

Do sound symbolism effects for written words relate to individual phonemes or to phoneme features?

Padraic Monaghan^{1,2,3} & Matthew Fletcher¹

¹Lancaster University, UK

²University of Amsterdam, The Netherlands

³Max Planck Institute for Psycholinguistics, The Netherlands

Correspondence to:

Padraic Monaghan

Department of Psychology

Lancaster University

Lancaster LA1 4YF

UK

Tel: +44 1524 593813

Fax: +44 1524 593744

Email: p.monaghan@lancaster.ac.uk

Abstract

The sound of words has been shown to relate to the meaning that the words denote, an effect that extends beyond morphological properties of the word. Studies of these sound-symbolic relations have described this iconicity in terms of individual phonemes, or alternatively due to acoustic properties (expressed in phonological features) relating to meaning. In this study, we investigated whether individual phonemes or phoneme features best accounted for iconicity effects. We tested 92 participants' judgments about the appropriateness of 320 nonwords presented in written form, relating to 8 different semantic attributes. For all 8 attributes, individual phonemes fitted participants' responses better than general phoneme features. These results challenge claims that sound symbolic effects for visually presented words can access broad, cross-modal associations between sound and meaning, instead the results indicate the operation of individual phoneme to meaning relations. Whether similar effects are found for nonwords presented auditorially remains an open question.

Keywords: sound symbolism, iconicity, cross-modal correspondences, phonology, psycholinguistics

Do Sound Symbolism Effects for Written Words Relate to Individual Phonemes or to
Phoneme Features?

Introduction

In contrast to the traditional assumption in linguistics that the relationship between the spoken form of a word and its meaning is arbitrary, there has been growing interest and evidence of the widespread existence of systematicity and iconicity in spoken language (Blasi, Wichmann, Hammerstom, Stadler, & Christiansen, 2016; Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Hinton, Nichols, & Ohala, 1994; Monaghan, Shillcock, Christiansen, & Kirby, 2014; Nuckolls, 1999; Perry, Perlman, & Lupyan, 2015; Sapir, 1929). Systematicity refers to a non-arbitrary mapping between distinctions in speech sound and distinctions in the meaning of their references, and can be established in terms of determining statistical relations between sound and meaning. Iconicity, on the other hand, refers to a goodness of fit between the sound properties of a word and the object to which it refers, and can be tested by participants' judgments of appropriateness of a label for a meaning, or guesses at the meaning of a given label (Cuskley, 2013; Dingemanse et al., 2015; Perniss, Thompson, & Vigliocco, 2010; Taylor & Taylor, 1965).

There are now numerous studies indicating that particular sounds are associated with certain distinctions in meaning in terms of participants' judgments of matches between labels and their referents. This has been exemplified in classic demonstrations that different sounding nonwords, such as *maluma* and *takete*, are judged to relate to rounded and angular shapes, respectively (Brand, Monaghan, & Walker, 2018; Fort, Martin, & Peperkamp, 2015; Köhler, 1929; Nielsen & Rendall, 2011, 2013). Such sound symbolic associations between forms of words and

meanings have now been shown for a range of semantic distinctions, for reviews see French (1977) and Lockwood and Dingemans (2015).

However, investigations of sound symbolism have traditionally been explored by relating particular sounds to meaning attributes using very small numbers of stimuli. These small stimulus numbers make it difficult to determine exactly which aspect of the sounds are relating to meaning, and may result in confounds between different phonological properties of the stimuli. For instance, for the nonwords *maluma* and *takete*, relating to rounded and angular shapes, there are differences in the phonological form of the words in terms of the consonants' place of articulation, manner of articulation, and voicing, as well as vowel height and position. Exceptionally, some studies have aimed to isolate precisely which properties of the sound relate to meaning. Nielsen and Rendall (2013) controlled the stimuli relating to angular and rounded objects and determined that both sonorant consonants and rounded vowels related more to rounded shapes, Fort et al. (2015) found consonant features drove judgments more strongly than vowel properties, and d'Onofrio (2014) found that both vowel position, consonantal voicing, and place of articulation all related differently to rounded compared to angular shapes. In a series of studies, Klink (2000, 2001, 2003), Klink and Athaide (2012), and Klink and Wu (2013) measured participants' judgments about sets of written nonwords that varied in terms of either manner of articulation (plosives versus fricatives), or voicing of consonants, or vowel position of the pronunciation of these nonwords. Klink (2000, 2001) showed that fricatives were more likely than plosives to relate to attributes of small, light, and feminine, and that voicing was related to large, and masculine, results that were confirmed in the subsequent studies (Klink & Athaide, 2012; Klink & Wu, 2013).

There has been a preponderance of sound symbolism studies that have focused on manner of articulation (Fort et al., 2015; Monaghan, Mattock, & Walker, 2012; Nielsen & Rendall, 2011, 2013), but *place* of articulation has rarely been considered, though see d’Onofrio (2014) for an exception, where labial and velar consonants were found to relate more closely to rounded shapes and alveolar consonants were found to relate more closely to angular shapes. In all these studies, however, place of articulation may be confounded with manner due to the small sets of experimental stimuli used. Furthermore, the relative contribution of different phoneme features – i.e., manner and place of articulation and voicing – have not been assessed simultaneously. Examining effect sizes of results in Klink’s (2000, 2001) and Klink and Wu’s (2013) studies suggest that voicing may be a more powerful effect than manner of articulation (when comparing fricatives to plosives), but there has been no direct comparison in those studies. A first aim of the current study was thus to compare the extent to which manner of articulation, place of articulation, and voicing of consonants related to participants’ judgments about the effectiveness of certain speech sounds, prompted by written presentation of words, relating to meaning attributes.

Alongside studies of phoneme features of speech sounds, a parallel tradition in sound symbolism studies has investigated the role of particular phonemes in reflecting meaning attributes of words, characterised by work on phonaesthemes (Bergen, 2004). For example, in English, specific consonant combinations are commonly associated with certain meanings; words beginning with /fl/ are frequently associated with movement (e.g. *fly, fling, flap, ...*) and words beginning with /gl/ are often linked with vision (e.g. *glow, glare, glitter, ...*) (Jespersen, 1922). Bergen (2004) demonstrated that such phonaesthemes had observable effects on participants’ lexical

access and processing. Otis and Sagi (2008) quantified the extent to which several proposed phonaesthemes corresponded to meaning distinctions in English. They found evidence of statistical correspondences between phoneme clusters and meanings for 27 of the 46 phonaesthemes that were tested, supported by broader analyses by Abramova, Fernandez, and Sangati (2013), Cassani, Chuang, and Baayen (2019), and Monaghan, Lupyan, and Christiansen (2014). The existence of such phonaesthemes appears to be general across languages (Dingemanse et al., 2015). Furthermore, Blasi et al. (2016) found that particular groups of phonemes were associated with certain meanings across a wide range of languages, such as high front vowels and /tʃ/ tending to occur in words for small, and /r,ɹ,ɻ,ʀ,r/ featuring in words expressing round.

These studies of individual phonemes, or phoneme clusters, relating to meanings raise the question of whether the phoneme to meaning correspondences are due to the phoneme features of the phonemes, or due to the unindividuated phonemes themselves. Alternative theories of the origins and effects of sound symbolism in language make different predictions here (Spence, 2011). Theories that relate to cross-modal correspondences between dimensions of sound and dimensions of non-auditory perceptual domains are consistent with accounts of sound symbolism residing in general characteristics of speech, reflected in phoneme features. For instance, the observation that vowel position relates to small versus large expressives across languages (Ultan, 1978) has been interpreted as due to differences in perceptions of pitch between front and back vowels, which in turn are symbolic of size due to associations between sound pitch and size (Ohala, 1994; Ultan, 1978; Walker, 2016).

In contrast, if the relation between sound and meaning is driven by individual phonemes rather than their phonological properties, then this is problematic for

perspectives that relate particular speech properties to general cross-modal correspondences driven by acoustic characteristics (such as perception of pitch, Klink, 2000; Ohala, 1994; Walker, 2016) that are then exapted as a special case into language processing. For example, the distinction between *maluma* and *takete* has been characterised as the distinction between continuant and obstruent consonants, which are acoustically realised in terms of the suddenness of the onset and offset envelope of speech (Rhodes, 1994), which can also be taken to suggest a physical analogy between temporal properties of sound and visual features. However, if the effects are found not to be related to general characteristics of the phonemes but rather due to judgments about particular phonemes, then this means that effects are better described in terms of speech-specific associations, rather than more general cross-modal associations (Sidhu & Pexman, 2018).

It is generally assumed that written words can access these broad cross-modal associations. For instance, the literature on sound symbolism in brand names tends to assume that consumers' decisions are affected by acoustic properties of written brands (Klink, 2000, 2001, 2003; Klink & Athaide, 2012; Klink & Wu, 2013; Shrum, Lowrey, Luna, Lerman, & Liu, 2012). A second aim of our study was to directly compare whether phoneme features, or individual phonemes, drive participants' iconicity judgments about sound to meaning relations for novel written words.

Recently, Westbury, Hollis, Sidhu, and Pexman (2018) made valuable progress in addressing these issues of the importance of variation in stimuli and contrasting the role of phonemes and phoneme features in explaining sound symbolism responses. They investigated relations between several phonemes in nonwords and several semantic attributes: rounded and sharp, large and small, and masculine and feminine. Further attributes of concrete/abstract and high/low valence were also tested,

indirectly, by judging the appropriateness of spoken and written nonwords as names of referents that were abstract or concrete, or high or low in valence. Participants were required to make a forced choice between whether a nonword was effective or not in promoting that attribute. The results were then analysed by constructing binary logistic regression models with either phoneme features (manner and place of articulation features for consonants, and height and position features for vowels), letters and pairs of letters, phonemes and pairs of phonemes, or combinations of all these predictors. For each model, all predictors were initially introduced, then individual predictors that did not significantly contribute to distinguish the binary categories were removed, until only significant contributors were left. Westbury et al. (2018) found that for each of the three antonym pairs, the most accurate (in terms of most hits and fewest false alarms) statistical models tended to be those based on phonemes and pairs of phonemes, or the combination models.

However, Westbury et al.'s (2018) study did not explicitly compare which phoneme features might account for most variance in participants' responses, but counting the number and weights of particular phoneme feature parameters can go some way to indicating where those effects may lie. For instance, for their statistical model of the attribute *round*, three place of articulation features (labiodental, alveolar, velar) were found to be significantly negatively related, and one manner of articulation feature (nasal) was found to be a significant positive predictor. Furthermore, their analyses did not compare the extent to which phoneme features could explain the variance in the data in comparison to individual phonemes. Instead, the analyses resulted in hybrid models that incorporated phoneme features, individual phonemes, pairs of phonemes, and orthographic letters and pairs of letters.

Determining the predictive power of phoneme features, or individual phonemes is not yet discernible from these data.

Comparing analyses based on phoneme features versus analyses of individual phonemes requires taking into account differences in the degrees of freedom introduced by each analysis. Phoneme features vary on few dimensions, and so an analysis using features adds only a few degrees of freedom to the model fitting the data. However, considering phonemes individually can add substantially more degrees of freedom to the model if the stimuli vary over many phonemes. Yet, it is possible to utilise model comparisons to determine which of alternative models can provide a better fit to the data, taking into account different degrees of freedom. We thus can compare models based on phoneme features to models based on individual phonemes in accounting for the observed data.

In the current study, we assessed the predictive role of consonant manner of articulation, place of articulation and voicing, individually and combined, in participants' judgments about the relation between properties of written nonwords and their appropriateness in terms of promoting particular meaning attributes. We also compared models based on phoneme features to models fitted with individual phonemes as predictors, to determine whether these sound symbolism effects were better characterised by properties over sets of phonemes, or by individual phonemes. We did not directly manipulate the vowels within the nonwords used in the study, but controlled for their contribution to the participants' behaviour in the analyses by including the identity of the template as a random effect in the analyses. Our focus here is on the consonantal properties of speech, but future research could also investigate the role of vowels, in addition to the effects of the consonants reported here.

Method

Participants

One-hundred and twenty participants who were students or associates of Lancaster University completed the study. After screening for responses (see below) 28 participants were removed from the analysis. Of the 92 remaining participants, 41 were male, 50 female, and 1 who self-identified as transgender, with age $M = 28.8$ years, $SD = 12.1$, range 18-73 years. All participants had University-level English, of whom the majority were native English speakers ($n = 77$), and the remainder spoke Mandarin ($n = 4$), Greek ($n = 2$), Arabic ($n = 1$), Konkani ($n = 1$), Malayan ($n = 1$), Somali ($n = 1$), Spanish ($n = 1$), Tamil ($n = 1$), Thai ($n = 1$), Russian ($n = 1$), and Welsh ($n = 1$) as first languages. Participants were recruited online via the distribution of an anonymous survey link through social media.

Materials

We tested eight different semantic attributes in four antonym pairs: small/large, soft/hard, fast/slow, and masculine/feminine which had been shown to relate to different sound symbolic properties in previous work (Cuskley, 2013; Klink, 2000; Lockwood & Dingemanse, 2015; Westbury et al., 2018). In terms of iconicity, higher-frequency sounds have been related to smaller sized referents (Ohala, 1984; Sapir, 1929), which have also been related to properties of softness, faster speed, and femininity (Klink, 2000). Fricatives and unvoiced consonants are perceived as higher in frequency than plosives, and voiced consonants, respectively (Hinton et al., 1994), and so a prediction would be that voicing and plosives would relate more closely to large, hard, slow, and masculine properties. For each semantic attribute, 40 nonwords were constructed which comprised four nonword templates with each of ten letters

(corresponding to different phonemes) beginning the nonword. Table 1 shows the templates used for each attribute. The nonword templates were taken from Klink (2000), and were designed to contain a range of vowels and other consonants to ensure the generalisability of the effects observed. The templates were bisyllabic, with one exception, and spanned at least four different vowels. The templates for each semantic attribute also included at least five different consonants varying in place and manner of articulation. The onset phonemes were selected to vary in manner of articulation (plosives versus fricatives), place of articulation (bilabial, labiodental, alveolar, velar) and voicing. The first phoneme was selected as the sound symbolism inbed, as word-initial phonemes have been shown to influence sound symbolism judgments to a greater degree than word-medial phonemes (Klink & Wu, 2013). The phonemes were /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /z/. All combinations of initial phonemes and word templates were nonwords in English.

Table 1. The semantic attributes, word template stimuli and catch trial words used in the experiment.

Attribute	Word Template				Catch Trial
Smallness	_itav	_olaw	_olud	_urley	tiny
Largeness	_etib	_elom	_upah	_uxir	huge
Softness	_avuz	_eley	_ewok	_undel	fluffy
Hardness	_aleck	_anup	_obal	_ovem	rock
Fast Speed	_emov	_ivad	_oyan	_uzam	zoom
Slow Speed	_elan	_ifin	_ilax	_olaw	sloth
Masculinity	_alaf	_emeg	_ewik	_oluf	man
Femininity	_arucky	_ockle	_ulay	_uxen	woman

For example, for the attribute small, ten nonwords ending in *_itav* and each beginning with one of the ten phonemes were used (e.g., *pitav*, *bitav*, ..., *sitav*, *zitav*), along with ten other nonwords each ending *_olaw*, *_olud*, or *_urley*. Order of nonwords for each attribute were randomised for each participant, and order of the semantic attributes were also randomised. The use of different templates enables us to determine whether the effects of phoneme features, or individual phonemes, are general across a range of word stimuli. If a particular onset consonant and template combination was adversely affecting the results then the fixed effects of phoneme features or individual phonemes would not be significant when also including the intercept and slope as a random effect of template in the analyses. The analyses using random effects of nonword template also enables us to determine that the use of different templates for each nonword did not adversely affect the observed results.

Eight items could be interpreted as pseudohomophones (*kockle* (*cockle*), *bundel* (*bundle*), *bobal/gobal* (*bobble*, *gobble*), *burley/gurley/kurley/surley* (*burly*, *girly*, *curly*, *surly*). Another contributor to possible links between particular nonwords and semantic attributes is the similarity in the form of the words, in terms of whether the nonword and the attribute began with the same letter or not. A further possible contributor to performance is repeat of the phonemes in nonwords such as *fifin* or *zuzam*. We repeated the following analyses excluding the pseudohomophones, the shared phoneme between nonword and semantic attribute, and nonwords where the first phoneme was reduplicated. The key results remained the same (see Supplementary Materials 2).

In order to determine whether participants were responding according to the meaning of the stimuli, one catch trial per attribute was also included as a semantically-related real word. These are shown in Table 1.

Procedure

Participants were required to respond according to how appropriate a brand name was for promoting the idea of each semantic attribute, the precise instructions were “How appropriate are these novel brand names for promoting the idea of X?”, where *X* was the semantic attribute in question. We gave no further guidance on how to interpret appropriateness as we did not want to lead participants to respond explicitly and strategically to graphical or auditory features of the nonwords. Each semantic attribute was tested in turn, with 40 nonwords presented as potential brand names for that attribute. Participants responded to each nonword using an ordinal 7-point Likert scale (from 1 = “Not Appropriate” on the left to 7 = “Very Appropriate” on the right), and they responded by selecting a radio button corresponding to the number of their response. For each participant, the order of the nonwords was completely randomised. We added one catch trial to each list of 40 nonwords, which was a real word relating in meaning to the attribute. This was included to ensure that the participant was responding to the attribute. After responding to the 40 nonwords and the catch trial, the next attribute was presented with 40 new nonwords and the next catch trial. The stimuli were presented visually on a computer using Qualtrics software.

We omitted participants whose responses used exactly the same appropriateness rating for 80% or more of the brand names, because these participants stood in stark contrast with the variation demonstrated by the majority of participants,

and suggested that they were not responding on the basis of the individual stimuli. This removed 28 of the participants. All remaining participants responded appropriately to the catch-trial items.

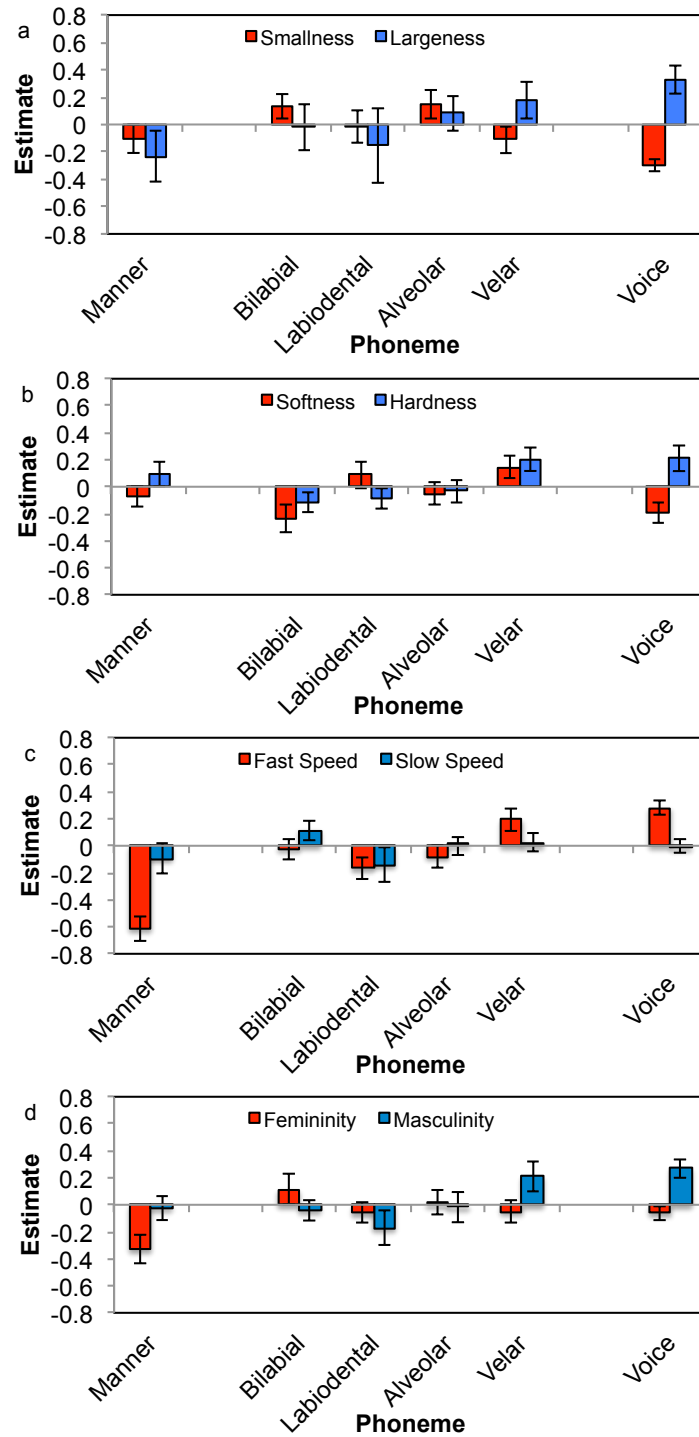


Figure 1. Estimates from the MLE model of the appropriateness of phoneme features relating to each semantic attribute. Values below zero indicate that nonwords containing the phonological feature are judged to be negatively related to the attribute, values above zero indicate that the phonological feature is judged to relate positively.

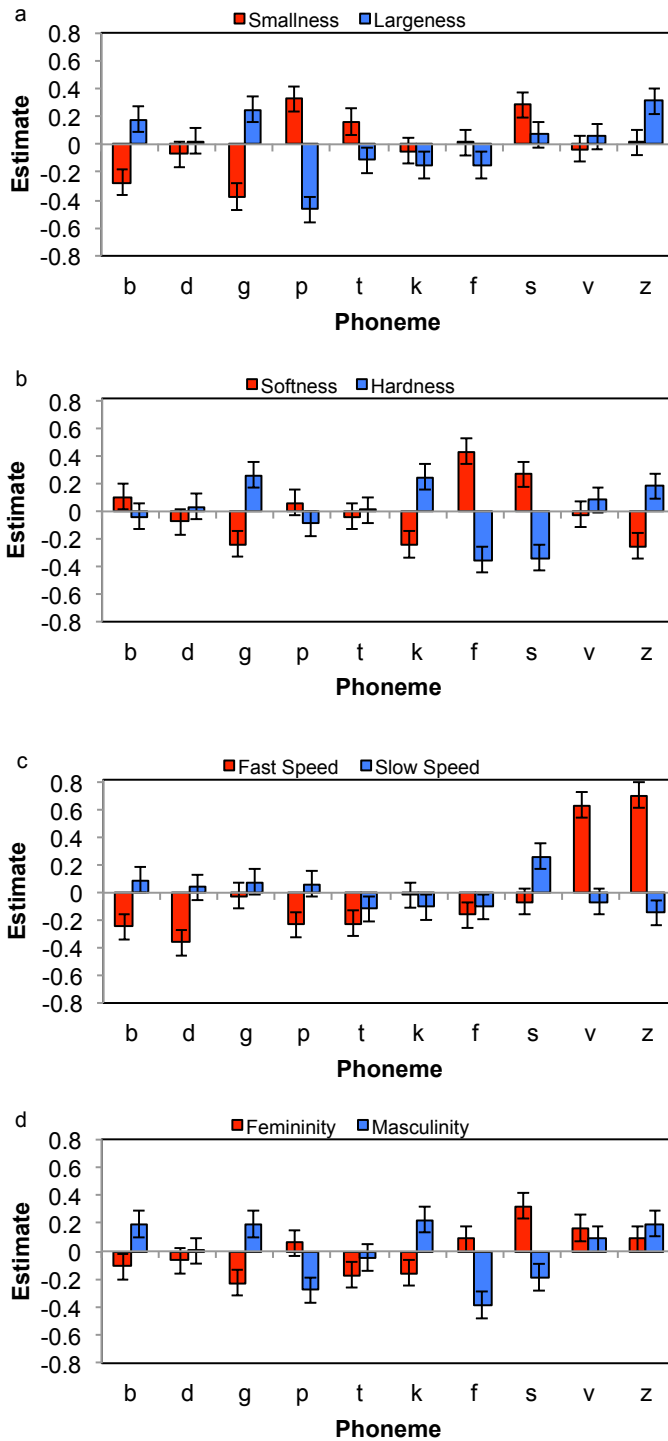


Figure 2. Estimates for appropriateness of each phoneme relating to each semantic attribute. Estimates above zero indicate that nonwords containing the phoneme are judged to relate negatively to the attribute, positive values indicate judgments that the phoneme does relate to the attribute.

Results

In the analyses, each meaning attribute was considered individually, as pairs of antonyms do not necessarily operate in opposite ways in terms of relations to sound properties (e.g., Westbury et al., 2018). For each attribute a series of linear mixed effects models were constructed, with participants' appropriateness rating for each nonword as the dependent variable. Participant and nonword template were entered as random effects, with intercept and random slopes for each fixed effect (Baayen, Davidson, & Bates, 2008). When the model failed to converge, we omitted random slopes for participants, and if the model still failed to converge we omitted random slopes for nonword template. We used the nlptr optimiser for convergence (Johnson, 2018). For the phoneme features model, we tested the model containing fixed effects of manner, place, and voicing to models where one of the phoneme features was removed, determining whether this significantly reduced model fit using log-likelihood comparison tests. For the phoneme model, we tested whether the fixed effect of phoneme improved model fit compared to a model containing only random effects.

To compare phoneme features and phonemes in terms of their explanation of variance in the data, a model with fixed effects of phoneme features and random intercepts only (so no slopes for the fixed effects) was compared to a model with

fixed effects of phonemes and the same random effects structure. This is because model comparison has to be made with the same random effects structure, and including random slopes for either phonemes or phoneme features may have introduced bias into the model comparison. The full model results are reported in Supplementary Materials 1.

Smallness and largeness

For smallness, manner of articulation did not significantly improve model fit, $\chi^2(1) = 1.1525, p = .283$. Place of articulation marginally improved model fit, $\chi^2(3) = 7.6509, p = .05381$. Voicing improved model fit, $\chi^2(1) = 12.209, p < .001$. Figure 1a shows the centred estimates for the features for smallness, in the red columns. Values that are greater than zero indicate that participants judged the feature to be more appropriate for expressing smallness than their mean response to the nonwords, values less than zero indicate judgments that the feature was less appropriate. To convert the estimates to responses on the 7 point Likert scale, the estimates can be added to the intercept from the model (see Supplementary Materials 1). For example, for the phoneme feature model, the intercept was 3.58, so the mean rating for nonwords containing voiced consonants was $3.58 - 0.30 = 3.28$. Words containing unvoiced consonants were judged to be more appropriate, and words containing bilabial, labiodental, or alveolar consonants were marginally more appropriate than words containing velars. For the phoneme model, phonemes improved model fit compared to the model containing only random effects, $\chi^2(9) = 51.248, p < .0001$. Figure 2a shows the estimates for each phoneme in predicting the attribute.

Comparisons of model fit for the phoneme feature model and the phoneme model demonstrated a significant improvement in model fit for the phoneme model,

$\chi^2(4) = 16.718, p = .002$. Thus, judgments about nonwords relating to smallness are better explained by individual phonemes rather than more general phoneme features.

For largeness, for the phoneme feature analysis, manner of articulation did not significantly improve model fit, $\chi^2(1) = 1.5809, p = .209$. Place of articulation did not improve fit, $\chi^2(3) = 4.1983, p = .241$, but voicing did improve fit, $\chi^2(1) = 5.9163, p = .015$, with voicing relating to largeness, operating in the opposite direction to judgments about smallness, see Figure 1a. For the phoneme analysis, phonemes improved model fit compared to a model containing only random effects, $\chi^2(9) = 102.86, p < .001$, with /b, g, z/ relating positively and /p, t, k, f/ relating negatively to largeness, see Figure 2a.

The phoneme model fit was significantly better than the phoneme feature model fit, $\chi^2(4) = 17.807, p = .001$, so, as for the smallness comparison, responses were best explained in terms of individual phonemes rather than general phoneme features.

Softness and hardness

For softness, analysing the phoneme features, manner of articulation was not significant, $\chi^2(1) = .769, p = .380$, but place of articulation, $\chi^2(3) = 11.102, p = .011$, and voicing, $\chi^2(1) = 4.378, p = .036$, were significant, with unvoiced and velars judged softer than voiced and bilabials, see Figure 1b. The Figure illustrates the estimates for each phoneme feature in terms of appropriateness of nonwords containing the feature reflecting the attribute of softness.

For the analysis of individual phonemes, phonemes explained significant variance, $\chi^2(9) = 103.79, p < .001$, with /f, s/ relating positively, and /g, k, z/ relating negatively to softness, see Figure 2b. In a comparison between the phoneme model and the phoneme feature model, the phoneme model was found to be a better fit to the data, $\chi^2(4) = 33.776, p < .001$.

For hardness, manner of articulation was not significant, $\chi^2(1) = 2.073, p = .150$, but place of articulation, $\chi^2(3) = 25.237, p < .001$, and voicing, $\chi^2(1) = 25.397, p < .001$, were significant, with velars and voicing relating positively to hardness. Phonemes contributed significantly to explaining variance in the phoneme model, $\chi^2(9) = 95.191, p < .001$, demonstrating the inverse pattern to softness (see Figure 2b), and once again the comparison between the phoneme model and the phoneme feature model revealed that the phoneme model was a better fit to the data than the phoneme feature model, $\chi^2(4) = 28.233, p < .001$.

Fast and slow speed

For fast speed, all phoneme features were significant. For manner of articulation, $\chi^2(1) = 76.169, p < .001$, for place of articulation, $\chi^2(3) = 18.490, p < .001$, and voicing, $\chi^2(1) = 39.886, p < .001$, with fricatives, voicing, and velars relating positively to the fast speed attribute. Figure 1c shows the appropriateness of each phoneme feature relating to the attribute fastness. For the phoneme model, phonemes were significant, $\chi^2(9) = 247.03, p < .001$, with /v, z/ relating positively and /b, d, p, t/ relating negatively to fast speed. In model comparisons, the phoneme model was a better fit to the data than the phoneme feature model, $\chi^2(4) = 87.787, p < .001$.

For slow speed, none of the phoneme features related significantly, manner, $\chi^2(1) = 1.922, p = .167$, place, $\chi^2(3) = 7.285, p = .063$, voicing, $\chi^2(1) < .001, p = .980$.

For the phoneme model, phonemes significantly contributed to explaining variance, $\chi^2(9) = 30.291, p < .001$, with /s/ relating positively and /t, z/ relating negatively to slow speed, see Figure 2c. Again, in a comparison of the phoneme and phoneme feature models, the phoneme model explained the data better than the phoneme feature model, $\chi^2(4) = 22.786, p < .001$.

Femininity and masculinity

For femininity, manner of articulation was significant, $\chi^2(1) = 5.756, p = .016$, see Figure 1d, with fricatives relating positively. Place of articulation, $\chi^2(3) = 3.871, p = .276$, and voicing, $\chi^2(1) = 1.233, p = .267$. Phonemes in the phoneme model were significant, $\chi^2(9) = 62.382, p < .001$, see Figure 2d, with /g, t, k/ relating negatively and /s, v/ relating positively to femininity. The comparison of models showed that the phoneme model was a better fit than the feature model, $\chi^2(4) = 10.034, p = .040$.

Finally for masculinity, manner of articulation was not significant, $\chi^2(1) = .137, p = .711$, but place of articulation, $\chi^2(3) = 22.706, p < .001$, and voicing, $\chi^2(1) = 41.396, p < .001$, were both significant predictors, with velars and voicing relating positively. Phonemes explained significant variance in the phoneme model, $\chi^2(9) = 98.377, p < .001$, with /b, g, k, z/ relating positively, and /f, s/ relating negatively to masculinity. As with all the other model comparisons, the phoneme model was a better fit to the data than the phoneme feature model, $\chi^2(4) = 26.643, p < .001$.

Discussion

We tested the extent to which phoneme features or individual phonemes better explained participants' preferences for the appropriateness of written nonwords in expressing a range of attributes. We found that, for eight different attributes drawn

from four antonym pairs, various phoneme features and individual phonemes explained significant variance in participants' responses. Thus, effects of sound symbolism across a broad set of attributes were effectively captured in terms of the sounds contained in the words. The only exception was that no phoneme features significantly predicted variance in responses to slow speed. Thus, previous explorations of effects of sound symbolism in relating to semantic attributes (Bergen, 2004; Klink & Wu, 2014; Lockwood & Dingemanse, 2015; Nielsen & Rendall, 2013) are shown to be extensive and prevalent across a range of different meanings.

As noted by Westbury et al. (2018) the relations between particular speech sound properties and antonyms were sometimes diametrically opposed but not always. One possible contribution to this may have been the spatial arrangement of the Likert scale responses (not at all appropriate always appeared on the left of the display, and very appropriate always appeared on the right of the display), which may have affected complete reversal of effects. For instance, participants may have a preference for small at the left of (any) scale. For smallness and largeness, we observed opposing effects for velar place of articulation (more related to largeness) and voicing (more related to largeness), consistent with Klink (2000, 2001, 2003). At the phoneme level, we found /b, g/ relating positively to largeness and negatively relating to smallness, and /p, t/ relating positively to smallness and negatively to largeness. Thus, the voicing effect appeared to be specific to only some phoneme contrasts.

For the softness and hardness antonym pair, again the features of velar and voicing related positively to hardness and negatively to softness. For the phonemes, /g,k,z/ were positively related to hardness and negatively related to softness, and /f,s/

related positively to softness and negatively to hardness, similar to trends in the results of Klink (2000).

For fast and slow speed, no phoneme features related significantly to slow speed, but for the individual phoneme model, /v,z/ related positively to fast and negatively to slow speed, and /b,d,s/ showed trends for positive relations to slow and negative to fast speed.

Finally, for the antonym pair of femininity and masculinity, no phoneme features related more closely to femininity than masculinity, but for the individual phoneme model, /s/ related positively to femininity and negatively to masculinity, and /b,g,k/ showed the opposite effect. Previous observations of relations between femininity and manner of articulation thus seem to be driven by particular phonemes rather than phoneme features (Klink, 2000).

Taken together, these analyses demonstrate that place of articulation of consonants relate to numerous meaning attributes in terms of iconicity, in addition to the established effects previously observed of relations between manner of articulation and meaning relations. However, the focus of previous studies on manner of articulation effects, e.g., continuants versus obstruents, or frication versus plosive (Fort et al. 2015; Monaghan et al., 2012; Nielsen & Rendall, 2011, 2013, though these particular studies relate to different attributes than those tested here) are shown in the current study to be less prevalent than effects of voicing, and alternate somewhat with the effects of place of articulation in the analyses. These mixed results demonstrate the volatility of analyses based on phoneme features, and point instead to the value of considering individual phonemes in explaining the data. Indeed, the current results highlight that design of sound symbolism studies that aim to test phoneme features

must include sufficient numbers of different phonemes to be assured that apparent feature effects are not instead due to the contribution of individual phonemes, an issue that may have led researchers to over-generalise their observed effects beyond the immediate items employed.

Interestingly, certain phonemes appeared to be particularly strongly related to several of the meaning attributes. For example, /s/ related positively to small, soft, slow, and feminine, and related negatively to hard, and masculine, and /g/ was positively related to large, hard, and masculine, and negatively related to small, soft, and feminine. These relations may be due to correspondences that already occur between meaning attributes – small and soft may be more related than small and hard, see Westbury et al. (2018) and Walker (2016) for further discussion. Nevertheless, the results highlight that phonemes are not limited to relate to only particular meanings, but can reflect iconicity between one sound and very many meaning relations.

Indeed, comparisons across the statistical models also enabled us to determine whether phoneme features or individual phonemes better explained participants' preferences for relations between written nonwords and meaning attributes. We observed that, for all meaning attributes tested in this study, participants' judgments about sound to meaning relations were best accounted for by individual phonemes rather than phoneme features. Though particular phoneme features may significantly relate to judgments, these do not appear to be general across all phonemes with that feature. For instance, for the largeness attribute, voicing is highly significant, but the voiced phonemes /d, z/ do not relate significantly positively to largeness, and nor do the unvoiced phonemes /t, s/ relate significantly negatively to largeness. Instead, particular phonemes, that are somewhat independent of their phoneme features, appear to symbolise the meaning.

One possibility is that interactions between phoneme features may be driving the effects. Our analyses considered phoneme features only as main effects. This was because an analysis including interactions between phoneme features would be identical to a phoneme-level analysis – each phoneme is individuated by its particular manner, place, and voicing features which would be distinguished in interactions between features. Though these explanations cannot be distinguished analytically, they can be distinguished theoretically. For explanations grounded in cross-modal associations between particular acoustic characteristics and meaning attributes, we do not expect that place of articulation should substantially affect the influence of voicing or of manner of articulation (e.g., Ohala, 1984; Rhodes, 1994). Phoneme feature accounts of sound symbolism would require an account of the acoustic effects of such interactions in explaining behaviour.

The finding that individual phonemes rather than phoneme features better reflect behavioural judgments about iconicity create tension for theories proposing that general cross-modal correspondences underwrite relations between speech sounds and meanings in studies of written word iconicity. Instead, the effects seem to be driven by more particular correspondences between individual phonemes and meanings, consistent with studies of phonaesthemes in language. This enables us to consider which of the several proposed mechanisms of sound symbolism characterised by Sidhu and Pexman (2018) best relates to our data. If sound symbolism for written words is expressed at the phoneme rather than the phoneme (or acoustic) feature level, then two of the mechanisms considered by Sidhu and Pexman (2018) can be ruled out as explanations for the current effects. The first proposed mechanism – that there are learned statistical associations between sounds and meanings – is consistent with our data. Participants may acquire the association

between particular phonemes and meaning distinctions, even though the actual relations between phonemes and meanings in natural language are minimal (Monaghan et al., 2014). The second proposed mechanism that there are shared properties between speech and meaning is difficult to reconcile with our data, as any such shared properties should be expressed better in phoneme features than phonemes. The third mechanism, that associations are due to similar neural activation (Marks, 1987) is also difficult to align with the current results, as it requires an explanation for why particular phonemes might generate neural activation that is distinct from the acoustic properties of the phoneme.

Though we have tested several semantic attributes in the current study – eight attributes in four antonym pairs – and measured the sound symbolic effects with a broad range of phonemes and phoneme features, we cannot be certain that our conclusions generalise to other observed sound symbolic effects, for three key reasons. First, it is possible that other sound symbolic distinctions that we have not tested may relate to speech sounds at the phoneme feature level, rather than the phoneme level, such as the angular/rounded distinction (Brand et al., 2018; Nielsen & Rendall, 2011, 2013). Second, the results may be differently realised if the referents were visually presented rather than just described by semantic attributes. It is possible that participants may be more drawn to the phoneme features rather than the individual phonemes if the referent is visually present, thus creating a more direct link between the perceptual modalities than in our study where the links are mediated by language (see Walker, 2016, for discussion of this point). Third, we have tested 10 different phonemes, but have not investigated all possible manner and place features that occur in segmental phonology. For instance, some studies of the bouba/kiki effect have contrasted sonorants with plosives, and these may result in greater differences

than those documented for fricatives versus plosives (e.g., Klink, 2000). Further work is clearly necessary to establish whether phoneme effects accounting for apparent phoneme features effects fricatives versus plosives and voiced versus unvoiced consonants apply too to approximants and nasals, and whether the four places of articulation indicate effects that generalise to all places.

The results show that written stimuli access phoneme-level over phoneme or acoustic feature level sound symbolic effects, and this have important implications for the extent to which sound symbolism effects are involved in, for instance, written brand names (e.g., Klink & Wu, 2013; Shrum et al., 2012). There is the possibility that visual features of the orthography of the nonwords may contribute to the effects. In order to measure the extent to which letter shape was affecting the results of our analyses, we included letter shape feature (angular or rounded) as defined in Cuskley et al. (2017) as an additional fixed factor in our analyses. As reported in Supplementary Material 3, the results indicate no substantial changes in the results, though there was in some cases an additional contribution from letter shape to account for variance (for soft, hard, fast, feminine and masculine), confirming Cuskley et al.'s (2017) and Westbury et al.'s (2018) observations that letters can contribute to behaviour in addition to the phoneme features that those letters denote. Though, for each of these attributes, the model containing only phonemes was still the best fitting model.

Conducting the study with auditory rather than visual stimuli would enable us to determine whether phonemes or phoneme features underwrite sound symbolism effects more broadly still. There is evidence to suggest that phoneme feature effects may be reduced in written compared to spoken nonwords. Cuskley, Simner, and Kirby (2017) showed that the effect of voicing was enhanced in spoken compared to written

presentation, when testing an angular/rounded distinction, and Bremner, Caparos, Davidoff, de Fockert, Linnell, and Spence (2013) found that the bouba-kiki effect was smaller in illiterate compared to literate participants. Nevertheless, the fact that we observed substantial phonological feature effects for semantic attributes that are consistent with studies tested with a variety of different methods (e.g., Lockwood & Dingemanse, 2015) suggests that the manner of testing may not be critical in revealing the observed effects. We found effects of phoneme features in all the semantic attributes we tested, and these were independent of the visual characteristics of the written nonwords. In all these cases we have tested, our conclusion stands: that phonological feature distinctions are better described in terms of phonemes to explain the goodness of fit between a label and the semantic attribute to which it refers. At the very least, our results highlight that studies presuming effects in terms of phonological or acoustic features need also to test the alternative explanation that results reside instead at the phoneme level.

The phenomena we have investigated have focused on segmental phonology – phoneme features, relating to acoustic properties, and phonemes. However, it is important to note the wealth of additional sound symbolic effects that are observable in suprasegmental phonology. For instance, there are large effects of duration, pitch, and timbre of speech that are also related to fit between sound and meaning (Nygaard, Herold, & Namy, 2009; Shintel, Nusbaum, & Okrent, 2006), effects that are observable even with limited, or even without any, segmental phonology present (e.g., Perlman, Dale, & Lupyan 2015). The presence of these suprasegmental sound symbolic effects thus demonstrate that cross-modal associations, that cannot be reduced to manner or place features of phonemes, are present and prevalent in speakers' interpretations of speech.

The comparison presented here between phoneme features and individual phoneme effects thus presents an opportunity for gauging different potential explanations for iconicity in language and deciding between alternative mechanisms proposed to account for sound symbolism. The results suggest caution should be taken before concluding that phoneme features, or acoustic features associated with segmental phonology, are driving the effects. It is imperative to at least test whether effects are instead better explained in terms of phonemes. Yet it remains the case that a more comprehensive analysis of speech sounds and meaning attributes, across both written and spoken forms, will be required before the precise mechanism or mechanisms driving relations between sound and meaning can be established.

Note

The data and R script for analyses are available via OSF:

https://osf.io/ujq76/?view_only=85a2b39b6fd342b99955eb7f91268b40

References

- Abramova, E., Fernández, R., & Sangati, F. (2013). Automatic labeling of phonesthemic senses. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1696-1701). Austin, TX: Cognitive Science Society.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.
- Bergen, B.K. (2004). The psychological reality of phonaesthemes. *Language*, *80*, 290-311.
- Blasi, D. B., Wichmann, S., Hammerstom, H., Stadler, P. F., and Christiansen, M. H. (2016). Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences of the United States of America*, *113* (39) 10818-10823.
- Brand, J., Monaghan, P., & Walker, P. (2018). The changing role of sound symbolism for small versus large vocabularies. *Cognitive Science*, *42*, 578-590.
- Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165–172.
- Cassani, G., Chuang, Y., and Baayen R. H. (2019). On the semantics of non-words and their lexical category. *PsyArXiv*, April 3, 1-49. <https://psyarxiv.com/qgsef/>

Cuskley, C. (2013). Mappings between linguistic sound and motion. *Public Journal of Semiotics*, 5(1), 39-62.

Cuskley, C., Simner, J., & Kirby, S. (2017). Phonological and orthographic influences in the bouba–kiki effect. *Psychological Research*, 81, 119-130.

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

D’Onofrio, A. (2014). Phonetic detail and dimensionality in sound-shape correspondences: Refining the *bouba-kiki* paradigm. *Language and Speech*, 57, 367–393.

Fort, M., Martin, A., & Peperkamp, S. (2015). Consonants are more important than vowels in the bouba-kiki effect. *Language and Speech*, 58, 247–266.

French, P. L. (1977). Toward an explanation of phonetic symbolism. *Word*, 28, 305-322.

Hinton, L., Nichols, J., & Ohala, J. L. (1994). Introduction: Sound-symbolic processes. In J. L. Ohala, L. Hilton, & J. Nichols (Eds.), *Sound symbolism* (pp. 1-11). Cambridge, England: Cambridge University Press.

Jespersen, O. (1922). *Language: Its nature, development and origin*. London: Allen and Unwin.

Johnson, S.G. (2018). The NLOpt nonlinear-optimization package, <http://ab-initio.mit.edu/nlopt>. Downloaded 25 June 2018.

- Klink, R. R. (2000). Creating brand names with meaning: The use of sound symbolism. *Marketing Letters*, *11* (1) 5-20.
- Klink, R. R. (2001). Creating meaningful new brand names: A study of semantics and sound symbolism. *Marketing Letters*, *11* (1), 27-34.
- Klink, R. R. (2003). Creating meaningful brands: The relationship between brand name and brand mark. *Marketing Letters*, *14* (3), 143-157.
- Klink, R. R., and Athaide, G. A. (2012). Creating brand personality with brand names. *Marketing Letters*, *23* (1), 109-117.
- Klink, R. R., and Wu, L. (2013). The role of position, type, and combination of sound symbolism imbeds in brand names. *Marketing letters*, *25* (1), 13-24.
- Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.
- Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: A review of behavioural, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology*, *6*, 1246.
- Monaghan, P., Christiansen, M.H., & Lupyan, G. (2014). The systematicity of the sign: Modeling activation of semantic attributes from nonwords. *Proceedings of the 36th Cognitive Science Society Conference*.
- Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in word learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *38*, 1152-1164.
- Monaghan, P., Shillcock, R.C., Christiansen, M.H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B*, 369 20130299.

- Nielsen, A. K., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology*, *65*(2), 115–124.
- Nielsen, A. K., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology*, *67*(2), 153–163.
- Nuckolls, J.B. (1999). The case for sound symbolism. *Annual Review of Anthropology*, *28*, 255-282.
- Ohala, J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In Hinton, L., Nichols, H., & Ohala, J. (Eds.), *Sound symbolism* (pp.325-347). Cambridge, UK: Cambridge University Press.
- Nygaard, L.C., Herold, D.S., & Namy, L.L. (2009). The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science*, *33*, 127-146.
- Otis, K., & Sagi, E. (2008). Phonesthemes: A corpus-based analysis. In V. Sloutsky, B. Love, & K. McRae (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 65-70). Austin, TX: Cognitive Science Society.
- Perniss, P., Thompson, R.L. & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology* *1*, 1– 15.
- Perry, L.K., Perlman, M. & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLoS ONE*, *10*, e0137147.

- Rhodes, R. (1994). Aural images. In Hinton, L., Nichols, J., & Ohala, J. (Eds.), *Sound symbolism* (pp.276-292). Cambridge: Cambridge University Press.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12* (3), 225-239.
- Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, *55*, 167-177.
- Shrum, L. J., Lowrey, T. M., Luna, D., Lerman, D. B., & Liu, M. (2012). Sound symbolism effects across languages: Implications for global brand names. *International Journal of Research in Marketing*, *29*, 275-279.
- Sidhu, D. M., and Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychological Bulletin & Review*, *25*, 1619-1643.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*, 971-995.
- Taylor, I. K., & Taylor, M.M. (1965). Another look at phonetic symbolism. *Psychological Bulletin* *64*, 413-427.
- Ultan, R. (1978). Size-sound symbolism re-examined. *Psychological Bulletin*, *60* (2), 200-209.
- Walker, P. (2016). Cross-sensory correspondences and symbolism in spoken and written language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42* (9). 1339-1361.

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