic words provide a benefit to word recognition, enhancing the extent to which phonology is recruited by the task may enhance this effect. In Experiment 2 we presented the same items in a *phonological* lexical decision task (PLDT: does this letter string *sound* like a word?). Phonology is explicitly emphasized in the PLDT, as participants are asked to make responses based on the phonology rather than the spelling of each item.

2. Experiment 1

2.1 Methods

2.1.1. Participants. Participants were 80 undergraduate students (61 female; *M* age = 21.01; *SD* = 3.38) at the University of Calgary who participated in exchange for course credit. Participants reported English fluency and normal or corrected to normal vision.

2.1.2. Materials and Procedure. Stimuli were 120 real words and 120 nonwords. The real words were chosen from the iconicity ratings collected by Perry et al. (2015) and Winter et al. (2017), in which words were rated on a scale from -5 (the word sounds like the opposite of its meaning) to 5 (the word was highly iconic), with 0 indicating that the word was arbitrary. In order to sample broadly from different types of words, we selected and matched words as per a factorial design, though in the analyses we treated iconicity as a continuous variable. To that end, we selected words corresponding to three categories: non-iconic words, onomatopoeia, and non-onomatopoeic iconic words. Non-iconic words had iconicity ratings ≥ -0.50 and ≤ 0.50 (*M* = 0.12; Perry et al., 2015; Winter et al., 2017). Onomatopoeia had iconicity ratings ≥ 2.50 (*M* = 3.42) and had phonologies that imitated their meanings. Non-onomatopoeic iconic words had iconicity ratings ≥ 2.50 (*M* = 2.99) but had phonologies that we judged not to directly imitate their meanings (e.g., *twist*, *fluff, slime*). The three types of words were matched on length, log subtitle word frequency (Brysbaert & New, 2009), number of morphemes, orthographic Levenshtein distance (OLD; Yarkoni, Balota, & Yap, 2008), phonological Levenshtein distance (PLD; Yarkoni et al., 2008), mean bigram frequency, age of acquisition (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012), concreteness (Brysbaert, Warriner, & Kuperman, 2014), and initial syllable rime phonological consistency (Yap, 2007). In addition, each type contained 26 nouns and 14 verbs. See Table 1 for properties of each type. The 120 nonwords did not contain any pseudohomophones (i.e., nonwords that share phonology with a real word; e.g., *brane*) and were matched with the 120 real words on length, number of orthographic neighbours, mean bigram frequency and number of syllables. See Appendix for stimuli.

(Table 1)

Participants completed an LDT in which their task was to categorize a presented letter string as a word or a nonword. Each trial began with a fixation cross for 400 msec, followed by a blank screen for 200 msec, after which the target letter string was presented. Participants categorized stimuli as nonwords or words via keyboard press. Their response triggered a 550 msec blank screen, after which the next trial began. If participants made an error they saw the word "Incorrect", and heard a brief sound, during this blank screen. Participants wore sound attenuating headphones during the task. Stimuli were presented in a random order, in two blocks, with a break between blocks. Participants saw an equal proportion of each stimulus type in each block.

In addition, participants were randomly assigned to one of two presentation conditions: clear or degraded (40 participants in each). In the clear condition stimuli were presented normally. In the degraded condition we followed the approach taken by Yap, Lim and Pexman (2015) to visually degrade the letter strings, rapidly alternating between the letter string and a mask of random symbols of the same length. This condition was run with a refresh rate of 144 Hz.

2.2. Results

2.2.1. Statistics. We used the packages "lme4" [version 1.1-18-1] (Bates, Maechler, Bolker, & Walker, 2015), "afex" [0.23-0] (Singmann, Bolker, & Westfall, 2015), and "RePsychLing" [0.0.4] (Baayen, Bates, Kliegl, & Vasishth, 2015) to perform our statistical analysis in R [3.5.1] (R Core Team, 2016). We took a confirmatory approach and fit models including all fixed effects of interest. We developed each model's random effects structure using the approach suggested by Bates, Kliegl, Vasishth, and Baayen (2015). In brief, we began by fitting the model with all random slope terms for each fixed effect, and removed correlations among random effects if this did not converge. We then performed a principal components analysis on the random effects and simplified the structure based on the suggested number of components (Baayen et al., 2015). We also tested the inclusion of correlations among random effects, and the effects themselves, using likelihood ratio tests. The detailed procedure for model selection, along with code used for the entire process, can be found in the online supplementary materials: https://osf.io/ue7sv/. We generated p-values using the package "lmerTest" [3.0.1] (Kuznetsova, Brockhoff, & Christensen, 2017). The "prediction" package [0.3.6] was used to generate marginal predictions. Throughout all results, we only report analyses based on real word trials.

We took the following approach to cleaning the data in these and all analyses of reaction time. First, we excluded all incorrect responses. Then trials with a reaction time less than 200 msec or greater than 3000 msec were removed. We then removed trials that were more than 2.5 standard deviations away from a participant's mean. No more than 5.13% of trials were ever removed by this process. We took the same approach to cleaning the accuracy data except that incorrect responses were not excluded.

2.2.2. Reaction time. We ran a linear mixed effects model[3] that predicted reaction time using condition (effects coded; -1 = clear presentation, +1 = degraded presentation), continuous iconicity (Perry et al., 2015; Winter et al., 2018) and their interaction. Length, frequency (log subtitle word frequency) and OLD were also included as control variables. This model revealed that there was no interaction between condition and iconicity (b = 1.82, p = .41). However, there was a main effect of iconicity such that reaction times were faster in response to words with higher iconicity (b = -13.65, p = .002). In particular, iconicity values of -1.5 and +1.5 SD resulted in predicted reaction times of 724.22 and 683.41 msecs, respectively. In addition, there was a main effect of condition, such that reaction times were faster in the clear condition (b = -45.77, p = .02). In particular, the clear and degraded conditions resulted in predicted reaction times of 659.19 and 750.72, respectively. See Table 2 for a model summary.

(Table 2)

2.2.3. Accuracy. We ran a logistic mixed effects model that predicted accuracy and included the same fixed and random effects as in the main analysis of reaction time. This model revealed that there was no interaction between condition and iconicity (b = -0.01, p = .72). However, there was a main effect of iconicity such that correct responses were 1.25 times more likely for each 1 SD increase in iconicity (b = 0.22, p = .005). There was not, however, a main effect of condition (b = 0.13, p = .12). See Table 3 for a model summary.

(Table 3)

[3] `lmer(RT~Length+Frequency+OLD+Iconicity*Condition+(1|Subject)+(1|Word))` Note that all subsequent models use this same general structure with changes to the included fixed or random effects noted in text. The specific models used in each analysis can be found in the following OSF repository: https://osf.io/ue7sv/

Deleted: that the
Deleted:
Deleted: linked to in text

2.3. Discussion

Participants were faster and more accurate when responding to words that were higher in iconicity. This suggests that iconicity does provide a benefit to word recognition. Interestingly, the effect of iconicity did not interact with condition, suggesting that any increased use of phonology due to visual degradation did not enhance the effect of iconicity. In the next experiment we explored the role of phonological encoding further, using a task that explicitly prioritizes phonology, by requiring participants to attend to the phonologies of each word (and not the spellings) in order to make a correct response.

3. Experiment 2

3.1 Methods

3.1.1. Participants. Participants were 48 undergraduate students at the University of Calgary who participated in exchange for course credit. Participants reported English fluency and normal or corrected to normal vision. We tested more participants than we intended to analyze (i.e., 40), because previous work with the PLDT has found that some participants do not comply with the instructions to emphasize phonology (e.g., Pexman, Lupker & Reggin, 2002). This noncompliance is evident in accuracy on pseudohomophone trials. Thus, we included the 40 participants (25 female; *M* age = 23.15; *SD* = 5.41) with the highest accuracy on pseudohomophone trials (i.e., correctly categorizing stimuli like *brane* as sounding like a real word; see below for details).[4]

3.1.2. Materials and Procedure. In addition to the stimuli from Experiment 1, we added 60 pseudohomophones and 60 additional nonwords. These pseudohomophones

---

[4] Note that the to-be-reported significant effects remain significant when including all 48 participants.

were created by altering the spelling of an existing word (e.g., *cough* to *koff*). As in Experiment 1, we selected and matched words as per a factorial design, though we treated iconicity as a continuous variable in the analyses. To that end, pseudohomophones were created using the phonologies of 20 non-iconic, 20 onomatopeia, and 20 non-onomatopoeic iconic words. Stimuli requiring a "word" response and stimuli requiring a "nonword" response were matched on length, orthographic neighbourhood size, mean bigram frequency and number of syllables. See Appendix for stimuli.

Participants completed a PLDT in which their task was to categorize a letter string as sounding like a word (i.e., having the phonology of a real word) or a nonword (i.e., not having the phonology of a real word). Except for the different decision criterion, trials were presented in the same manner as in the clear condition of Experiment 1.

3.2 Results

Data were analyzed using the same approach as in Experiment 1.

3.2.1. Reaction time. We ran a linear mixed effects model that predicted reaction time using iconicity as a predictor. The model also included length, frequency (log subtitle word frequency) and PLD (as this experiment prioritized phonology) as control variables. This model revealed an effect of iconicity such that reaction times were faster in response to words with higher iconicity (b = -22.48, p < .001). In particular, iconicity values of -1.5 and +1.5 SD resulted in predicted reaction times of 780.33 and 712.90 msecs, respectively. See Table 4 for a model summary.

(Table 4)

We ran a supplementary analysis in which we combined the present results with those of the clear condition in Experiment 1 and tested for an interaction between task

(LDT vs PLDT) and iconicity, to examine the impact of explicitly directing participants to focus on phonology. The model also included length, frequency and OLD as control variables; it also included a random item slope for the effect of task. The interaction between task and iconicity was non-significant (b = -5.24, p = .06).

Finally, we ran a supplementary analysis on the pseudohomophone trials. We ran a model that predicted reaction time using the iconicity of the word on which each pseudohomophone was based. The model also included length, orthographic neighbourhood size, and mean bigram frequency of the pseudohomophones, and base word frequency (log subtitle frequency), as control variables. This model revealed an effect of base word iconicity, such that reaction times were faster in response to pseudohomophones that were based on words with higher iconicity (b = -61.78, p = .04). In particular, base word iconicity values of -1.5 and +1.5 SD resulted in predicted reaction times of 1196.96 and 1011.61, respectively.

3.2.2. Accuracy. We ran a logistic mixed effects model that predicted accuracy using the same fixed and random effects as in the main analysis of reaction time. This model found no effect of iconicity on response accuracy (b = 0.18, p = .13), see Table 5.

We ran a supplementary analysis in which we combined the present results with those of the clear condition in Experiment 1 and tested for an interaction between task and word type. The model also included length, frequency and OLD as control variables. This model found no interaction between task and iconicity (b = 0.00, p = .97).

Finally, we again ran a supplementary analysis on the pseudohomophone trials. We ran a model that predicted accuracy using the same fixed and random effects as in the analysis of pseudohomophone reaction time. This model revealed an effect of base word iconicity, such that correct responses were 1.80 times more likely for each 1 SD increase in base word iconicity ($b = 0.59$, $p = .004$).

3.3. Discussion

In Experiment 2, we found that participants responded faster to words with higher iconicity. They did not, however, respond more accurately. Interestingly, an iconicity benefit also emerged in the processing of pseudohomophones. As the correct identification of these items would have relied on phonology, this speaks to the role of phonology in iconicity. We next examined responses to these words when presented in the broader context of a megastudy that included a wide variety of words.

4. English Lexicon Project Analysis

We performed an analysis examining average reaction time and accuracy in the English Lexicon Project (*ELP*; Balota et al., 2007) LDT for the 120 words used in Experiments 1 and 2.

4.1 Results

We ran a linear model that predicted LDT reaction time in the ELP using iconicity as a predictor. The model also included length, frequency (log subtitle word frequency) and OLD as control variables. Note that this analysis was done at the item level, and therefore did not include random effects. This model found no effect of iconicity on reaction time ($b = 11.00$ $p = .054$). The corresponding logistic model found no effect of iconicity on accuracy ($b = 0.01$ $p = .46$).

In a supplementary analysis, we explored the possibility that list context might have led to the difference in results between the ELP data and the non-degraded LDT from Experiment 1 (which both used the same task). While ELP participants were presented with a wide variety of words, two-thirds of the words in Experiment 1 were iconic. We examined whether this list context in the non-degraded LDT from Experiment 1 increased participants' attention to iconicity through a linear mixed effects model that included trial number and its interaction with iconicity, in addition to previously used control variables. We found a significant interaction between trial number and iconicity (b = -8.82, p = .002), such that iconicity had a greater effect on later trials (see Figure 1). In particular, iconicity values of -1.5 and +1.5 SD resulted in predicted reaction times of 672.76 and 660.88 on trial 60 (of 240), and 688.20 and 629.91 on trial 180, respectively. The corresponding logistic mixed effects model found no interaction between condition and trial number (b = 0.06, p = .34).

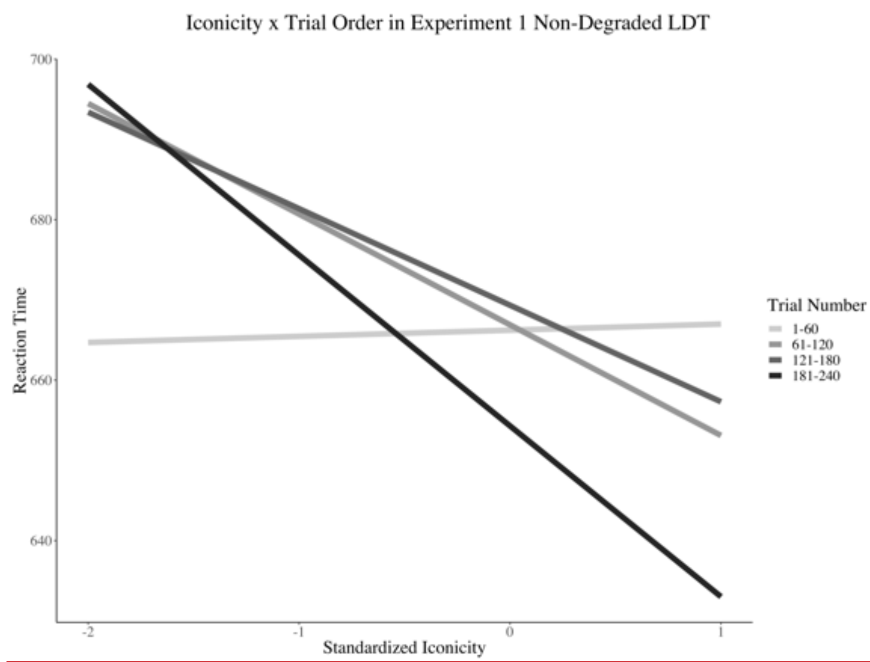

Figure 1. Plot showing the relationship between iconicity and predicted reaction time in the non-degraded LDT from Experiment 1. This was calculated separately for each quarter of the experiment (e.g., the first 60 trials that a participant saw represents the first quarter of the experiment). Each line shows the effect in a different quarter of the experiment.

Deleted: , for each quarter of trials separately.

4.2. Discussion

We found no effect of iconicity on LDT response times in the ELP. This contrasts with results from the non-degraded LDT in Experiment 1. One possibility is that list context plays a role. While two-thirds of the words in Experiment 1 were iconic, this was not true in the ELP. Speaking to this is the fact that the effect of iconicity on reaction time in the non-degraded LDT from Experiment 1 emerged over time, potentially as participants shifted their attention towards iconicity as a useful cue to a "word" response.

5. General Discussion

It had traditionally been assumed that the relationship between a word's form and its meaning is arbitrary (e.g., Hockett, 1963), but a more recent perspective is that this relationship contains aspects of arbitrariness and iconicity (Dingemanse et al., 2015; Perniss et al., 2010). Here we tested the possibility that recognition might be facilitated for relatively more iconic words. Indeed, we found that more iconic words were processed faster and more accurately in a visual LDT, and faster in a phonological LDT.

These experiments suggest that iconicity can confer an advantage in processing to typical individuals, in addition to people with aphasia (Meteyard et al., 2015). These results stand somewhat in contrast to a study on Dutch onomatopoeic words that found no effects in reaction time (Peeters, 2016). Note that stimuli in the Dutch study were presented auditorily, and that this difference in modality may interact with effects of iconicity (cf. Meteyard et al., 2015). In fact, Peeters speculated that the lack of an iconicity effect may have been due to onomatopoeia being more difficult to identify as words (as opposed to environmental sounds). Auditory presentation may have exaggerated this. It is also important to note that the study by Peeters (2016) did find a smaller N400 in response to onomatopoeic vs. non-onomatopoeic words, which was interpreted as being indicative of facilitated word retrieval, and thus consistent with the present results.

Deleted: both

Deleted: y and iconic elements

We found some evidence that list context plays a role in the effects of iconicity. Indeed, LDT reaction time and accuracy in the ELP were not affected by iconicity. Note that ELP participants would have received a much smaller proportion of iconic words than the participants we tested in Experiments 1 and 2. It may be that this greater proportion of iconic words served to shift participants' attention to iconicity. Indeed, the effect of iconicity on reaction time in the non-degraded LDT from Experiment 1 emerged over the course of that experiment. Thus, it seems that participants' use of iconicity in LDT is somewhat strategic; they can shift their response strategy to rely more heavily on iconicity when it is beneficial to do so. This is consistent with other evidence for attentional control and strategic reliance on relevant lexical and semantic variables in LDT (Balota, Paul, & Spieler, 1999; Hargreaves & Pexman, 2012).

The results from Experiments 1 and 2 demonstrate that iconicity can confer an advantage in language processing, in addition to aiding in language learning (e.g., Imai et al., 2008) and increasing the vividness of communication (e.g., Lockwood & Tuomainen, 2015). Vocal iconicity has been also argued to have played a role in language evolution (Perlman et al., 2015; Perlman & Lupyan, 2018). Various factors have been proposed to act on the relative balance of iconicity and arbitrariness in the lexicon over generations (for discussions of this topic see Perniss & Vigliocco, 2014; Sidhu & Pexman, 2018a; Winter et al., 2017). One may speculate that the facilitatory role of iconicity in processing is one such factor, and that it might increase the chances of iconic forms being maintained over time.

In addition to testing for a benefit of iconicity, we also hoped to gain insight into how such an effect would fit within existing models of word recognition. We speculated that iconicity may have an effect via links from phonology to semantics that are special in some way (see also Meteyard et al., 2015; Vigliocco & Kita, 2006; for evidence of phonemes activated cross modal information see Lockwood, Hagoort, & Dingemanse, 2016): iconic words could possess extra links (i.e., a quantitative difference) or links that are more direct in nature (i.e., a qualitative difference). To investigate this possibility, we examined whether increasing participants' reliance on phonology by degrading visual stimuli (in Experiment 1) or having participants respond based on phonology (Experiment 2) increased the effects of iconicity. We found no evidence that these manipulations increased effects of iconicity. It is worth noting however that, while non-significant (p = .06), the numerical trend that emerged when comparing the clear condition of Experiment 1 with the phonological lexical decision task in Experiment 2, with iconicity playing a larger role in the latter.

The lack of a phonological effect speaks against the possibility that iconic words have *extra* links from phonology to semantics. Were this the case, increasing participants' reliance on phonology should have allowed these extra links to facilitate processing. However, it still may be the case that iconic words have more *direct* links from phonology to semantics, and that these links facilitate processing regardless of the extent to which words are processed phonologically. That is, while a null result with regards to the interaction between phonology and iconicity speaks against iconic words having *quantitatively* different links as compared to non-iconic words, it does not necessarily rule out the possibility that iconic words have *qualitatively* different links. Additionally, recall that in Experiment 2, pseudohomophones based on words with a higher iconicity were responded to faster and more accurately. This suggests some role of phonology as correctly identifying a pseudohomophone as sounding like a word in the PLDT largely depends on the processing of phonology as the stimulus has no extant mapping from orthography to semantics.

Of course, it is also simply possible that effects of iconicity do not arise from special links between phonology and semantics. If so, how else might we account for them? One possibility is that the semantic representations of iconic words are special in some way. For instance, Meteyard et al. (2015) speculated that iconic words may have additional connections from semantic representations to modality-specific features. Indeed, previous work has shown relationships between iconicity and sensory experience (Sidhu & Pexman, 2018; Winter et al. 2018) and concreteness (Lupyan & Winter, 2018). As iconicity ratings become available for a greater number of items, future research should explore the role of various semantic dimensions in processing advantages of iconic words.

6. Conclusion

We found that iconicity provides a benefit to visual word recognition in typical adults. Thus, iconicity contributes to language processing and should be considered in models of word recognition. On a larger scale, these findings demonstrate another benefit of iconicity to language.

Acknowledgements

The authors thank Mark Dingemanse and Darin Flynn for helpful correspondences about matters related to this work. The authors also thank Kristen Deschamps and Stella Heo for assistance running the experiments.

References

Baayen, H., Bates, D., Kliegle, R., & Vasishth, S. (2015). RePsychLing: Data sets from Psychology and Linguistics experiments. R package version 0.0.4.

Balota, D. A., Paul, S., & Spieler, D. H. (1999). Attentional control of lexical processing pathways during word recognition and reading. In S. Garrod & M. Pickering (Eds.), *Language Processing,* (pp. 15-57). East Sussex, UK: Psychology Press Ltd.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv:1506.04967 [stat]. Retrieved from http://arxiv.org/abs/1506.04967

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*, 1-48.

Brysbaert, M., & New, B. (2009). 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*, 977-990.

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

de Saussure, F. (1916). The nature of the linguistic sign. In P. Meisel, & H. Saussy (Eds.), *Course in General Linguistics* (pp. 65–70). New York, NY: Coumbia University Press.

Dingemanse, M. (2012a). Advances in the Cross-Linguistic Study of Ideophones. *Linguistics and Language Compass*, *6*, 654-672. doi:10.1002/lnc3.361

Dingemanse, M. (2012b). *Coerced iconicity in writing and speech*. Retrieved from https://semioticon.com/semiotix/2012/07/coerced-iconicity-in-writing-and-speech/

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics*, *3*.

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. doi:10.1016/j.tics.2015.07.013

Hargreaves, I. S., & Pexman, P. M. (2012). Does richness lose its luster? Effects of extensive practice on semantic richness in visual word recognition. *Frontiers in Human Neuroscience, 6: 234*. doi: 10.3389/fnhum.2012.00234

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662-720.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1331-1356.

Hockett, C. (1963). The problem of universals in language. In J. Greenberg (Ed.), *Universals of Language* (pp. 1-22). Cambridge, MA: M.I.T. Press.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*, 20130298.

Jaeger, B. (2017). r2glmm: Computes R Squared for Mixed (Multilevel) Models. R package version 0.1.2.

Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How sound symbolism is processed in the brain: A study on Japanese mimetic words. *PLoS ONE*, *9*, 1–8.

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

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*, 1-26.

Laing, C.E., Viham, M. & Portnoy, T.K. (2017). How salient are onomatopoeia in the early input? A prosodic analysis of infant-directed speech. *Journal of Child Language*, *44*, 1117-1139.

Leeper, T. J. (2018). prediction: Tidy, Type-Safe 'prediction()' Methods. R package version 0.3.6.

Deleted: ¶

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: Neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*, *2*, 1–15. doi: 10.1525/collabra.42

Formatted: Font: Italic

Lockwood, G., & Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. *Frontiers in Psychology*, *6*. doi:10.3389/fpsyg.2015.00933

Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. *Neuropsychologia*, *76*, 264-275.

Peeters, D. (2016). Processing consequences of onomatopoeic iconicity in spoken language comprehension. In *38th Annual Meeting of the Cognitive Science Society (CogSci 2016)* (pp. 1632-1647). Cognitive Science Society.

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

Perlman, M., & Lupyan, G. (2018). People can create iconic vocalizations to communicate various meanings to naïve listeners. *Scientific Reports*, *8*, 2634.

Perniss, P., Lu, J. C., Morgan, G., & Vigliocco, G. (2018). Mapping language to the world: The role of iconicity in the sign language input. *Developmental Science*, *21*.

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society of London*. *Series B, Biological Sciences*, *369*, 20130300.

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*. doi:10.1371/journal.pone.0137147

Perry, L. K., Perlman, M., Winter, B., Massaro, D. W., & Lupyan, G. (2018). Iconicity in the speech of children and adults. *Developmental Science*, *21*. doi:10.1111/desc.12572

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision.

Journal of Experimental Psychology: Learning, Memory, & Cognition, 27, 139-156.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28,* 572-584.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Revill, K. P., Namy, L. L., DeFife, L. C., & Nygaard, L. C. (2014). Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI. *Brain and Language*, *128*, 18-24.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*, 225-239.

Sidhu, D. M., & Pexman, P. M. (2018a). Lonely sensational icons: semantic neighbourhood density, sensory experience and iconicity. *Language, Cognition and Neuroscience*, *33*, 25-31.

Sidhu, D. M., & Pexman, P. M. (in press). Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review*, 1-25.

Singmann, H., Bolker, B., & Westfall, J. (2015). afex: Analysis of factorial experiments. R package version 0.23-0.

Vigliocco, G., & Kita, S. (2006). Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes*, *21*, 790-816.

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

Yap, M. J. (2007). Visual word recognition: Explorations of megastudies, multisyllabic words, and individual differences. Unpublished doctoral dissertation, Washington University in St. Louis, St. Louis, MO.

Yap, M. J., Lim, G. Y., & Pexman, P. M. (2015). Semantic richness effects in lexical decision: the role of feedback. *Memory & Cognition*, *43*, 1148-1167.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*, 971-979.